

# Symplectic Geometry

Spring School, June 7–14, 2004, Utrecht

**J.J. Duistermaat**

Department of Mathematics, Utrecht University, Postbus 80.010, 3508 TA Utrecht, The Netherlands. E-mail: duis@math.uu.nl

## Contents

<b>1</b>	<b>Symplectic Linear Algebra</b>	<b>2</b>
1.1	Symplectic Forms . . . . .	2
1.2	Orthogonal Complements in the Dual Space . . . . .	3
1.3	Orthogonal Complements for a Bilinear Form . . . . .	3
1.4	Isotropic Subspaces . . . . .	4
1.5	Standard Form of the Symplectic Form . . . . .	4
1.6	The Lagrangian Grassmannian . . . . .	5
1.7	The Symplectic Linear Group . . . . .	6
1.8	Exterior Algebra . . . . .	7
1.9	Hermitian Forms . . . . .	9
1.10	Historical Remarks . . . . .	10
1.11	Exercises . . . . .	10
<b>2</b>	<b>Symplectic Manifolds</b>	<b>11</b>
2.1	Definition . . . . .	11
2.2	The Cotangent Bundle . . . . .	12
2.3	Reduction . . . . .	13
2.4	Complex Projective Varieties . . . . .	14
2.5	Almost Complex Structure . . . . .	15
2.6	Cohomology Classes . . . . .	16
2.7	Exercises . . . . .	16
<b>3</b>	<b>Hamiltonian Systems</b>	<b>18</b>
3.1	Flows of Vector Fields . . . . .	18
3.2	Lie Derivatives . . . . .	18
3.3	Hamiltonian Vector Fields . . . . .	21
3.4	The Legendre Transform . . . . .	22
3.5	Poisson Brackets . . . . .	26
3.6	Darboux’s Lemma . . . . .	27
3.7	Hamiltonian Group Actions . . . . .	28
3.8	Poisson Structures . . . . .	30
3.9	Exercises . . . . .	31

<b>4</b>	<b>Hamilton-Jacobi Theory</b>	<b>33</b>
4.1	Lagrange Manifolds . . . . .	33
4.2	Lie's View on First Order PDE . . . . .	34
4.3	An Initial Value Problem . . . . .	35
4.4	Ray Bundles . . . . .	36
4.5	High Frequency Waves and Fourier Integral Operators . . . . .	38
4.6	Some History . . . . .	40
4.7	Exercises . . . . .	40

# 1 Symplectic Linear Algebra

## 1.1 Symplectic Forms

Let  $E$  be a finite-dimensional vector space over a field  $k$ , which later usually will be  $\mathbf{R}$ . A *symplectic form* on  $E$  is a nondegenerate two-form  $\sigma$  on  $E$ . Here the word "two-form" means that  $\sigma$  is an antisymmetric bilinear form on  $E$ . A *bilinear form* on  $E$  is a mapping  $\sigma : E \times E \rightarrow k$  such that, for every choice of  $u \in E$ ,  $v \mapsto \sigma(u, v) : E \rightarrow k$  is a linear form and, for every choice of  $v \in E$ ,  $\sigma(u, v)$  depends linearly on  $u$ . The bilinear form  $\sigma$  is called *antisymmetric* if

$$\sigma(v, u) = -\sigma(u, v), \quad u, v \in E. \tag{1.1}$$

The bilinear form  $\sigma$  is called *nondegenerate* if  $\sigma(u, v) = 0$  for every  $v \in E$  implies that  $u = 0$ .

As usual, we identify a bilinear form  $\sigma$  on  $E$  with the linear mapping  $u \mapsto (v \mapsto \sigma(u, v))$  from  $E$  to the dual space  $E^*$  of  $E$ , this linear map will also be denoted by  $\sigma$ . That is, we write

$$(\sigma(u))(v) = \sigma(u, v), \quad u, v \in E.$$

The mapping which assigns to  $v \in E$  the linear form  $\alpha \mapsto \alpha(v)$  on  $E^*$  induces a linear isomorphism from  $E$  onto  $(E^*)^*$ , which is used to identify  $(E^*)^*$  with  $E$ . Then the dual (= transposed) mapping  $\sigma^*$  of the linear map  $\sigma : E \rightarrow E^*$  is a linear mapping from  $(E^*)^* = E$  to  $E^*$  and the antisymmetry of  $\sigma$  is equivalent to the condition that  $\sigma^* = -\sigma$ .

The nondegeneracy of  $\sigma$  means that the linear mapping  $\sigma : E \rightarrow E^*$  has zero kernel (= null space), and because  $\dim E^* = \dim E$ , this is equivalent to the condition that the linear mapping  $\sigma : E \rightarrow E^*$  is bijective.

More generally, any linear mapping from  $E$  to  $E^*$  corresponds in the above fashion to a unique bilinear form on  $E$ , and the linear mapping  $E \rightarrow E^*$  is bijective (= an isomorphism) if and only if the bilinear form is nondegenerate. In the case that the bilinear form is an inner product, i.e. symmetric and positive definite, then it is nondegenerate and we obtain the usual identification of  $E$  with  $E^*$  by means of the inner product. In this way we may think of a symplectic form as an antisymmetric analogue of an inner product.

**Example 1.1** On  $k^{2n} = k^n \times k^n$  we define  $\sigma$  by

$$\sigma((q, p), (q', p')) = \sum_{j=1}^n p_j q'_j - p'_j q_j. \tag{1.2}$$

It is easy to verify that  $\sigma$  is a nondegenerate antisymmetric bilinear form on  $k^n \times k^n$ . It is called the *standard symplectic form* on  $k^n \times k^n$ . A coordinate free version is the symplectic form on  $E = F \times F^*$  defined by

$$\sigma((x, \xi), (y, \eta)) = \xi(y) - \eta(x), \quad x, y \in F, \quad \xi, \eta \in F^*,$$

defined for any finite-dimensional vector space  $F$  over  $k$ . ⊗

## 1.2 Orthogonal Complements in the Dual Space

If  $L$  is a linear subspace of  $E$ , then the *orthogonal complement* or *annihilator*  $L^0$  of  $L$  in  $E^*$  is defined as the set of all  $\alpha \in E^*$  such that  $\alpha(v) = 0$  for every  $v \in L$ . Clearly  $L^0$  is a linear subspace of  $E^*$  and  $\dim L^0 = \dim E^* - \dim L = \dim E - \dim L =$  the codimension of  $L$  in  $E$ .

Similarly the orthogonal complement or annihilator  $A^0$  in  $E$  of a linear subspace  $A$  of  $E^*$  is defined as the common kernel of all the linear forms  $\alpha \in A$ , i.e. the set of all  $v \in E$  such that  $\alpha(v) = 0$  for every  $\alpha \in A$ .  $A^0$  is a linear subspace of  $E$  and  $\dim A^0 = \dim E - \dim A$ .

If  $L$  is a linear subspace of  $E$ , then obviously  $L \subset (L^0)^0$ , and because  $\dim(L^0)^0 = \dim E - \dim L^0 = \dim E - (\dim E - \dim L) = \dim L$ , we conclude that  $L = (L^0)^0$ . Similarly  $A = (A^0)^0$  for any linear subspace  $A$  of  $E^*$ .

If  $L$  and  $M$  are linear subspaces of  $E$ , then obviously  $L \subset M$  implies  $M^0 \subset L^0$ . Therefore in general  $L^0 \subset (L \cap M)^0$  and  $M^0 \subset (L \cap M)^0$ , which implies that  $L^0 + M^0 \subset (L \cap M)^0$ , and similarly  $(L + M)^0 \subset L^0 \cap M^0$ . Taking orthogonal complements these inclusions imply that

$$L \cap M = ((L \cap M)^0)^0 \subset (L^0 + M^0)^0 \subset (L^0)^0 \cap (M^0)^0 = L \cap M,$$

It follows that both inclusions are equalities and therefore  $(L \cap M)^0 = L^0 + M^0$ . Similarly we have  $(L + M)^0 = L^0 \cap M^0$ .

## 1.3 Orthogonal Complements for a Bilinear Form

Let  $\sigma$  be a nondegenerate bilinear form on  $E$ , not necessarily antisymmetric. If  $L$  is a linear subspace of  $E$ , then the  $\sigma$ -orthogonal complement  $L^\sigma$  of  $L$  in  $E$  is defined as

$$L^\sigma := \{u \in E \mid \sigma(u, v) = 0 \text{ for every } v \in L\}. \quad (1.3)$$

Clearly  $L^\sigma$  is a linear subspace of  $E$  and  $L^\sigma = \sigma^{-1}(L^0) = (\sigma^*(L))^0 =$ . Note that  $L^\sigma = L^{\sigma^*}$  if  $\sigma$  is symmetric or antisymmetric. That is, in these case we can interchange the role of  $u$  and  $v$  in the definition (1.3).

Because  $\sigma : E \rightarrow E^*$  is a linear isomorphism, the rules for annihilators in the dual spaces imply the rules

$$\dim L^\sigma = \dim E - \dim L, \quad (1.4)$$

$$L \subset M \implies M^\sigma \subset L^\sigma, \quad (1.5)$$

$$(L \cap M)^\sigma = L^\sigma + M^\sigma \quad \text{and} \quad (L + M)^\sigma = L^\sigma \cap M^\sigma \quad (1.6)$$

for the  $\sigma$ -orthogonal complements. If  $k = \mathbf{R}$  and  $\sigma$  is an inner product, then the  $\sigma$ -orthogonal complement of  $L$  is usual orthogonal complement denoted by  $L^\perp$ , and we recognize (1.4), (1.5) and (1.6) as familiar properties of the orthogonal complementation. Note that for the inner product we have the additional property that  $L \cap L^\perp = \{0\}$ , which implies that  $E$  is equal to the direct sum  $L \oplus L^\perp$  of  $L$  and  $L^\perp$ .

## 1.4 Isotropic Subspaces

In the sequel  $(E, \sigma)$  will be a *symplectic vector space*, i.e.  $E$  is a finite-dimensional vector space and  $\sigma$  is a symplectic form on  $E$ .

A linear subspace  $L$  is called *isotropic with respect to  $\sigma$*  if  $L \subset L^\sigma$ , that is  $\sigma(u, v) = 0$  for all pairs of vectors  $u, v \in L$ . A maximal isotropic linear subspace of  $E$  is called a *Lagrange plane*. Because of the finite-dimensionality of  $E$ , any strictly increasing sequence of isotropic subspaces terminates at a maximal one, which shows that every isotropic subspace is contained in at least one Lagrange plane.

The antisymmetry of  $\sigma$  implies that  $\sigma(v, v) = -\sigma(v, v)$ , hence  $2\sigma(v, v) = 0$ . Therefore, if the characteristic of  $k$  is not equal to two, the antisymmetry (1.1) implies that

$$\sigma(v, v) = 0, \quad v \in E. \quad (1.7)$$

Conversely, if (1.7) holds, then  $0 = \sigma(u + v, u + v) = \sigma(u, u) + \sigma(u, v) + \sigma(v, u) + \sigma(v, v) = \sigma(u, v) + \sigma(v, u)$ , which implies (1.1). Therefore, if  $\text{char } k \neq 2$ , such as for  $k = \mathbf{R}$ , then (1.1) is equivalent to (1.7). If  $\text{char } k = 2$ , then everything what follows remains true if we replace the antisymmetry condition (1.1) by the stronger condition (1.7).

The condition (1.7) implies that *every one-dimensional linear subspace of  $E$  is isotropic*. This is very different from the situation for an inner product, where  $\{0\}$  is the only isotropic subspace.

If  $L \subset L^\sigma$  and  $L \neq L^\sigma$ , then for every  $v \in L^\sigma \setminus L$  we have that  $(L + kv)^\sigma = L^\sigma \cap (kv)^\sigma$  contains  $L$  because  $L^\sigma \supset L$  and  $kv \subset L^\sigma$  implies  $(kv)^\sigma \supset L$ . It also contains  $v$  because  $v \in L^\sigma$  and  $v \in (kv)^\sigma$ . It follows that the linear subspace  $(L + kv)^\sigma$  contains  $L + kv$ , which means that  $L' := L + kv$  is isotropic,  $L \subset L'$  and  $\dim L' = \dim L + 1$ . We conclude that  $L$  is a maximal isotropic linear subspace if and only if  $L = L^\sigma$ . This is equivalent to the condition that  $L$  is isotropic and  $\dim L = \dim L^\sigma = \dim E - \dim L$ , or equivalently  $\dim E = 2 \dim L$ .

In particular the dimension of a symplectic vector space must be even, say equal to  $2n$ . We have that  $\dim L \leq n$  for every isotropic linear subspace  $L$  of  $E$  and that the maximal isotropic subspaces are the isotropic subspaces which have dimension equal to  $n$ .

## 1.5 Standard Form of the Symplectic Form

In order to identify the symplectic form with the standard one of Example 1.1, we start with an  $L \in \mathcal{L}(E, \sigma)$  and introduce a second  $M \in \mathcal{L}(E, \sigma)$  such that  $L \cap M = \{0\}$ , which then implies that  $E = L \oplus M$ . The existence of such a Lagrange plane  $M$  follows from the fact that, if  $M$  is an isotropic subspace of  $(E, \sigma)$  such that  $M \neq M^\sigma$  and  $L \cap M = \{0\}$ , then there exists a  $v \in M^\sigma \setminus M$  such that  $L \cap M' = \{0\}$  if  $M' := M + kv$ . If  $L \cap M' \neq \{0\}$  then there exist  $m \in M$  and  $c \in K$  such that  $l = m + cv$  is a nonzero element of  $L$ , which means that  $v \in L + M$ . If this would hold for every  $v \in M^\sigma \setminus M$ , then  $M^\sigma \subset L + M$ , which implies that  $L \cap M^\sigma = L^\sigma \cap M^\sigma = (L + M)^\sigma \subset (M^\sigma)^\sigma = M$ , which in view of  $L \cap M = \{0\}$  means that  $L \cap M^\sigma = \{0\}$ . However,  $\dim L = n$  and  $\dim M^\sigma = \dim E - \dim M > \dim E - n$ , hence  $\dim L + \dim M^\sigma > \dim E$ , which implies that  $\dim(L \cap M^\sigma) = \dim L + \dim M^\sigma - \dim E > 0$ , and we arrive at a contradiction.

The restriction  $S$  to  $M$  of the mapping  $m \mapsto (\sigma m)|_L$  is a linear isomorphism from  $M$  onto  $L^*$ . Indeed,  $S m = 0$  means that  $m \in L^\sigma = L$ , which in view of  $L \cap M = \{0\}$  implies that  $m = 0$ . Now let  $e_i$  be a basis of  $L$  and let  $\epsilon_j$  be the corresponding dual basis of  $L^*$ , determined by the conditions that  $\epsilon_j(e_i)$  is equal to zero and one when  $i \neq j$  and  $i = j$ , respectively. Let  $f_j$  be the basis of  $M$  such that  $S f_j = \epsilon_j$ . Then we have  $\sigma(e_i, e_{i'}) = 0$ ,  $\sigma(f_j, f_{j'}) = 0$  and  $-\sigma(e_i, f_j) = \sigma(f_j, e_i) = \epsilon_j(e_i) = \delta_{ij}$ .

The  $e_i$  and  $f_j$  together form a basis on which  $\sigma$  has a standard form. Such a basis is called a *symplectic basis of  $(E, \sigma)$* .

More precisely, if we write  $u = \sum_{i=1}^n p_i e_i + q_j f_j$ ,  $v = \sum_{i=1}^n p'_i e_i + q'_j f_j$ , then it follows that  $\sigma(u, v)$  is equal to the right hand side of (1.2). This means that the pull-back of  $\sigma$  under the linear isomorphism  $(p, q) \mapsto \sum_{i=1}^n p_i e_i + q_j f_j$  from  $k^n \times k^n$  onto  $E$  is equal to the standard symplectic form on  $k^n \times k^n$ .

For an arbitrary antisymmetric bilinear form  $\sigma$  on a finite-dimensional vector space  $E$  over a field  $k$ , a basis on which  $\sigma$  has a standard form is obtained as follows. If  $N = \ker \sigma$ , then the equation  $\sigma_{E/N}(u + N, v + N) = \sigma(u, v)$  leads to a well-defined antisymmetric bilinear form on  $E/N$  which is nondegenerate, this is called the *induced symplectic form on  $E/N$* . Write  $r = \dim N$ ,  $m = \frac{1}{2} \dim(E/N)$ , and choose  $n_1, \dots, n_r, e_1, \dots, e_m, f_1, \dots, f_m$  in  $E$  such that the  $n_1, \dots, n_r$  form a basis of  $N$  and the  $e_1 + N, \dots, e_m + N, f_1 + N, \dots, f_m + N$  form a symplectic basis of  $(E/N, \sigma_{E/N})$ . Then these  $r + 2m$  vectors form a basis of  $E$ , for which  $\sigma(n_i, n_j) = \sigma(n_i, e_j) = \sigma(n_i, f_j) = 0$ ,  $\sigma(e_i, e_j) = \sigma(f_i, f_j) = 0$ , and  $\sigma(e_i, f_j) = -\sigma(f_j, e_i) = \delta_{ij}$ .

The fact that antisymmetric bilinear forms on vector spaces of the same dimension and with the same nullity (= dimension of the kernel) have the same normal form, differs from the situation for symmetric bilinear forms. If  $k$  has the property that every element of  $k$  has a square root in  $k$ , then all symmetric bilinear forms with the same nullity have the same normal form, but for general fields the classification of symmetric bilinear forms can be very complicated. For  $k = \mathbf{R}$  the situation still is relatively simple, as a real symmetric bilinear form is determined by its nullity  $n_0$ , and its positive and negativity index = the dimension  $n_+$  and  $n_-$  of any maximal linear subspace on which the form is positive definite and negative definite, respectively.

## 1.6 The Lagrangian Grassmannian

The set  $\mathcal{L} = \mathcal{L}(E, \sigma)$  of all Lagrange planes in the symplectic vector space  $(E, \sigma)$  is called the *Lagrangian Grassmannian* of  $(E, \sigma)$ . It is a non-empty algebraic subvariety of the Grassmann manifold  $G_n(E)$  of all  $n$ -dimensional linear subspaces of  $E$ .

If  $L \in \mathcal{L}$  and  $k \in \mathbf{Z}_{\geq 0}$ , then we denote by  $\mathcal{L}_{L,k}$  the set of all  $M \in \mathcal{L}$  such that  $\dim L \cap M = k$ .  $\mathcal{L}_{L,0}$  is an open subset of  $\mathcal{L}$ , and in the previous subsection we have seen that it is not empty. Interchanging the roles of  $L$  and  $M$  we obtain that the  $\mathcal{L}_{L,0}$  for  $L \in \mathcal{L}$  form an open covering of  $\mathcal{L}$ , i.e. for every  $M \in \mathcal{L}$  there exists an  $L \in \mathcal{L}$  such that  $M \in \mathcal{L}_{L,0}$ .

Let  $L, M \in G_n(E)$  be such that  $E = L \oplus M$ . For every  $L' \in G_n(E)$  such that  $L' \cap M = \{0\}$  there is a unique linear mapping  $A : L \rightarrow M$  such that  $L' = \{x + Ax \mid x \in L\}$ . This leads to a bijective mapping from the open subset  $\{L' \in G_n(E) \mid L' \cap M = \{0\}\}$  of  $G_n(E)$  onto the  $n^2$ -dimensional vector space  $\text{Lin}(L, M)$  of all linear mappings from  $L$  onto  $M$ , and the matrix coefficients of  $A$  with respect to any bases in  $L$  and  $M$  define a coordinatization of the aforementioned open subset of  $G_n(E)$ . These are the standard coordinatizations of  $G_n(E)$ . The coordinate changes are rational mappings and in this way  $G_n(E)$  is exhibited as a smooth  $n^2$ -dimensional manifold.

Now assume that  $L, M \in \mathcal{L}(E, \sigma)$ . We have that  $L' \in \mathcal{L}(E, \sigma)$  if and only if, for every  $x, y \in L$ ,

$$0 = \sigma(x + Ax, y + Ay) = \sigma(Ax, y) + \sigma(x, Ay) = \sigma(Ax, y) - \sigma(Ay, x),$$

where we have used in the second identity that  $\sigma(x, y) = 0$  and  $\sigma(Ax, Ay) = 0$  because  $L$  and  $M$  are isotropic. This means that the bilinear form  $(x, y) \mapsto \sigma(Ax, y)$  on  $L$  is symmetric. Because  $A \mapsto ((x, y) \mapsto \sigma(Ax, y))$  is a linear isomorphism from  $\text{Lin}(L, M)$  onto the space of all bilinear

forms on  $L$ , we see that in the aforementioned coordinatization the Lagrangian Grassmannian appears as the space of all symmetric bilinear forms on  $L$ , viewed as a linear subspace of the space of all bilinear forms on  $L$ . This exhibits  $\mathcal{L}(E, \sigma)$  as a smooth  $\frac{1}{2}n(n+1)$ -dimensional linear submanifold of the  $n^2$ -dimensional manifold  $G_n(E)$ .

If  $\widetilde{M}$  is another Lagrange plane which is transversal to  $L$  and  $L' \in \mathcal{L}$  is transversal to both  $M$  and  $\widetilde{M}$ , then we have a second coordinatization of  $L'$  by means of the  $\widetilde{A} \in \text{Lin}(L, \widetilde{M})$  such that  $L' = \{y + \widetilde{A}y \mid y \in L\}$ . On the other hand there exists  $B \in \text{Lin}(M, L)$  such that  $\widetilde{M} = \{z + Bz \mid z \in M\}$ . The elements of  $L'$  are of the form  $x + Ax = y + \widetilde{A}y$  for unique  $x, y \in L$ , and  $\widetilde{A}y = z + Bz$  for a unique  $z \in M$ . From  $x + Ax = y + z + Bz$ , with  $x, y, Bz \in L$  and  $Ax, z \in M$  it follows that  $z = Ax, y = x - Bz = x - BAx$ , hence  $x + Ax = (x - BAx) + \widetilde{A}(x - BAx)$ . Taking the symplectic product with  $u \in L$ , this leads to

$$\sigma(Ax, u) = \sigma(\widetilde{A}(x - BAx), u) = \sigma(\widetilde{A}x, u) - \sigma(\widetilde{A}BAx, u).$$

For small  $A$ , which corresponds to  $L'$  close to  $L$ , we see that the symmetric bilinear form on  $L$  defined by the  $M$  differs from the symmetric bilinear form on  $L$  defined by  $\widetilde{M}$  by a term which vanishes of second order in  $A$ . In this way the tangent space  $T_L \mathcal{L}$  of  $\mathcal{L}$  at the element  $L \in \mathcal{L}$  is canonically identified with the space  $\text{Symm}^2(L)$  of all symmetric bilinear forms on  $L$ .

## 1.7 The Symplectic Linear Group

Let  $(E, \sigma), (F, \tau)$  be a symplectic vector spaces over a field  $k$ . A linear mapping  $A : E \rightarrow F$  is called a *symplectic linear mapping from  $(E, \sigma)$  to  $(F, \tau)$* , if  $\tau(Au, Av) = \sigma(u, v)$  for all  $u, v \in E$ , or  $A^* \tau A = \sigma$ . Because  $\sigma : E \rightarrow E^*$  is injective,  $A$  is injective, hence  $\dim E \leq \dim F$ , and we have that  $A$  is bijective if and only  $\dim E = \dim F$ , in which case  $A$  is called a *symplectic linear isomorphism from  $(E, \sigma)$  onto  $(F, \tau)$* . If  $\dim E < \dim F$ , then  $A$  is a symplectic linear isomorphism from  $(E, \sigma)$  onto the symplectic vector subspace  $(A(E), \tau|_{A(E) \times A(E)})$  of  $(F, \tau)$ .

A symplectic linear mapping from  $(E, \sigma)$  to itself is called a *symplectic linear transformation* in  $(E, \sigma)$ . The symplectic linear transformations form an algebraic subgroup of the group  $\text{GL}(E)$  of all linear transformations in  $E$ , which is called the *symplectic linear group*  $\text{Sp}(E, \sigma)$ .

If  $e_1, \dots, e_n, f_1, \dots, f_n$  is a symplectic basis, then a linear mapping  $A$  is symplectic if and only if  $Ae_1, \dots, Ae_n, Af_1, \dots, Af_n$  is a symplectic basis. Because any basis of a Lagrange plane can be extended to a symplectic basis, it follows that the symplectic linear group  $\text{Sp}(E, \sigma)$  acts transitively on the Lagrangian Grassmannian  $\text{Sp}(E, \sigma)$ . If, for given  $L \in \mathcal{L}(E, \sigma)$ ,  $\text{Sp}(E, \sigma)_L := \{A \in \text{Sp}(E, \sigma) \mid A(L) = L\}$  denotes the stabilizer subgroup of  $L$  in  $\text{Sp}(E, \sigma)$ , then the Lagrangian Grassmannina  $\mathcal{L}(E, \sigma)$  is identified with the homogeneous space  $\text{Sp}(E, \sigma) / \text{Sp}(E, \sigma)_L$ .

If we write  $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$  on a symplectic basis, in which  $\alpha, \beta, \gamma, \delta$  are  $n \times n$ -matrices, then  $A \in \text{Sp}(E, \sigma)$  if and only if  $\alpha^* \gamma - \gamma^* \alpha = 0, \beta^* \delta - \delta^* \beta = 0$  and  $\alpha^* \delta - \beta^* \gamma = I$ . The first two equations mean that  $\alpha^* \gamma = \epsilon$  and  $\beta^* \delta = \eta$  are symmetric, and we see  $2(\frac{1}{2}n(n-1)) + n^2 = 2n^2 - n$  independent equations. If the first  $n$  vectors of the symplectic basis span  $L$ , then  $A \in \text{Sp}(E, \sigma)_L$  if and only if  $\gamma = 0$ , and the equations are that  $\delta = (\alpha^*)^{-1}$  and  $\delta^* \beta = \alpha^{-1} \beta$  is symmetric. It follows that  $\dim \text{Sp}(E, \sigma)_L = n^2 + \frac{1}{2}n(n+1)$ , and therefore  $\dim \text{Sp}(E, \sigma) = n^2 + n(n+1) = 2n^2 + n$ . It follows that the codimension of  $\text{Sp}(n, E)$  in  $\text{GL}(E)$  is equal to  $2n^2 - n$ , in agreement with the number of independent equations for the matrices  $\alpha, \beta, \gamma, \delta$ .

The equation  $A^* \sigma A = \sigma$  for  $A \in \text{Sp}(E, \sigma)$  implies that  $A^* = \sigma A \sigma^{-1}$ , which means that the linear isomorphism  $\sigma : E \rightarrow E^*$  conjugates the linear mapping  $A : E \rightarrow E$  with the linear mapping

$(A^*)^{-1} : E^* \rightarrow E^*$ . This implies that  $A$  has the same eigenvalues as  $(A^*)^{-1}$  with the same algebraic multiplicities. Because  $A^*$  has the same eigenvalues with the same algebraic multiplicities as  $A$ , it follows that if  $\lambda$  is an eigenvalue of  $A$  with algebraic multiplicity  $m$ , then  $1/\lambda$  is also an eigenvalue of  $A$  with algebraic multiplicity  $m$ . If  $k = \mathbf{R}$ , then a passage to the complexification leads to the conclusion that the complex eigenvalues of  $A$  appear in foursomes  $\lambda, \bar{\lambda}, 1/\lambda, 1/\bar{\lambda}$  with  $\lambda \in \mathbf{C}$ ,  $\lambda \notin \mathbf{R}$ ,  $\lambda \bar{\lambda} \neq 1$ , or in real pairs  $\lambda, 1/\lambda$ , or in complex conjugate pairs on the unit circle, in each case with equal algebraic multiplicities.

The Lie algebra  $\mathfrak{sp}(E, \sigma)$  of  $\mathrm{Sp}(E, \sigma)$  consists of the linear mappings  $A : E \rightarrow E$  such that  $\sigma(Au, v) + \sigma(u, Av) = 0$  for all  $u, v \in E$ , or  $\sigma A + A^* \sigma = 0$ . These  $A$  are called the *infinitesimal symplectic linear transformations* in  $(E, \sigma)$ . Note also that the mapping which assigns to  $A \in \mathfrak{sp}(E, \sigma)$  the bilinear form  $\sigma A : (u, v) \mapsto \sigma(Au, v)$  is a linear isomorphism from  $\mathfrak{sp}(E, \sigma)$  onto the space  $\mathrm{Symm}^2(E)$  of all symmetric bilinear forms on  $E$ . This shows that  $\dim \mathrm{Sp}(E, \sigma) = \frac{1}{2} 2n(2n+1) = 2n^2 + n$ , in agreement with our previous dimension calculations.

The equation  $A^* = -\sigma A \sigma^{-1}$  implies that if  $\lambda$  is an eigenvalue of  $A$  with multiplicity  $m$ , then  $-\lambda$  is also an eigenvalue of  $A$  with multiplicity  $m$ . When  $k = \mathbf{R}$ , this implies that the complex eigenvalues of  $A$  appear in foursomes  $\lambda, \bar{\lambda}, -\lambda, -\bar{\lambda}$  with  $\lambda \in \mathbf{C}$ ,  $\lambda \notin \mathbf{R}$ ,  $\lambda \notin i\mathbf{R}$ , or in real pairs  $\lambda, -\lambda$ , or in complex conjugate pairs on the imaginary axis, in each case with equal algebraic multiplicities.

For  $k = \mathbf{R}$  a complete list of normal forms of infinitesimal symplectic linear transformations has been given by Williamson [30], and one has a corresponding list for the symplectic linear transformations, see also [2] and [4].

## 1.8 Exterior Algebra

Let  $E$  be any  $d$ -dimensional vector space over  $k$ . The space of all antisymmetric  $p$ -linear forms on  $E$  is denoted by  $\Lambda^p E^*$ . For  $\alpha \in \Lambda^p E^*$  and  $\beta \in \Lambda^q E^*$ , one defines the exterior product  $\alpha \wedge \beta \in \Lambda^{p+q} E^*$  by

$$(\alpha \wedge \beta)(v_1, \dots, v_{p+q}) = \sum_{\pi} \mathrm{sgn} \pi \alpha(v_{\pi(1)}, \dots, v_{\pi(p)}) \beta(v_{\pi(p+1)}, \dots, v_{\pi(p+q)}) \quad (1.8)$$

for all  $v_1, \dots, v_{p+q} \in E$ . Here the sum is over all equivalence classes of permutations  $\pi$  of the indices  $\{1, \dots, p+q\}$ , where  $\pi$  and  $\pi \circ \psi$  are equivalent if  $\psi$  maps the subsets  $\{1, \dots, p\}$  and  $\{p+1, \dots, p+q\}$  of indices to themselves. (Note that terms with equivalent permutations are equal.) The signature  $\mathrm{sgn} \pi$  of the permutation is  $+1$  or  $-1$  when  $\pi$  consists of an even or odd number of transpositions, respectively.

With this product,

$$\Lambda E^* := \bigoplus_{p \geq 0} \Lambda^p E^*$$

becomes an algebra, which is called the *exterior algebra* of  $E^*$ . Here we have used the convention that  $\Lambda^0 E^* = k$ . Note also that  $\Lambda^1 E^* = E^*$  is a linear subspace of  $\Lambda E^*$ .

The exterior product is *associative*, meaning that  $(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma)$ . It is *anticommutative* in the sense that  $\alpha \wedge \beta = (-1)^{pq} \beta \wedge \alpha$  if  $\alpha \in \Lambda^p E^*$  and  $\beta \in \Lambda^q E^*$ .

Let  $e_i$  be any basis of  $E$  and  $\epsilon_j$  de corresponding dual basis of  $E^*$ , characterized by  $\epsilon_j(e_i) = \delta_{ij}$ . For any strictly increasing function  $I : \{1, \dots, p\} \rightarrow \{1, \dots, d\}$ , write  $\epsilon_I = \epsilon_{I(1)} \wedge \dots \wedge \epsilon_{I(p)}$  and  $e_I = (e_{I(1)}, \dots, e_{I(p)})$ . Then  $\epsilon_I(e_J) = \delta_{IJ}$ . Therefore, if  $\alpha \in \Lambda^p E^*$ , then  $\alpha = \sum_I \alpha(e_I) \epsilon_I$ , which follows from applying both sides to  $e_J$  and observing that the numbers  $\alpha(e_J)$  determine  $\alpha$ .

Conversely, if  $\sum_I c_I \epsilon_I = 0$ , then application to  $e_J$  yields  $c_J = 0$  for every  $J$ , and it follows that the  $\epsilon_I$  form a basis of  $\Lambda^p E^*$ , which in turn implies that

$$\dim \Lambda^p E^* = \binom{d}{p}. \quad (1.9)$$

In particular  $\Lambda^p E^* = \{0\}$  when  $p > n$  and the space  $\Lambda^d E^*$  of *oriented volume forms* on  $E$  is one-dimensional.

If  $\sigma \in \Lambda^2 E^*$  is a symplectic form on  $E$  and  $(e_1, \dots, e_n, f_1, \dots, f_n)$  is a symplectic basis, with corresponding dual basis  $(\epsilon_1, \dots, \epsilon_n, \phi_1, \dots, \phi_n)$ , then we see from (1.2) that

$$\sigma = \sum_{i=1}^n \epsilon_i \wedge \phi_i \quad (1.10)$$

and therefore the  $n$ -th exterior power

$$\sigma^n = \sigma \wedge \dots \wedge \sigma = n! \epsilon_1 \wedge \phi_1 \wedge \epsilon_2 \wedge \phi_2 \wedge \dots \wedge \epsilon_n \wedge \phi_n \quad (1.11)$$

of  $\sigma$  is a nonzero volume form on  $E$ . (Note that the exterior product is commutative on the subalgebra generated by the two-forms  $\epsilon_i \wedge \phi_i$ .) Here we assume that the field  $k$  has characteristic zero or, in the case of nonzero characteristic,  $\text{char } k > n$ . This implies that also all the intermediate powers  $\sigma^m \in \Lambda^{2m} E^*$ ,  $0 \leq m \leq n$ , are nonzero.

With the identification of  $E$  with  $(E^*)^*$ , the exterior algebra  $\Lambda E$  of  $E$  is defined as the algebra of antisymmetric multilinear forms on  $E^*$ . There is a natural identification of  $\Lambda E$  with the dual space of  $\Lambda E^*$  and vice versa, as follows.

If  $v \in (\Lambda^p E^*)^*$ , then

$$i(v)(\alpha_1, \dots, \alpha_p) := v(\alpha_1 \wedge \dots \wedge \alpha_p), \quad \alpha_i \in E^*,$$

defines an antisymmetric  $p$ -linear form on  $E^*$ , and therefore belongs to  $\Lambda^p E$ . This defines a linear mapping  $i : (\Lambda^p E^*)^* \rightarrow \Lambda^p E$ . Furthermore, if  $i(v) = 0$ , then  $v$  annihilates all  $p$ -fold exterior products of one-forms and therefore  $v = 0$  because the  $\epsilon_I$  form a basis of  $\Lambda^p E^*$ . We conclude that  $i$  is injective and because

$$\dim(\Lambda^p E^*)^* = \binom{d}{p} = \dim \Lambda^p E,$$

we conclude that  $i$  is a linear isomorphism.

We have a similar linear isomorphism  $j : (\Lambda^p E)^* \rightarrow \Lambda^p E^*$ . If  $v_1 \wedge \dots \wedge v_p = i(v)$  and  $\alpha_1 \wedge \dots \wedge \alpha_p = j(\alpha)$ , then

$$\begin{aligned} \alpha(i(v)) &= j(\alpha)(v_1, \dots, v_p) = (\alpha_1 \wedge \dots \wedge \alpha_p)(v_1, \dots, v_p) = \sum_{\pi} \text{sgn } \pi \prod_{k=1}^p \alpha_k(v_{\pi(k)}) \\ &= (v_1 \wedge \dots \wedge v_p)(\alpha_1, \dots, \alpha_p) = i(v)(\alpha_1, \dots, \alpha_p) = v(j(\alpha)), \end{aligned}$$

which shows that the mappings  $i$  and  $j$  are each others adjoints.

One uses  $i$  and  $j$  to identify  $(\Lambda^p E^*)^*$  with  $\Lambda^p E$  and  $(\Lambda^p E)^*$  with  $\Lambda^p E^*$ . With these identifications, one has the formulas  $v(\alpha_1 \wedge \dots \wedge \alpha_p) = v(\alpha_1, \dots, \alpha_p)$  for  $v \in (\Lambda^p E^*)^* = \Lambda^p E$  and  $\alpha_i \in E^*$ , and similarly  $\alpha(v_1 \wedge \dots \wedge v_p) = \alpha(v_1, \dots, v_p)$  for  $\alpha \in (\Lambda^p E)^* = \Lambda^p E^*$  and  $v_i \in E$ .



## 1.9 Hermitian Forms

Let  $E$  be an  $n$ -dimensional vector space over  $\mathbf{C}$  and let  $h : E \times E \rightarrow \mathbf{C}$  be a Hermitian form on  $E$ , i.e. for every  $u \in E$  the mapping  $v \mapsto h(u, v)$  is complex antilinear ( $h(u, cv) = \bar{c}h(u, v)$ ), for every  $v \in E$  the mapping  $u \mapsto h(u, v)$  is complex linear, and  $h(v, v) > 0$  for every nonzero element  $v$  of  $E$ .

From now on we regard  $E$  as a  $2n$ -dimensional vector space over  $\mathbf{R}$ . Then the real part  $g := \operatorname{Re} h$  of  $h$  is an inner product on  $E$ , and therefore a nondegenerate symmetric bilinear form. Because

$$\operatorname{Im} h(u, v) = -\operatorname{Re}(ih(u, v)) = -\operatorname{Re} h(iu, v) = -\operatorname{Re} h(v, iu) = \operatorname{Re}(ih(v, u)) = -\operatorname{Im} h(v, u),$$

we see that the imaginary part  $\sigma := \operatorname{Im} h$  is an antisymmetric bilinear form on  $E$ . These equations also show that  $\sigma = -g \circ J$ , if  $J : E \rightarrow E$  is the real linear transformation in  $E$  defined by  $J(u) = iu$ ,  $u \in E$ , the complex multiplication by means of the complex number  $i$ . Because  $J : E \rightarrow E$  and  $g : E \rightarrow E^*$  are injective,  $\sigma : E \rightarrow E^*$  is injective as well, and we conclude that  $\sigma$  is a symplectic form on  $E$ . Note that  $J^2 = -1$  and therefore  $g$  can also be expressed in terms of  $\sigma$  by means of the formula  $g = \sigma \circ J$ , and the Hermitian form is equal to  $h = \sigma \circ J + i\sigma$ .

In general, if  $E$  is a vector space over  $\mathbf{R}$ , then a *complex structure* in  $E$  is defined as a real linear mapping  $J : E \rightarrow E$  such that  $J^2 = -1$ . This makes  $E$  into a complex vector space if we define  $(a + ib)v = av + J(bv)$  for any  $a, b \in \mathbf{R}$  and  $v \in E$ , and it follows that the real dimension of  $E$  is equal to  $2n$  if  $n$  denotes the dimension of  $E$  as a complex vector space.

That every symplectic form is equal to the imaginary part of a Hermitian form with respect to a suitable complex structure, can be seen by bringing the symplectic form into the standard form (1.2), writing  $z_j = q_j + ip_j$  and taking the standard Hermitian form

$$h(z, z') = \sum_{j=1}^n z_j \bar{z}'_j \tag{1.12}$$

in  $\mathbf{C}^n$ .

The vectors  $e_1, \dots, e_n \in E$  form an  $h$ -orthonormal basis of the complex vector space  $E$ , if and only if they form a  $g$ -orthonormal basis of a Lagrange plane  $L$ . Let  $U(E, h)$  denote the *unitary group* of all complex linear transformations  $A : E \rightarrow E$  such that  $A^*(h) = h$ , in which the Hermitian form  $A^*h$  on  $E$  is defined by  $(A^*h)(u, v) = h(Au, Av)$  for all  $u, v \in E$ . Note that

$$U(E, h) = \operatorname{GL}_{\mathbf{C}}(E) \cap O(E, g) = \operatorname{GL}_{\mathbf{C}}(E) \cap \operatorname{Sp}(E, \sigma) = \operatorname{Sp}(E, \sigma) \cap O(E, g),$$

in which  $\operatorname{GL}_{\mathbf{C}}(E)$  denotes the group of complex linear transformations on  $E$  and  $O(E, g)$  denotes the group of  $g$ -orthogonal real linear transformations in  $E$ . This is based on the fact that a real linear transformation  $A$  in  $E$  is complex linear if and only if  $A \circ J = J \circ A$ , and  $g = \sigma \circ J$  and  $h = g + i\sigma$ .

Because  $U(E, h)$  acts transitively on the set of all  $h$ -orthonormal bases of  $E$ , the compact Lie group  $U(E, h)$  also acts transitively on the Lagrangian Grassmannian. Because for any  $A \in U(E, h)$  we have that  $A(L) = L$  if and only if  $A$  is the complex linear extension to  $E$  of a  $g$ -orthogonal transformation in  $L$ , this leads to an identification of  $\mathcal{L}(E, \sigma)$  with the homogeneous space  $U(E, h)/U(E, h)_L$ , in which  $U(E, h)_L$  is isomorphic to  $O(L, g)$ . This is the meaning of the equation  $\Lambda(n) = U(n)/O(n)$  in [1].

## 1.10 Historical Remarks

The name “symplectic group” has been introduced in 1939 by Hermann Weyl in [29, p. 165]. There he explained that he formerly advocated the name “complex group” in allusion to Plücker’s linear line complexes, see Exercise 1.3. And that this name became more and more embarrassing through collision with the word “complex” in the connotation of complex number. He therefore proposed to replace it by the Greek adjective “symplectic”. This immediately took hold in the literature. Weyl concluded by remarking that Dickson calls the group “Abelian linear group”, in homage to Abel who first studied it. We could add here that this name is equally embarrassing, as nowadays Abelian group = commutative group.

The names “Lagrange plane” and “Lagrangian Grassmannian” have been introduced by Arnol’d [1], after the name “Lagrange manifold”, introduced by Maslov [23, p. 115] in 1965 for a manifold of which all tangent spaces are Lagrange planes.

## 1.11 Exercises

**Exercise 1.1** In the notation of Subsection 1.9, prove that  $J \in \text{Sp}(E, \sigma)$ . Prove that if  $L$  is a Lagrange plane, then  $J(L)$  is a Lagrange plane which is  $g$ -orthogonal to  $L$ , and therefore satisfies  $J(L) \cap L = \{0\}$ .  $\circlearrowright$

**Exercise 1.2** It was the idea of Plücker (1846), to consider the projective lines in the three-dimensional space as the elements (points) of a new space. The three-dimensional projective space is defined as the space  $\mathbb{P}(E)$  of all one-dimensional linear subspaces  $l$  of the four-dimensional vector space  $E$ . A projective line is equal to the set  $\mathbb{P}(L)$  of all one-dimensional linear subspaces  $l$  of a two-dimensional linear subspace  $L$  of  $E$ . In this way, Plücker’s space is identified with the Grassmann manifold  $G_2(E)$  of all two-dimensional linear subspaces  $L$  of the four-dimensional vector space  $E$ .

Let  $a, b \in E$  be linearly independent. Prove that  $u = a \wedge b$  is a nonzero element of  $\Lambda^2 E$  such that  $u \wedge u = 0$ . Prove that every nonzero  $u \in \Lambda^2 E$  such that  $u \wedge u = 0$  arises in this way, and that  $x \in E$  is equal to a linear combination of  $a$  and  $b$ , if and only if  $u \wedge x = 0$ . Prove that the relation between  $L \in G_2(E)$  and  $u \in \Lambda^2 E$ , that  $u = a \wedge b$  for a basis  $a, b$  of  $L$ , is equivalent to  $u \wedge x = 0$  for every  $x \in L$ . Prove that this relation defines a bijection between  $G_2(E)$  and the quadric  $Q$  in the five-dimensional projective space  $\mathbb{P}(\Lambda^2 E)$  defined by the equation  $u \wedge u = 0$ . Prove that  $(u, v) \mapsto u \wedge v$  is a nondegenerate symmetric bilinear form on  $\Lambda^2 E$  with values in the one-dimensional vector space  $\Lambda^4 E$ , and that, as a consequence, the quadric  $Q$  is smooth.

The coordinates of  $\Lambda^2 E$  are called *Plücker coordinates* on  $G_2(E)$ . Strictly speaking these should be regarded as projective coordinates and, as functions on the projective coordinate patches of  $\mathbb{P}(\Lambda^2 E)$ , be restricted to the quadric  $Q$ .  $\circlearrowright$

**Exercise 1.3** Plücker [26] defined a *line complex of degree  $m$*  as the intersection of  $Q$  with an algebraic hypersurface in  $\mathbb{P}(\Lambda^2 E)$  of degree  $m$ , defined by the equation  $F(u) = 0$  in which  $F$  is a homogeneous polynomial of degree  $m$  on  $\Lambda^2 E$ . He defined a *linear line complex* as a line complex of degree one.

Prove that a linear line complex corresponds to a nonzero two-form  $\sigma \in \Lambda^2 E^*$  on  $E$ , unique up to a nonzero scalar multiple, such that  $L \in G_2(E)$  belongs to the linear line complex if and only if  $L$

is  $\sigma$ -isotropic. Prove that if  $\sigma$  is nondegenerate, then the linear line complex defined by  $\sigma$ , viewed as a subset of  $G_2(E)$ , is equal to the manifold  $\mathcal{L}(E, \sigma)$  of all the Lagrange planes with respect to the symplectic form  $\sigma$ .

Prove that if  $\sigma$  is degenerate, then  $\ker \sigma \in G_2(E)$ . In this case Plücker called the corresponding linear line complex *special* and called  $\ker \sigma$  the *axis* of the special linear line complex. Prove that the special linear line complex is equal to the set of all  $L \in G_2(E)$  such that  $L \cap \ker \sigma \neq 0$ . In terms of projective lines: the special line complex is equal to the set of all projective lines which intersect its axis.  $\circlearrowright$

**Exercise 1.4** We use the notation of Subsection 1.6.

Let  $L \in \mathcal{L}(E, \sigma)$ ,  $I \in G_m(L)$ . Prove that  $\sigma$  induces a symplectic form on the  $(2n - 2m)$ -dimensional vector space  $I^\sigma/I$ . Prove that  $L' \in \mathcal{L}(E, \sigma)$  and  $L' \cap L = I$  if and only if  $L'/I$  is a Lagrange plane in  $I^\sigma/I$  which is transversal to  $L/I$ . Prove that the mapping  $L' \mapsto L' \cap L$  exhibits  $\mathcal{L}_{L,,m}$  as a smooth bundle over  $G_m(L)$  of which each fiber is an affine space of dimension  $\frac{1}{2}(n - m)(n - m + 1)$ . Prove  $\mathcal{L}_{L,,m}$  is a smooth submanifold of  $\mathcal{L}(E, \sigma)$  of dimension equal to  $\frac{1}{2}n(n + 1) - \frac{1}{2}m(m + 1)$ . Prove that the closure of  $\mathcal{L}_{L,,m}$  in  $\mathcal{L}(E, \sigma)$  is equal to  $\bigcup_{l=m}^n \mathcal{L}_{L,l}$ .  $\circlearrowright$

**Exercise 1.5** Let  $\gamma$  be a differentiable curve in  $\mathcal{L} = \mathcal{L}(E, \sigma)$  such that  $\gamma(t_0) \in \mathcal{L}_{L,1}$ . Prove that  $\gamma(t)$  intersects  $\mathcal{L}_{L,1}$  transversally at  $t = t_0$ , i.e.  $\gamma'(t_0) \notin T_{\gamma(t_0)} \mathcal{L}_{L,1}$ , if and only if the restriction to  $\gamma(t_0) \cap L$  of the symmetric bilinear on  $\gamma(t_0)$ , which is assigned to  $\gamma't(t_0) \in T\mathcal{L}$  as in Subsection 1.6, is nonzero. We will call the intersection positive or negative according to whether this restriction is positive or negative definite, respectively.

It is known that by slight perturbation every closed curve  $\gamma$  in  $\mathcal{L}$  can be made such that all intersections with  $\mathcal{L}_{L,1}$  are transversal. prove that this implies that  $\gamma$  intersects  $\mathcal{L}_{L,1}$  only finitely many times. Write  $i(\gamma)$  for the number of positive intersections minus the number of negative intersections.

It is also known that by slight perturbation any homotopy of closed curves, viewed as a mapping from a two-dimensional cylinder to  $\mathcal{L}$  can be made to miss each of the manifolds  $\mathcal{I}_{L,m}$  with  $m \geq 2$ , each of which has codimension  $\geq 3$  in  $\mathcal{L}$ . This shows that  $i(\gamma)$  only depends on the homotopy class of  $\gamma$ , and  $i$  induces a homomorphism from the fundamental group  $\pi_1(\mathcal{L})$  of  $\mathcal{L}$  to  $\mathbf{Z}$ .

Prove that  $\mathcal{L}_{L,1}$  is connected and that  $\mathcal{L}_{L,0}$  is simply connected. Prove that if  $i(\gamma) = 0$ , then  $\gamma$  is contractible in  $\mathcal{L}$ . Finally, find a smooth closed curve  $\gamma$  in  $\mathcal{L}$  such that  $i(\gamma) = 1$  and prove that  $i : \pi_1(\mathcal{L}) \rightarrow \mathbf{Z}$  is an isomorphism.

**Remark** Maslov [23, p. 147–149] called the set  $\Sigma = \overline{\mathcal{L}_{L,1}}$  and the integer  $i(\gamma)$  the *singular set* and the *index* of the curve  $\gamma$ , respectively. Arnol'd [1] observed that  $\Sigma$  defines an oriented codimension one cycle in  $\mathcal{L}$ , that the index is equal to the topological intersection number of the one-dimensional cycle  $\gamma$  with  $\Sigma$  and that the index defines an isomorphism of the fundamental group of  $\mathcal{L}$  with  $\mathbf{Z}$ .

$\circlearrowright$

## 2 Symplectic Manifolds

### 2.1 Definition

Let  $M$  be a finite-dimensional smooth manifold. A *symplectic form* on  $M$  is a smooth differential form  $\sigma$  of degree two on  $M$  such that

- i) For every  $m \in M$  the bilinear form  $\sigma_m$  on  $T_m M$  is nondegenerate, and
- ii)  $\sigma$  is closed, i.e.  $d\sigma = 0$ .

Condition i) means that, for every  $m \in M$ ,  $\sigma_m$  is a symplectic form on  $T_m M$ . This implies that  $\dim M = \dim T_m M$  is even, say equal to  $2n$ , cf. Subsection 1.4. A *symplectic manifold* is defined as a pair  $(M, \sigma)$ , in which  $M$  is a finite-dimensional smooth manifold and  $\sigma$  is a symplectic form on  $M$ .

**Example 2.1** Probably the simplest example is  $\mathbf{R}^n \times \mathbf{R}^n$  provided with the "constant" standard symplectic form of (1.2), which in differential form notation is equal to

$$\sigma = \sum_{j=1}^n dp_j \wedge dq_j. \quad (2.1)$$

Here  $p_j$  and  $q_j$  are viewed as (coordinate) *functions* on  $\mathbf{R}^n \times \mathbf{R}^n$ . ⊙

## 2.2 The Cotangent Bundle

An important generalization of the previous example is obtained by starting with an arbitrary  $n$ -dimensional smooth manifold  $X$ . The *cotangent bundle*  $T^*X$  of  $X$  is defined as the vector bundle over  $X$  of which the fiber at the point  $x \in X$  is equal to the dual  $T_x^*X := (T_x X)^*$  of  $T_x X$ , the space of all linear forms  $\xi$  on the tangent space  $T_x X$  of  $X$  at the point  $x$ .

Let  $\pi$  denote the projection from  $M := T^*X$  onto  $X$ . which sends every element of  $T_x^*X$  to  $x$ . Then, for every  $\xi \in T_x^*X$ , the tangent map  $T_\xi \pi$  is a linear mapping from  $T_\xi M$  onto  $T_x X$ , and if we subsequently apply the linear form  $\xi \in (T_x X)^*$  to it, we obtain the linear form

$$\tau_\xi := \xi \circ T_\xi \pi \quad (2.2)$$

on  $T_\xi M$ . This defines a special smooth differential form  $\tau$  of degree one on  $M$ .

If  $\alpha$  is any smooth differential form of degree one on  $X$ , then it can also be viewed as a smooth mapping  $\alpha : X \rightarrow T^*X$  such that  $\pi \circ \alpha$  is equal to the identity on  $X$ . It follows that

$$(\alpha^* \tau)_x = \tau_{\alpha(x)} \circ T_x \alpha = \alpha(x) \circ T_{\alpha(x)} \pi \circ T_x \alpha = \alpha(x) \circ T_x (\pi \circ \alpha) = \alpha(x),$$

where in the first, second, third and last identity we used the definition of pullback of a differential form, the definition of  $\tau$ , the chain rule for differentiation and  $\pi \circ \alpha = \text{Id}$ , respectively. The equation

$$\alpha = \alpha^* \tau \quad \text{for every one-form } \alpha \text{ on } X \quad (2.3)$$

says that every one-form on  $X$  is equal to the pullback of  $\tau$  by means of the one-form viewed as a mapping from  $X$  to  $T^*X$ . For this reason  $\tau$  is called the *tautological one-form* on the cotangent bundle.

The exterior derivative

$$\sigma := d\tau \quad (2.4)$$

of the tautological one-form is a two-form on  $T^*X$ , which is closed because  $d(d\omega) = 0$  for every differential form  $\omega$  (of any degree). Moreover, in local coordinates  $(x_1, \dots, x_n)$  on  $X$ , with corresponding dual coordinates  $(\xi_1, \dots, \xi_n)$ , the equation (2.2) takes the form

$$\tau = \sum_{i=1}^n \xi_i dx_i, \quad (2.5)$$

and therefore

$$\sigma = \sum_{i=1}^n d\xi_i \wedge dx_i, \quad (2.6)$$

which shows that  $\sigma$  is equal to the standard symplectic form if we substitute  $x_i = q_i$  and  $\xi_i = p_i$ . This shows that  $\sigma := d\tau$  is a symplectic form on  $T^*X$ , which is called the *canonical symplectic form* of the cotangent bundle.

### 2.3 Reduction

Let  $N$  be a smooth manifold and let  $\omega$  be a closed smooth two-form on  $N$  for which the kernel has constant rank, i.e. there exists a nonnegative integer  $k$  such that, for every  $n \in N$ ,  $\dim(\ker \omega_n) = k$ . This implies that the  $K_n := \ker \omega_n$ ,  $n \in N$ , define a smooth vector subbundle  $K$  of the tangent bundle  $TN$  of  $N$ . (In the present differential geometric terminology, a smooth vector subbundle of the tangent bundle of  $N$  is also called a *distribution* on  $N$ , not to be confused with the distributions in Analysis. In the 19-th century literature a smooth vector subbundle of the tangent bundle of  $N$  is called a *Pfaffian system* in  $N$ .)

A general smooth vector subbundle  $K$  of the tangent bundle  $TN$  of any finite-dimensional smooth manifold  $N$  is called *integrable* if for each  $n_0 \in N$  there exists an open neighborhood  $N_0$  of  $n_0$  in  $N$  and a smooth fibration of  $N_0$ , such that, for each  $n \in N_0$ ,  $K_n$  is equal to the tangent space of the fiber through  $n$ . The *theorem of Frobenius* says that  $K$  is integrable if and only if  $[X, Y] \subset K$  holds for any pair of smooth vector fields  $X, Y$  on  $N$  such that  $X \subset K$  and  $Y \subset K$ .

Returning to our two-form  $\omega$  with kernel of constant rank, the claim is that the closedness of  $\omega$  implies that its kernel  $K := \ker \omega$  is integrable. For the proof we use that  $[X, Y]$  is equal to the Lie derivative  $\mathcal{L}_X Y$  of  $Y$  with respect to the vector field  $X$ . In view of the Leibniz formula for Lie derivatives, it follows that

$$i_{[X, Y]} \omega = \mathcal{L}_X (i_Y \omega) - i_Y (\mathcal{L}_X \omega),$$

whereas the homotopy formula for the Lie derivative yields that

$$\mathcal{L}_X \omega = i_X(d\omega) + d(i_X \omega). \quad (2.7)$$

Therefore,  $i_Y \omega = 0$ ,  $i_X \omega = 0$  and  $d\omega = 0$  imply that  $i_{[X, Y]} \omega = 0$ .

Now suppose that  $\pi : N \rightarrow M$  is a fibration with connected fibers, such that, for each  $n \in N$ ,  $\ker \omega_n$  is equal to the tangent space  $\ker T_n \pi$  of the fiber through the point  $n$ . Fix  $m \in M$ . Then there exists, for each  $n \in \pi^{-1}(\{m\})$ , a unique two-form  $\sigma_{m, n}$  on  $T_m M$ , such that  $(T_n \pi)^* \sigma_{m, n} = \omega_n$ . However, for each smooth vector field  $X$  such that  $X \subset K$  we see from (2.7) that  $\mathcal{L}_X \omega = 0$ , which implies that  $(e^{tX})^* \omega = \omega$ , whereas on the other hand  $\pi \circ e^{tX} = \pi$ . This implies that  $\sigma_{m, n'} = \sigma_{m, n}$  if  $n' = e^{tX}(n)$ . Because the compositions of the flows  $e^{tX}$  with  $X \subset K$  act locally transitively on the fibers, and therefore transitively because the fibers are connected, the conclusion is that  $\sigma_{m, n} = \sigma_m$  does not depend on the choice of  $n \in \pi^{-1}(\{m\})$ .

The  $\sigma_m$ ,  $m \in M$ , define a smooth two-form on  $M$  such that  $\omega = \pi^*\sigma$ . Because

$$\ker T_n \pi = \ker \omega_n = T_n \pi^{-1}(\ker \sigma_m),$$

it follows that  $\ker \sigma_m = \{0\}$ , which means that  $\sigma_m$  is a symplectic form. Furthermore,  $\pi^*(d\sigma) = d(\pi^*\sigma) = d\omega = 0$ , which in view of the surjectivity of the linear mappings  $T_n \pi$ ,  $n \in N$  implies that  $d\sigma = 0$ . In other words,  $(M, \sigma)$  is a symplectic manifold, which is called the *reduced symplectic manifold* of the pair  $(N, \omega)$ .

## 2.4 Complex Projective Varieties

Let  $E$  be an  $n$ -dimensional complex vector space with Hermitian form  $h$ , which we provide with real inner product  $g = \operatorname{Re} h$  and the symplectic form  $\sigma = \operatorname{Im} h$  as in Subsection 1.9, which are related by  $g = \sigma \circ J$ . Let

$$S := \{z \in E \mid h(z, z) = g(z, z) = 1\}$$

be the unit sphere in  $E$  with respect to the inner product  $g$ . Then, for each  $z \in S$ ,

$$T_z S = \{v \in E \mid g(z, v) = 0\} = \{v \in E \mid \sigma(i z, v) = 0\}.$$

It follows that  $i z \in T_z S$ , and  $i z$  belongs to the kernel of the restriction to  $T_z S$ . Because  $T_z S$  has real codimension one in  $E$ , its symplectic orthogonal complement is real one-dimensional, and therefore equal to  $\mathbf{R} i z$ . On the other hand  $\mathbf{R} i z$  is equal to the tangent space of the circle

$$C_z := \{c z \mid c \in \mathbf{C}, |c| = 1\} = (\mathbf{C} z) \cap S$$

through the point  $z$ . If we denote the identity mapping from  $S$  to  $E$  by  $\iota$ , then  $\iota^*\sigma$  is the restriction to  $S$  of the two-form  $\sigma$ , and the fibration of  $S$  by the integral curves of the kernels of  $\iota^*\sigma$  is equal to the fibration of  $S$  by the circles  $C_z$ ,  $z \in S$ . On the space  $M$  of these circles we have the reduced symplectic form  $\hat{\sigma}$ , the unique two-form  $\hat{\sigma}$  on  $M$  such that  $\iota^*\sigma = \pi^*\hat{\sigma}$ , if  $\pi : S \rightarrow M$  denotes the projection defined by  $\pi(z) = C_z$ ,  $z \in S$ .

On the other hand the  $\mathbf{C} z$ ,  $z \in S$ , are the complex one-dimensional linear subspaces  $l$  of  $E$ , which form the elements of the complex  $(n - 1)$ -dimensional projective space  $\mathbf{CP}(E)$  of  $E$ . The mapping  $l \mapsto l \cap S$  is a diffeomorphism from  $\mathbf{CP}(E)$  onto  $M$ , which can be used in order to identify  $M$  with  $\mathbf{CP}(E)$ . Note that  $\mathbf{CP}(E)$  is a complex analytic manifold, with a complex multiplication  $J_l$  by  $i$  in each tangent space  $T_l(\mathbf{CP}(E))$ . It is easily verified that  $\hat{g} := \hat{\sigma} \circ J$  is the inner product on  $T_l(\mathbf{CP}(E))$  which corresponds to the restriction of  $g$  to the  $g$ -orthogonal complement of  $i z$  in  $T_z S$ .

If  $E = \mathbf{C}^n$  and  $h$  is the standard hermitian structure on  $\mathbf{C}^n$ , then the Hermitian inner product  $\hat{h} := \frac{1}{\pi}(\hat{g} + i\hat{\sigma})$  is called the *Fubini-Study metric* on  $\mathbf{CP}(E) = \mathbf{CP}_{n-1}$ . Here the factor  $1/\pi$  is inserted in order to arrange that the integral of  $\omega := \frac{1}{\pi}\hat{\sigma}$  over any complex projective line in  $\mathbf{CP}_{n-1}$  is equal to one.

More generally, if  $M$  is a complex analytic manifold, then a *Kähler structure* on  $M$  is a smooth Hermitian inner product  $h$  on its tangent bundle, such that its imaginary part, the two-form  $\sigma := \operatorname{Im} h$ , is closed. This implies that  $\sigma$  is a symplectic form on  $M$ , called the *Kähler form* of the *Kähler manifold*  $(M, h)$ . The Fubini-Study metric is a Kähler structure on the complex projective space.

If  $\iota : V \rightarrow M$  is a complex analytic submanifold of a Kähler manifold  $(M, h)$ , then the restriction  $\iota^*h$  of  $h$  to  $TV$  is a Kähler structure on  $V$ . This exhibits every smooth complex projective variety as a Kähler manifold, by providing it with the restriction to its tangent bundle of the Fubini-Study metric of the projective space of which it is a subvariety. This is a very rich source of examples of compact symplectic manifolds. Vice versa, this introduces symplectic differential geometry into complex algebraic geometry.

## 2.5 Almost Complex Structure

An *almost complex structure*  $J$  on a smooth manifold  $M$  is a complex structure  $J_m$  on each tangent space  $T_m M$ , depending smoothly on  $m \in M$ . As in Subsection 1.9, this turns each tangent space into a complex vector space. If  $n$  is the complex dimension of  $T_m M$  with respect to the complex structure  $J_m$ , then the real dimension is equal to  $2n$ .

An *almost complex manifold* is a pair  $(M, J)$  in which  $M$  is a smooth manifold and  $J$  is an almost complex structure on  $M$ . The almost complex structure is called *integrable* if for every  $m_0 \in M$  there is coordinate system in an open neighborhood of  $m_0$  in  $M$ , in which  $m \mapsto J_m$  is constant. If we use the constant complex structure in order to identify  $\mathbf{R}^{2n}$  with  $\mathbf{C}^n$ , we obtain a system of local coordinatizations for which the coordinate changes are complex analytic mappings. With these coordinatizations,  $M$  is a complex analytic manifold for which the  $J_m$  are the multiplications by  $i$  in the tangent spaces.

For each  $m \in M$  one has the antisymmetric bilinear mapping  $[J, J]_m$  from  $T_m M \times T_m M$  to  $T_m M$ , which is defined by

$$[J, J](v, w) = [Jv, Jw] - J[Jv, w] - J[v, Jw] - [v, w], \quad (2.8)$$

in which  $v$  and  $w$  are smooth vector fields on  $M$  and the brackets in the right hand side are the Lie brackets of vector fields. The *theorem of Newlander and Nirenberg* says that the almost complex structure  $J$  is integrable, if and only if  $[J, J] = 0$ , cf. [25], [13], [22].

**Lemma 2.2** *Let  $\sigma$  be a symplectic form on  $M$ . Then there exists an almost complex structure  $J$  on  $M$  such that  $h = \sigma \circ J + i \sigma$  is a Hermitian structure on  $TM$ . If  $\sigma$  is invariant under the action of a group  $G$  on  $M$ , and  $M$  carries a  $G$ -invariant Riemannian structure, then  $J$  can be chosen to be  $G$ -invariant as well.*

**Proof** There exists a Riemannian structure  $g$  in  $M$ . Such a Riemannian structure exists in local coordinates. Let  $\xi_j$  be a smooth partition of unity subordinate to a locally finite covering of  $M$  by means of open subsets  $M_j$  on which we have a Riemannian structure  $g_j$ . This means that the  $\xi_j$  are smooth real valued functions on  $M$  such that  $\xi_j \geq 0$ , the support of  $\xi_j$  is contained in  $M_j$  and  $\sum_j \xi_j = 1$ . Then  $g = \sum_j \xi_j g_j$  is the desired Riemannian structure on  $M$ .

Define, for each  $m \in M$ ,  $A_m := \sigma_m^{-1} g_m$ . Then  $g_m \circ A_m = g_m \circ \sigma_m^{-1} \circ g_m$  is antisymmetric, or  $A_m$  is  $g_m$ -antisymmetric, and there exists a  $g_m$ -orthonormal basis in  $T_m M$  on which the matrix of  $A_m$  consists of  $2 \times 2$ -matrices  $\begin{pmatrix} 0 & -a_j \\ a_j & 0 \end{pmatrix}$  along the diagonal, with  $a_j > 0$ . Let  $B_m$  be the linear transformation in  $T_m M$  of which the matrix consists of the  $2 \times 2$ -matrices  $\begin{pmatrix} 1/a_j & 0 \\ 0 & 1/a_j \end{pmatrix}$  along the diagonal. Then  $J_m := A_m \circ B_m = \sigma_m^{-1} \circ g_m \circ B_m$  is a complex structure on  $T_m M$ ,

$\widehat{g}_m := g_m \circ B_m = \sigma_m \circ J_m$  is an inner product on  $T_m M$  and therefore  $\sigma_m \circ J_m + i \sigma$  is a Hermitian form on  $T_m M$ .

At first sight this construction seems to depend on the choice of the  $g_m$ -orthonormal basis, but  $B_m^2 = -A_m^{-2}$ , which shows that actually  $B_m$  is equal to the unique positive definite square root of the positive definite linear transformation  $-A_m^{-2}$  (all with respect to the inner product  $g_m$ ). This makes  $J_m$  globally well-defined and depending smoothly on  $m \in M$ .

If  $\sigma$  and  $g$  are  $G$ -invariant, then the uniqueness of the  $B_m$ ,  $m \in M$ , make that also  $B$  and  $J$  are  $G$ -invariant.  $\square$

It cannot always be arranged that in addition  $J$  is integrable. In other words, not every symplectic form is equal to a Kähler form on a complex analytic manifold. However, the weaker almost complex structure is sufficient for many purposes.

## 2.6 Cohomology Classes

It follows from the observation in Subsection 1.8 that the  $n$ -th power of a symplectic form is nonzero that  $\sigma^n$  is a nowhere vanishing volume form on  $M$ .

Assume in the sequel that  $M$  is compact and connected. Then the de Rham cohomology class  $[\sigma^n] \in H^{2n}(M)$  of the nowhere vanishing volume form  $\sigma^n$  is nonzero, and therefore generates the one-dimensional vector space  $H^{2n}(M)$ . Because  $[\sigma^n] = [\sigma]^n$ , this in turn implies that the element  $[\sigma]^k \in H^{2k}(M)$  is nonzero for every  $1 \leq k \leq n$ .

The fact that  $H^{2k}(M) \neq 0$ , or more precisely that there exists an  $s \in H^2(M)$  such that  $s^k \neq 0$  for every  $1 \leq k \leq n$ , puts a quite severe topological restriction on a compact smooth manifold for allowing a symplectic form. For instance, if  $M$  is a  $2n$ -dimensional sphere, then  $H^p(M) = 0$  for all  $p$  except  $p = 0$  and  $p = 2n$ , and therefore the only sphere which can carry a symplectic form is the two-dimensional one.

If  $M$  is the complex  $n$ -dimensional complex projective space, then  $H^p(M) = 0$  except when  $p = 2k$ ,  $0 \leq k \leq n$ , in which case  $\dim H^{2k}(M) = 1$ . In other words, in this case the whole cohomology ring is generated by the cohomology class  $[\omega]$  of the Kähler form  $\omega$  defined by the Fubini-Study metric. This is even true for the cohomology ring with values in  $\mathbf{Z}$ , cf. [9, pp. 60, 150].

## 2.7 Exercises

**Exercise 2.1** let  $X$  and  $Y$  be smooth manifolds and let  $\phi : X \rightarrow Y$  be a local diffeomorphism. Define the *induced transformation*  $\Phi : T^*X \rightarrow T^*Y$  by

$$\Phi(x, \xi) := \left( \phi(x), ((T_x \phi)^*)^{-1}(\xi) \right), \quad x \in X, \xi \in (T_x X)^*.$$

Prove that

$$\Phi^* \tau_{T^*Y} = \tau_{T^*X}$$

and that

$$\Phi^* \sigma_{T^*Y} = \sigma_{T^*X}.$$

In other words, the induced mapping is a *canonical transformation*, in the sense that it preserves the canonical symplectic forms.  $\circlearrowright$



**Exercise 2.2** Consider the standard coordinatization

$$\varphi : (z_1, \dots, z_n) \mapsto \mathbf{C} (1, z_1, \dots, z_n)$$

of the open subset of the  $n$ -dimensional complex projective space which consist of the one-dimensional complex linear subspaces  $l$  which are not contained in the  $n$ -dimensional linear subspace defined by the equation  $z_0 = 0$ . Write

$$w = \psi(z) := (1 + (z, z))^{-1/2} (1, z_1, \dots, z_n),$$

$w_j = u_j + i v_j$  with  $u_j, v_j \in \mathbf{R}$ ,  $0 \leq j \leq n$ , and  $z_j = x_j + i y_j$  with  $x_j, y_j \in \mathbf{R}$ ,  $1 \leq j \leq n$ . Prove that the pullback of the Fubini-Study Kähler form  $\omega$  under  $\varphi$  satisfies

$$\begin{aligned} \pi \omega &:= \psi^* \left( \sum_{j=0}^n dv_j \wedge du_j \right) = (1 + (z, z))^{-1} \sum_{j=1}^n dy_j \wedge dx_j \\ &\quad - (1 + (z, z))^{-2} \sum_{j,k=1}^n (x_j x_k + y_j y_k) dy_j \wedge dx_k + x_j y_k (dx_j \wedge dx_k + dy_j \wedge dy_k). \end{aligned}$$

(This may be compared with [9, p. 30, 31].) How easy is it to verify by direct computation that the two-form in the right hand side is closed?

Verify that for  $n = 1$  we have

$$\pi \omega = (1 + x^2 + y^2)^{-2} dy \wedge dx,$$

and that the integral of  $\omega$  over the complex projective line is equal to 1. ⊙

**Exercise 2.3** Let  $J$  be an almost complex structure on the manifold  $M$ . Prove that  $[J, J](v, Jv) = 0$  for any smooth vector field  $v$  on  $M$  and prove that  $J$  is integrable if  $\dim M = 2$ . Now assume that  $\dim M = 2$ , that  $M$  is connected and that  $\sigma$  is a nowhere vanishing area form = symplectic form on  $M$ . Prove that  $g := \sigma \circ J$  is a symmetric bilinear form on each tangent space which is invariant under the linear transformations  $e^{tJ_m}$ . Prove that  $g$  is either positive definite or negative definite. In other words,  $M$  is a complex analytic "curve" and either  $\sigma$  or  $-\sigma$  is equal to the Kähler form of a Kähler structure on  $M$ . ⊙

**Exercise 2.4** For which compact oriented surfaces is the cohomology ring generated by the class of a symplectic form? ⊙

**Exercise 2.5** If  $n = 2$  in Subsection 2.4, identify the projection  $\pi : S \rightarrow \mathbf{CP}(E)$  with the Hopf fibration. ⊙

### 3 Hamiltonian Systems

#### 3.1 Flows of Vector Fields

Let  $M$  be a smooth manifold. If  $v$  is a smooth vector field on  $M$  then the general theory of systems of ordinary differential equations, see for instance [3], implies the following statements. For every  $m \in M$ , there is a unique maximal solution  $\gamma = \gamma_m : I_m \rightarrow M$  of the differential equation  $d\gamma(t)/dt = v(\gamma(t))$ , such that  $\gamma(0) = m$ . The domain of definition  $I_m$  of this maximal solution  $\gamma$  is an open interval in  $\mathbf{R}$  containing 0. If  $s := \sup I_m < \infty$  or  $i := \inf I_m > -\infty$ , then there exists for every compact subset  $K$  of  $M$  an  $\epsilon > 0$  such that  $\gamma(t) \notin K$  for every  $t \in ]s - \epsilon, s[$  or  $t \in ]i, i + \epsilon[$ , respectively. In other words, the only way the maximal solutions do not exist for all time is that they run out of every compact subset of  $M$  in a finite time. This implies that if  $\gamma_m(t)$  stays within a compact subset of  $M$  for all  $t \in I_m$ , then  $I_m = \mathbf{R}$ . The set  $D := \{(t, m) \in \mathbf{R} \times M \mid t \in I_m\}$  is an open subset of  $\mathbf{R} \times M$  which contains  $\{0\} \times M$ , and  $A : (t, m) \mapsto \gamma_m(t)$  is a smooth mapping from  $D$  to  $M$ . If  $s \in I_m$  and  $t \in I_{\gamma_m(s)}$ , then  $s + t \in I_m$  and  $\gamma_m(s + t) = \gamma_{\gamma_m(s)}(t)$ . This follows from the uniqueness of the solutions, because, as a function of  $t$ ,  $\gamma_m(s + t)$  and  $\gamma_{\gamma_m(s)}(t)$  satisfy the same differential equation and have the same initial value.

We say that the vector field  $v$  is *complete* if  $D = \mathbf{R} \times M$ , in which case  $A : \mathbf{R} \times M \rightarrow M$  is a smooth action of the additive group  $(\mathbf{R}, +)$  on  $M$ . The mapping  $m \mapsto \gamma_m(t) : M \rightarrow M$  is called the *time  $t$  flow of the vector field  $v$*  and will be denoted by  $e^{tv}$ . This notation reminds of the defining equation  $de^{tv}/dt = v \circ e^{tv}$  and of the group homomorphism property  $e^{(s+t)v} = e^{tv} \circ e^{sv}$ . Because  $e^{tv} \circ e^{-tv} = \text{Id} = e^{-tv} \circ e^{tv}$ , the time  $t$  flow is a diffeomorphism of  $M$ , i.e. it is bijective from  $M$  to  $M$  and has a smooth inverse (equal to  $e^{-tv}$ ).

If  $D$  is a proper subset of  $\mathbf{R} \times M$ , then the flow  $e^{tv}$  is defined on the open subset

$$M_t := \{m \in M \mid (t, m) \in D\}$$

of  $M$ , and  $e^{(s+t)v}(m) = e^{tv} \circ e^{sv}(m)$  holds when  $m \in M_s$  and  $e^{sv}(m) \in M_t$ , in which case  $m \in M_{s+t}$ . Also,  $e^{tv}$  is a diffeomorphism from the open subset  $M_t$  of  $M$  onto the open subset  $M_{-t}$  of  $M$ , with inverse equal to  $e^{-tv}$ . If  $D \neq \mathbf{R} \times M$  we don't have a group action, but we have the same identities "as far as the objects appearing in the formulas are defined". Lie [21] called  $t \mapsto e^{tv}$  the *one-parameter group of transformations generated by the vector field  $v$* , where he did not worry about domains of definition.

#### 3.2 Lie Derivatives

Let  $\Omega^p(M)$  denote the space of smooth  $p$ -forms on the smooth manifold  $M$ , where  $\Omega^0(M) = \mathcal{F}(M)$  denotes the space of smooth real valued functions on  $M$  and  $\Omega^p(M) = \{0\}$  if  $p > \dim M$ . If  $M$  and  $N$  are smooth manifolds of any dimensions, and  $\varphi : M \rightarrow N$  is a smooth map, then for any  $\omega \in \Omega^p(N)$  the *pullback  $\varphi^*\omega$  of  $\omega$  under the map  $\varphi$*  is defined by

$$(\varphi^*\omega)_m(v_1, \dots, v_n) = \omega_{\varphi(m)}(\mathbf{T}_m \varphi v_1, \dots, \mathbf{T}_m \varphi v_n). \quad (3.1)$$

Note that  $\varphi^*$  defines a continuous linear operator from  $\Omega^p(N)$  to  $\Omega^p(M)$ , where the word "pullback" reminds of the fact that this goes in the opposite direction of the map  $\varphi : M \rightarrow N$ . For  $p = 0$  we have  $\varphi^*\omega = \omega \circ \varphi$ , which means that the pullback under  $\varphi$  is just the substitution  $n = \varphi(m)$  in  $\omega(n)$ , and  $\varphi^*$  is just a convenient notation for the linear mapping  $\omega \mapsto \omega \circ \varphi$ .

The transposition symbol  $*$  in the notation is a reminder to the transposition of  $\varphi$  in the formula  $\varphi^*\omega = \omega \circ \varphi$ . It also helps reminding that in the natural composition formula  $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$  the order is reversed: pullback is an *antihomomorphism* with respect to composition.

It is known that the exterior derivative  $d : \Omega^p(M) \rightarrow \Omega^{p+1}(M)$  of differential forms behaves naturally under smooth mappings, in the sense that

$$\varphi^*(d\omega) = d(\varphi^*\omega), \quad \omega \in \Omega^p(N), \quad \varphi : M \rightarrow N. \quad (3.2)$$

Let  $\mathcal{X}(M)$  denote the vector space of smooth vector fields on  $M$ . If  $v \in \mathcal{X}(M)$ , then the *Lie derivative*  $\mathcal{L}_v\omega$  of  $\omega \in \Omega^p(M)$  with respect to the vector field  $v$  is defined as

$$\mathcal{L}_v\omega := \left. \frac{d}{dt} (e^{tv})^* \omega \right|_{t=0}. \quad (3.3)$$

If in

$$(e^{tv})^* \circ (e^{sv})^* \omega = (e^{sv} \circ e^{tv})^* \omega = (e^{(t+s)v})^* \omega$$

we take the derivative with respect to  $s$  at  $s = 0$  in the left and right hand side, we obtain that

$$(e^{tv})^* \mathcal{L}_v\omega = \frac{d}{dt} (e^{tv})^* \omega.$$

This implies that  $\mathcal{L}_v\omega = 0$  if and only if  $t \mapsto (e^{tv})^* \omega$  is constant, hence equal to its value  $\omega$  at  $t = 0$ . We say that  $\omega$  is called *invariant under the flow of  $v$*  if, for every  $t \in \mathbf{R}$ ,  $(e^{tv})^* \omega = \omega$  in  $M_t$ . We have proved that  $\omega$  is invariant under the flow of  $v$  if and only if the Lie derivative of  $\omega$  with respect to  $v$  is equal to zero.

If  $v \in \mathcal{X}(M)$  and  $\omega \in \Omega^p(M)$ , then the *inner product*  $i_v\omega \in \Omega^{p-1}(M)$  of  $\omega$  with  $v$  is defined by

$$(i_v\omega)_m(v_2, \dots, v_n) = \omega_m(v(m), v_2, \dots, v_n). \quad (3.4)$$

In other words,  $i_v$  is the continuous linear operator from  $\Omega^p(M)$  to  $\Omega^{p-1}(M)$  of *inserting the vector field  $v$  at the first slot*.

If  $p = 0$ , then obviously  $\mathcal{L}_v\omega = i_v(d\omega)$ , which is the *derivative of the function  $\omega$  in the direction of the vector field  $v$* . For general  $p$  we have the *homotopy formula*

$$\mathcal{L}_v\omega = i_v(d\omega) + d(i_v\omega), \quad (3.5)$$

or  $\mathcal{L}_v = i_v \circ d + d \circ i_v$ . Note that in the first summand  $d$  and  $i_v$  is a linear operator from  $\Omega^p(M)$  to  $\Omega^{p+1}(M)$  and from  $\Omega^{p+1}(M)$  to  $\Omega^p(M)$ , respectively, whereas in the second summand  $i_v$  and  $d$  is a linear operator from  $\Omega^p(M)$  to  $\Omega^{p-1}(M)$  and from  $\Omega^{p-1}(M)$  to  $\Omega^p(M)$ , respectively. This makes the beautiful formula (3.5) easy to remember.

Let  $v \in \mathcal{X}(M)$  and let  $\varphi : M \rightarrow M$  be a smooth mapping for which, at each point  $m \in M$ ,  $T_m\varphi : T_m M \rightarrow T_{\varphi(m)} M$  is invertible. Then the *pullback*  $\varphi^*v \in \mathcal{X}(M)$  on  $M$  of  $v$  under  $\varphi$  is defined by

$$(\varphi^*v)(m) := (T_m\varphi)^{-1} v(\varphi(m)), \quad m \in M. \quad (3.6)$$

This definition has been arranged in such a way that

$$\varphi^*(i_v\omega) = i_{\varphi^*v} \varphi^*\omega$$

for any  $v \in \mathcal{X}(M)$  and  $\omega \in \Omega^p(M)$ . If  $w$  is another smooth vector field on  $M$ ,  $\varphi = e^{tw}$ , and we differentiate the left and right hand side with respect to  $t$  at  $t = 0$ , then we obtain that

$$\mathcal{L}_w(i_v \omega) = i_{[w, v]} \omega + i_v(\mathcal{L}_w \omega), \quad (3.7)$$

if we define the *Lie brackets*  $[w, v] \in \mathcal{X}(M)$  of the vector fields  $w$  and  $v$  on  $M$  by means of

$$[v, w] = \mathcal{L}_v w := \left. \frac{\partial}{\partial t} (e^{tv})^* w \right|_{t=0} := \left. \frac{\partial}{\partial t} \left( \frac{\partial}{\partial s} e^{-tv} \circ e^{sw} \circ e^{tv} \right) \right|_{s=0} \Big|_{t=0}. \quad (3.8)$$

Note that this yields the opposite sign compared with the usual definition for the Lie algebra of a Lie group, but it is probably better to conform with the generally accepted definition of Lie brackets of vector fields.

In local coordinates, the Lie brackets of the vector fields  $v$  and  $w$  can be computed as

$$[v, w](m) = (Dw)(m)v(m) - (Dv)(m)w(m). \quad (3.9)$$

If in the local coordinates the vector fields are linear, then this leads to  $[v, w] = w \circ v - v \circ w$ , where the right hand side is equal to the opposite of the commutator of  $w$  and  $v$ .

If in (3.7) we substitute  $\omega = df$  in which  $f$  is a smooth function, then we obtain that

$$\mathcal{L}_{[w, v]} f = [\mathcal{L}_w, \mathcal{L}_v] f, \quad (3.10)$$

in which  $[A, B] := A \circ B - B \circ A$  denotes the *commutator* of the linear operators  $A$  and  $B$ . It is customary in differential geometry to identify the smooth vector field  $v$  with the derivation  $D = \mathcal{L}_v$  of functions in the direction of  $v$ , and with this identification the Lie brackets of vector fields is defined as their commutator. This definition has been adopted in Lie [21, Vol. I], where the identity

$$(e^{tv})^* f = e^{t\mathcal{L}_v} f := \sum_{k=0}^{\infty} \frac{t^k}{k!} (\mathcal{L}_v)^k f,$$

which is valid if the vector field  $v$  and the function  $f$  are analytic, is presented as another motivation for the exponential notation for the one-parameter group of transformations generated by the vector field  $v$ .

Actually, since derivations are linear operators in the space of smooth functions, and linear operators are usually denoted by capital letters, the identification of vector fields with derivations has led to the custom in differential geometry and in Lie groups to denote vector fields and elements of the Lie algebra by capital letters, usually  $X$ . This in turn has led to the notation  $\mathcal{X}(M)$  for the Lie algebra of all smooth vector fields on  $M$ . Lie's "continuous groups" were subgroups  $G$  of the groups of diffeomorphism of a smooth manifold  $M$ , which depend smoothly on parameters. The corresponding infinitesimal transformations form a Lie subalgebra  $\mathfrak{g}$  of  $\mathcal{X}(M)$ .

If in the local coordinates the vector fields are linear, then (3.9) leads to  $[v, w] = w \circ v - v \circ w$ , where the right hand side is equal to the opposite of the commutator of  $w$  and  $v$ . This is another consequence of the opposite sign choice for the Lie brackets of vector fields as compared to the one in the Lie algebra of a Lie group.

### 3.3 Hamiltonian Vector Fields

Let  $(M, \sigma)$  be a symplectic manifold. Then, according to Subsection 3.2, the flow  $e^{t v}$  of the smooth vector field  $v \in \mathcal{X}(M)$  leaves the symplectic form  $\sigma$  invariant, if and only if

$$0 = \mathcal{L}_v \sigma = i_v(d\sigma) + d(i_v \sigma) = d(i_v \sigma),$$

i.e. if and only if the one-form  $i_v \sigma$  is closed. Here we have used the homotopy formula (3.5) for the Lie derivative in the second identity and the fact that  $\sigma$  is closed in the third identity.

In turn the condition that  $i_v \sigma$  is closed is locally equivalent to the condition that  $i_v \sigma$  is equal to the total derivative of a smooth function. In formula:

$$i_v \sigma = -df \tag{3.11}$$

for a locally defined smooth function  $f$ , where the minus sign is a matter of convention. If  $H^1(M) = 0$ , then there is a globally defined function  $f$  on  $M$  such that (3.11) holds, and if  $M$  is connected, then  $f$  is uniquely determined up to an additive constant.

Conversely, if  $f$  is a smooth function on  $M$ , then the fact that for every  $m \in M$  the linear mapping

$$\sigma_m : v \mapsto i_v \sigma_m : T_m M \rightarrow (T_m M)^*$$

is bijective shows that there is a unique vector field  $v$  on  $M$  which satisfies (3.11). Moreover,  $v$  is smooth because  $df$  and  $m \mapsto \sigma_m^{-1}$  are smooth. The smooth vector field  $v$  on  $M$  such that (3.11) holds is called the *Hamiltonian vector field*  $H_f$  on  $M$  defined by the function  $f$ . In the literature  $f$  is called the *Hamiltonian function* of the vector field  $v$  and denoted by  $H$ . However, we would like to stress that  $f$  can be any smooth function on  $M$ . It is quite remarkable that there are so many smooth vector fields whose flows leave  $\sigma$  invariant: one for every closed one-form on  $M$ , or smooth function on  $M$  modulo an additive constant.

A system of coordinates  $x_i, \xi_i$  in  $M$  is called a *canonical system of coordinates* if (2.6) holds. Let, in such a coordinate system,  $\dot{x}_i, \dot{\xi}_i$  denote the coordinates of the vector field

$$v = H_f = \left( \dot{x}_1, \dots, \dot{x}_n; \dot{\xi}_1, \dots, \dot{\xi}_n \right).$$

With these notations, the equation (3.11) reads

$$i_v \sigma = \sum_{i=1}^n \left( \dot{\xi}_i dx_i - \dot{x}_i d\xi_i \right) = - \sum_{i=1}^n \left( \frac{\partial f}{\partial x_i} dx_i + \frac{\partial f}{\partial \xi_i} d\xi_i \right),$$

from which obtain that

$$\dot{x}_i = \frac{\partial f(x, \xi)}{\partial \xi_i}, \quad \dot{\xi}_i = -\frac{\partial f(x, \xi)}{\partial x_i}, \quad 1 \leq i \leq n. \tag{3.12}$$

In other words, if  $dm(t)/dt = v(m(t))$  is the differential equation for the flow defined by the vector field  $v = H_f$ , then in canonical local coordinates we arrive at the system of ordinary differential equations (3.12), in which we replace  $\dot{x}_i$  and  $\dot{\xi}_i$  by  $dx_i(t)/dt$  and  $d\xi_i(t)/dt$ , respectively, and in the right hand side take the partial derivatives of  $f$  at  $x_i = x_i(t)$ ,  $\xi_i = \xi_i(t)$ . We recognize the resulting system of ordinary differential equations as the Hamiltonian system defined by the

function  $f$  as it appears in the textbooks in Classical Mechanics, where usually the position and momentum coordinates  $x_i$  and  $\xi_i$  are denoted by  $q_i$  and  $p_i$ , respectively.

Note that if we would have taken the other sign convention in (3.11), then the signs in (3.12) would be opposite to the standard ones in classical mechanics. Also note that the convention, to use Greek letters for the momentum coordinates corresponding to the Latin letters for the position coordinate, allows us to write for example the momentum coordinates corresponding to the position coordinates  $x, y, z$  as  $\xi, \eta, \zeta$ , respectively.

**Example 3.1** If  $X$  is an  $n$ -dimensional smooth manifold and  $v \in \mathcal{X}(X)$ , then the *momentum function* of the vector field  $v$  in the base manifold  $X$  is the function  $\mu_v$  on the cotangent bundle  $M := T^*X$ , defined by

$$\mu_v(x, \xi) = \xi(v(x)), \quad x \in X, \quad \xi \in (T_x X)^*. \quad (3.13)$$

Then the fiber derivative of  $\mu_v$  is equal to

$$\dot{x} = \frac{\partial \mu_v(x, \xi)}{\partial \xi} = v(x) \in T_x X = ((T_x X)^*)^*.$$

It follows that the projection  $\pi : (x, \xi) \rightarrow x : T^*X \rightarrow X$  intertwines the  $H_{\mu_v}$ -flow in  $T^*X$  with the  $v$ -flow in  $X$ , in the sense that

$$\pi \circ e^t H_{\mu_v} = e^{tv} \circ \pi.$$

More precisely, for every  $x \in X$  and  $\xi \in (T_x X)^*$  we have that

$$e^t H_{\mu_v}(x, \xi) = \left( e^{tv}(x), ((T_x e^{tv})^*)^{-1}(\xi) \right).$$

In other words, the flow in  $T^*X$  of the Hamiltonian system defined by the function  $\mu_v$  is equal to the flow in  $T^*X$  which is induced by the flow in  $X$  of the vector field  $v$ .  $\circlearrowright$

### 3.4 The Legendre Transform

Let  $L$  be a smooth real-valued function on an open subset  $U$  of the *tangent* bundle  $TX$  of an  $n$ -dimensional smooth manifold  $X$ , where we will write  $L(x, v) \in \mathbf{R}$  whenever  $x \in X$  and  $v \in T_x X$ . For any smooth curve  $\gamma : [a, b] \rightarrow X$  such that  $(\gamma(t), \gamma'(t)) \in U$  for all  $t \in [a, b]$ , define the integral

$$I(\gamma) = \int_a^b L(\gamma(t), \gamma'(t)) dt. \quad (3.14)$$

The *variational formula of Euler and Lagrange* states that if  $\gamma = \gamma_\epsilon$  depends smoothly on a parameter  $\epsilon$ , then

$$\begin{aligned} \frac{dI(\gamma_\epsilon)}{d\epsilon} &= - \int_a^b [L](t) \delta(t) dt \\ &+ \mu(\gamma(b), \gamma'(b)) \delta(b) - \mu(\gamma(a), \gamma'(a)) \delta(a). \end{aligned} \quad (3.15)$$

Here

$$\delta(t) := \frac{\partial \gamma_\epsilon(t)}{\partial \epsilon} \in T_{\gamma(t)} X \quad (3.16)$$

denotes the "variation with respect to  $\epsilon$ " of the curve  $\gamma_\epsilon(t)$ ,  $[L](t)$  is the linear form on  $T_{\gamma(t)} X$  which in local coordinates is given by

$$[L]_i := \left. \frac{d\mu_i(\gamma(t), \gamma'(t))}{dt} - \frac{\partial L(x, \gamma'(t))}{\partial x_i} \right|_{x=\gamma(t)} \quad (3.17)$$

and the linear form  $\mu(x, v) = \mu_L(x, v)$  on  $T_x X$  is defined by

$$\mu(x, v) := \frac{\partial L(x, v)}{\partial v} \in (T_x X)^*. \quad (3.18)$$

The formula (3.15) is obtained by differentiating with respect to  $\epsilon$  under the integral sign, and then performing an integration by parts on the term with the factor

$$\frac{\partial^2 \gamma_\epsilon(t)}{\partial \epsilon \partial t} = \frac{\partial^2 \gamma_\epsilon(t)}{\partial t \partial \epsilon}.$$

The linear form  $\mu(x, v)$  on the tangent space in (3.18), which is defined in a coordinate-independent way, is called the *momentum vector* assigned to the velocity vector  $v$  by means of the function  $L$ .

It is one of the basic observations of Lagrange [19, Tome I, Partie 2, Section IV] that, although the two summands in (3.17) transform in a quite complicated manner under a change of coordinates, the quantity  $[L](t)$  transforms as a covector, an element of  $(T_{\gamma(t)} X)^*$ . His argument is that for any  $w \in T_{\gamma(t_0)} X$  the real number  $-[L](t_0)w$  is equal to the limit for  $j \rightarrow \infty$  of the left hand side of (3.17), which is independent of any choice of coordinates, if we take  $\gamma_\epsilon^j(t)$  in such a way that  $\delta^j(t) = \partial \gamma_\epsilon^j(t) / \partial \epsilon$  is only nonzero for  $t$  in a shrinking neighborhood of  $t_0$  and is asymptotically equal to a large multiple of  $v$ , in such a way that the integral over  $t$  remains equal to  $w$ . (Those who are familiar with the theory of distributions will recognize  $\delta^j(t)$  as a sequence of smooth functions which approximate the Dirac delta function at the point  $t_0$  times  $w$ , where the approximation is in the distributional sense.)

The *velocity-to-momentum mapping*

$$\Phi = \Phi_L : (x, v) \mapsto (x, \mu(x, v)) : U \rightarrow T^* X$$

is a local diffeomorphism if and only if its tangent mapping is invertible, which is equivalent to *Legendre's condition* that

$$\text{the symmetric bilinear form } \frac{\partial \mu(x, v)}{\partial v} = \frac{\partial^2 L(x, v)}{\partial v^2} \text{ on } T_x X \text{ is nondegenerate.}$$

Here, in linear coordinates in  $T_x X$ , the bilinear form  $\partial^2 L(x, v) / \partial v^2$  has the symmetric matrix  $\partial^2 L(x, v) / \partial v_i \partial v_j$ ,  $1 \leq i, j \leq n$ , the Hessian of the function  $v \mapsto L(x, v)$ . By restricting  $U$  to open subsets on which  $\Phi$  is injective, we obtain a diffeomorphism from  $U$  onto an open subset  $V$  of the cotangent bundle  $T^* X$  of  $X$ .

A curve  $\gamma(t)$  is called a *stationary curve* for the integral  $I$  in (3.14), if it satisfies the *Euler-Lagrange equations*

$$[L](t) \equiv 0, \quad (3.19)$$

i.e. if the integral in (3.15) vanishes for any variation  $\delta(t)$  of  $\gamma(t)$ . If the Legendre condition holds, then the Euler-Lagrange equations can be written in local coordinates as a second order system of ordinary differential equations

$$\frac{d^2 \gamma_i(t)}{dt^2} = a_i(\gamma(t), \gamma'(t)), \quad 1 \leq i \leq n,$$

in which the components of the acceleration  $a_i(x, v)$  are smooth functions of  $x$  and  $v$ . Actually, it is more convenient to view this second order system as a first order system

$$x'(t) = v(t), \quad v'(t) = a(x(t), v(t)),$$

defined by the vector field  $(v, a(x, v))$  in the tangent bundle.

Now define the functions  $H$  on  $U$  and  $h$  on  $V$  subsequently by means of the equations

$$H(x, v) := \langle v, \mu(x, v) \rangle - L(x, v), \quad h = H \circ \Phi^{-1}. \quad (3.20)$$

Here  $\langle v, \xi \rangle$  is the customary, more symmetric notation for the value  $\xi(v)$  which the linear form  $\xi$  takes on the vector  $v$ . The function  $h$  on the open subset  $V$  of the cotangent bundle  $T^*X$  of  $X$  is called the *Legendre transform* of the function  $L$ .

Write  $v = v(x, \xi)$  for the solution  $v$  of the equation  $\mu(x, v) = \xi$ . Then

$$h(x, \xi) = \langle v(x, \xi), \xi \rangle - L(x, v(x, \xi)),$$

hence

$$\frac{\partial h(x, \xi)}{\partial \xi_i} = \left\langle \frac{\partial v(x, \xi)}{\partial \xi_i}, \xi \right\rangle + v_i(x, \xi) - \left\langle \frac{\partial v(x, \xi)}{\partial \xi_i}, \mu(x, v(x, \xi)) \right\rangle = v_i(x, \xi),$$

where in the first and second identity we have used the definition (3.18) of  $\mu$  and the equation  $\mu(x, v(x, \xi)) = \xi$ , respectively. Similarly we have

$$\frac{\partial h(x, \xi)}{\partial x_i} = \left\langle \frac{\partial v(x, \xi)}{\partial x_i}, \xi \right\rangle - \frac{\partial L(x, v)}{\partial x_i} \Big|_{v=v(x, \xi)} - \left\langle \frac{\partial v(x, \xi)}{\partial x_i}, \mu(x, v(x, \xi)) \right\rangle = - \frac{\partial L(x, v)}{\partial x_i} \Big|_{v=v(x, \xi)},$$

where again in the first and second identity we have used the definition (3.18) of  $\mu$  and the equation  $\mu(x, v) = \xi$ , respectively.

The Euler-Lagrange equations are

$$\frac{dx(t)}{dt} = v(t), \quad \frac{d\xi_i(t)}{dt} = \frac{\partial L(x, v)}{\partial x_i} \Big|_{v=v(x, \xi)},$$

where in the second equation we have substituted  $\xi = \mu(x, v) = \partial L(x, v)/\partial v$ . The point of the computation of the partial derivatives of the function  $h(x, \xi)$  is that the velocity-to-momentum mapping  $\Phi$  transforms the Euler-Lagrange equations into the Hamiltonian system

$$\frac{dx_i}{dt} = \frac{\partial h(x, \xi)}{\partial \xi_i}, \quad \frac{d\xi_i}{dt} = - \frac{\partial h(x, \xi)}{\partial x_i}$$

on  $V$  which is defined by the function  $h$ .

Conversely, if  $h$  is a smooth real-valued function on an open subset  $V$  of  $T^*X$  for which the *momentum-to-velocity mapping*

$$\Psi = \Psi_h : (x, \xi) \mapsto (x, v(x, \xi)), \quad v(x, \xi) := \frac{\partial h(x, \xi)}{\partial \xi}$$

is a diffeomorphism from  $V$  onto an open subset  $U$  of  $TX$ , then we can define subsequently

$$l(x, \xi) := \langle v(x, \xi), \xi \rangle - h(x, \xi), \quad L = l \circ \Psi^{-1}.$$



It is then not hard to verify that  $\Psi_h = \Phi_L^{-1}$  and  $h$  is equal to the Legendre transform of  $L$ , which shows that  $\Psi_L$  transforms the Hamiltonian system defined by the function  $h$  into the Euler-Lagrange equations for the function  $L$ .

We now turn to the relation with Classical Mechanics. Lagrange [19, Tome 1, partie 2, Section IV] actually made his observation, that  $[L]$  transforms under changes of local coordinates as a covector, in the case that  $L(x, v)$  is equal to the *kinetic energy*

$$T(x, v) = \frac{1}{2}m(x)(v, v)$$

of a classical mechanical system. Here  $m(x)$ , the *inertial mass tensor* is an inner product on  $T_x X$ . If we have local coordinates in which  $m(x) = m$  does not depend on  $x$ , then  $[T] = d(mv)/dt$  and we recognize the equation

$$[T] = F = \text{the force acting on the system} \tag{3.21}$$

as *Newton's equations of motion*. However, Lagrange observed that under arbitrary nonlinear changes of coordinates, such as the passage from rectangular coordinates to polar coordinates, the acceleration  $a = dv/dt$  transforms in a complicated way, not at all as a tensor, and the transformed equations of motion do not look like Newton's equations  $F = ma$  at all. (Also the inertial mass tensor in this case no longer is independent of the position.)

Because Lagrange had understood that in general a quantity of the form  $[L]$  transforms covariantly, he proposed to formulate the equations of motion for a general classical mechanical system as (3.21), in which  $T$  is the kinetic energy function viewed as a smooth function on the tangent bundle. Moreover, the force field  $F$  is (has to be equal to) a smooth mapping which assigns to each  $x \in X$  and  $v \in T_x X$  an element of  $(T_x X)^*$ , a linear form on  $T_x X$ .

The force field  $F$  is called *conservative*, if  $F(x, v) = F(x)$  does not depend on  $v$  and  $F(x) = -dV(x)$  for a *potential energy function*  $V$  which is a smooth real-valued function on  $X$ . In this case one has  $F = [V]$  and we recognize the equations of motion (3.21) as the Euler-Lagrange equations  $[L] = 0$ , in which  $L = T - V$ .

The momentum defined by  $L = T - V$  is equal to  $\mu(x, v) = m(x)v$ , in which the inertial mass tensor  $m(x)$  is regarded as a bijective linear mapping from  $T_x X$  onto  $(T_x X)^*$ . Using that  $v \mapsto T(x, v)$  is homogeneous of degree two, one obtains that its Legendre transform is equal to  $T \circ \Phi^{-1}$ , whereas the Legendre transform of  $V$  is equal to  $-V$ . This yields that the Legendre transform of  $L = T - V$  is equal to the *total energy function*  $h = T \circ \Phi^{-1} + V$ , viewed as a function of the positions and the momenta. In this way the equations of motion for a classical mechanical system with a conservative force field are equivalent to the Hamiltonian system defined by the total energy function, viewed as a function on the cotangent bundle  $T^*X$  rather than on the tangent bundle  $TX$ .

**Remark 3.1** Lagrange [19, Tome 1, Partie 2, Section V] explicitly introduced the velocity-to-momentum mapping  $\Phi = \Phi_L$  and the two-form  $\Phi_L^* \sigma$  on the tangent bundle  $TX$ , but without mentioning the canonical two-form  $\sigma$  of the cotangent bundle  $T^*X$ . He also proved that  $\Phi_L^* \sigma$  is invariant under the flow defined by the Euler-Lagrange equations. His proof is paraphrased in Exercise 3.4

The equivalence between the Euler-Lagrange equations of variational calculus and Hamiltonian systems has been found for  $L = T - V$  by Hamilton [11], and then it was soon realized by many

authors that his proof holds for an arbitrary function  $L$  on the tangent bundle which satisfies Legendre's condition. Earlier, the perturbation equations of a classical mechanical system in which the potential energy is perturbed were written in a Hamiltonian form by Lagrange [19, Tome I, p. 310]. He might have missed the general equivalence between Euler-Lagrange equations and Hamiltonian systems, because he probably was not aware of the Legendre transform.  $\circledast$

### 3.5 Poisson Brackets

If  $f, g \in \mathcal{F}(M)$ , then the *Poisson brackets*  $\{f, g\}$  of  $f$  and  $g$  are defined as the derivative of the function  $g$  in the direction of the Hamiltonian vector field  $H_f$  defined by the function  $f$ :

$$\{f, g\} := \mathcal{L}_{H_f}(g) = i_{H_f}(dg) = -i_{H_f}(i_{H_g}\sigma) = \sigma(H_f, H_g). \quad (3.22)$$

Here we used (3.11) and the antisymmetry of  $\sigma$  in the third and fourth identity, respectively. The right hand side shows that the Poisson brackets are antisymmetric in the sense that

$$\{g, f\} = -\{f, g\}, \quad f, g \in \mathcal{F}(M). \quad (3.23)$$

In canonical local coordinates the Poisson brackets are given by

$$\{f, g\}(x, \xi) = \sum_{i=1}^n \left( \frac{\partial f(x, \xi)}{\partial \xi_i} \frac{\partial g(x, \xi)}{\partial x_i} - \frac{\partial f(x, \xi)}{\partial x_i} \frac{\partial g(x, \xi)}{\partial \xi_i} \right). \quad (3.24)$$

It follows immediately that the following conditions a)–d) are equivalent:

- a)  $g$  is a constant of motion for the Hamiltonian system defined by the function  $f$ , in the sense that  $g$  is invariant under the  $H_f$ -flow.
- b)  $\{f, g\} = 0$ .
- c)  $\{g, f\} = 0$ .
- d)  $f$  is a constant of motion for the Hamiltonian system defined by the function  $g$ .

In the applications, one often has that the  $H_g$ -flow is a one-parameter group of symmetry for the function  $f$ , which means that d) holds. The conclusion, that in this case  $g$  is a constant of motion for the Hamiltonian system defined by the function  $f$ , is called *Noether's principle for Hamiltonian systems*. In many examples we have  $M = T^*X$  and  $g = \mu_v$ , the momentum function of a smooth vector field  $v$  in the base manifold  $X$ , cf. Example 3.1.

It also follows from (3.23) that  $\{f, f\} = 0$ , meaning that the function  $f$  is a constant of motion for the Hamiltonian system defined by  $f$ . If  $f$  is equal to the total energy of a classical mechanical system as in Subsection 3.4, then this is the *law of conservation of the total energy*.

The derivative of  $\{f, g\}$  is equal to

$$d\{f, g\} = d\mathcal{L}_{H_f}g = \mathcal{L}_{H_f}dg = -\mathcal{L}_{H_f}(i_{H_g}\sigma) = -i_{[H_f, H_g]}\sigma.$$

Here we used in the first, second, third and fourth identity the definition (3.22) of the Poisson brackets, the fact that exterior differentiation commutes with pullbacks and therefore with Lie

derivatives, the definition (3.11) of Hamiltonian vector fields, and formula (3.7) together with the fact that  $\mathcal{L}_{H_f}\sigma = 0$ .

This formula for the derivative of  $\{f, g\}$  just means that

$$[H_f, H_g] = H_{\{f, g\}}. \quad (3.25)$$

In words, the Lie brackets of the Hamiltonian vector fields of the functions  $f$  and  $g$  is again a Hamiltonian vector field, namely of the Poisson brackets of  $f$  and  $g$ .

If we now let act the left and right hand side of (3.25) on a third smooth function  $h$  on  $M$ , then we obtain, using (3.10), that

$$\{f, \{g, h\}\} - \{g, \{f, h\}\} = \{\{f, g\}, h\}.$$

Using the antisymmetry (3.23) at several places, this identity can be rewritten as the *Jacobi identity for Poisson brackets*:

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0. \quad (3.26)$$

(Note the cyclic permutation of  $f$ ,  $g$  and  $h$  in the left hand side of (3.26). Together, (3.26) and (3.25) mean that

**Theorem 3.2** *The space of smooth functions  $\mathcal{F}(M)$  on  $M$  is a Lie algebra with respect to the Poisson brackets, and the mapping which assigns to a smooth function its Hamiltonian vector field is a homomorphism of Lie algebras from  $\mathcal{F}(M)$  to the Lie algebra  $\mathcal{X}(M)$  of smooth vector fields on  $M$ . The kernel of this homomorphism is equal to the space of function which are constant on the connected components of  $M$ .*

**Remark 3.2** The Jacobi identity for the Poisson brackets (3.26) (in canonical coordinates) goes back to the article of Jacobi [18], which appeared posthumously in 1862. Jacobi mentioned that (3.26) implies the earlier *theorem of Poisson*, which states that if  $g$  and  $h$  are constants of motion for the Hamiltonian system defined by the function  $f$ , then  $\{g, h\}$  also is a constant of motion for the Hamiltonian system defined by  $f$ . It could very well be that Jacobi was led to (3.26) by means of an analysis of Poisson's proof. This observation of Jacobi may also have led to adoption of the name "Poisson brackets", which brackets appeared earlier in the work of Lagrange [19, Tome I, p. 315].

Inspired by the Jacobi identity for Poisson brackets, Lie [21, Vol. 1, Kap. 5, §26 and Vol. 2, Kap. 7, §44, 45] introduced the Jacobi identity for vector fields, and coined the name "Jacobi identity". ⊙

### 3.6 Darboux's Lemma

Let  $(M, \sigma)$  be a symplectic manifold. The *Darboux lemma* states that locally  $\sigma$  form can be brought into the canonical form (2.6). This means that for every  $m_0 \in M$  there exists an open neighborhood  $U$  of  $m_0$  in  $M$  and a diffeomorphism  $\Phi$  from  $U$  onto an open subset  $V$  of  $\mathbf{R}^n \times \mathbf{R}^n$  such that

$$\Phi^* \left( \sum_{i=1}^n d\xi_i \wedge dx_i \right) = \sigma \quad \text{on } U.$$

A proof can be given as follows. Let  $\phi$  be any smooth function defined in a neighborhood of  $m_0$  such that  $\phi(m_0) = 0$  and  $d\phi_{m_0} \neq 0$ . Choose any codimension one smooth submanifold  $S$  of  $M$  through  $m_0$  such that  $H_\phi(m_0) \notin T_{m_0} S$ . Then there is a locally unique smooth function  $f$  on a neighborhood of  $m_0$  such that  $\{\phi, f\} = \mathcal{L}_{H_\phi} f = 1$  and  $f = 0$  on  $S$ . It follows from (3.22) that the symplectic product of  $H_\phi$  and  $H_f$  is equal to one, which implies that at every points these vectors are linearly independent, which in turn implies that at every point  $d\phi$  and  $df$  are linearly independent. This implies in particular that

$$N := \{m \in U \mid f(m) = \phi(m) = 0\}$$

is a smooth codimension 2 submanifold of the open neighborhood  $U$  of  $m_0$  on which  $\phi$  and  $f$  are defined. Moreover, the restriction  $\sigma_N$  of  $\sigma$  to  $N$  is a symplectic form, because the symplectic orthogonal complement of  $T_n N$  is spanned by  $H_\phi(n)$  and  $H_f(n)$ , which space is complementary to the  $T_n N = \ker(d\phi_n) \cap \ker(df_n)$ , as is readily verified.

It follows from (3.25) and the fact that the Hamiltonian vector field of any constant function is equal to zero, that the vector fields  $H_\phi$  and  $H_f$  commute, which implies that their flows commute as well. Now define, for  $n \in N$  and  $t, \tau \in \mathbf{R}$ ,

$$\Phi(n, t, \tau) := e^t H_\phi \circ e^{-\tau} H_f(n) = e^{-\tau} H_f \circ e^t H_\phi(n).$$

Then  $\partial\Phi(n, t, \tau)/\partial t = H_\phi(\Phi(n, t, \tau))$  and  $\partial\Phi(n, t, \tau)/\partial\tau = -H_f(\Phi(n, t, \tau))$ . It follows that the value  $s$  which  $(\Phi^*\sigma)_{(n, t, \tau)}$  takes on the pair of vectors  $(\delta n, \delta t, \delta\tau)$  and  $(\delta n', \delta t', \delta\tau')$  is equal to  $\sigma_{\Phi(n, t, \tau)}(v, v')$ , in which

$$v = T_n \left( e^t H_\phi \circ e^{-\tau} H_f \right) \delta n + \delta t H_\phi - \delta\tau H_f,$$

and  $v'$  is given by the same formule with  $\delta n, \delta t, \delta\tau$  replaced by  $\delta n', \delta t', \delta\tau'$ , respectively. Using again that  $\sigma(H_\phi, H_f) = 1$  and that Hamiltonian flows preserve the symplectic form, we arrive at the conclusion that

$$s = \sigma_n(\delta n, \delta n') + \delta\tau \delta t' - \delta\tau' \delta t.$$

This shows that  $\Phi^*\sigma$  is equal to the direct sum of  $\sigma_N$  and the standard symplectic form in  $\mathbf{R}^2$ . The proof of Darboux's lemma now follows by induction on  $n$ .

**Remark 3.3** Weinstein [27] gave a proof of the Darboux lemma which is based on a deformation argument which has been introduced in normal form theory by Moser [24]. The proof given above is closer to the one given by Darboux. Both proofs have their merits.  $\circlearrowright$

If we combine the Darboux lemma with the reduction in Subsection 2.3, then one obtains that any closed two-form of constant rank has a local normal form.

### 3.7 Hamiltonian Group Actions

If a Lie group  $G$  acts on the smooth manifold  $M$ , we will denote for each  $g \in G$  the diffeomorphism  $m \mapsto gm$  of  $M$  by  $g_M$ . For each element  $X$  in the Lie algebra  $\mathfrak{g}$  of  $G$  we have the infinitesimal action

$$X_M := \left. \frac{d}{dt} (\exp(tX))_M \right|_{t=0}$$

of  $X$  on  $M$ , which is a smooth vector field on  $M$ . Note that, as a consequence,  $(\exp(tX))_M = e^{tX_M}$  for every  $t \in \mathbf{R}$ .

The definition (3.8) of the Lie brackets of vector fields implies that

$$[X, Y]_M = -[X_M, Y_M], \quad X, Y \in \mathfrak{g}, \quad (3.27)$$

i.e. the mapping  $X \mapsto X_M$  is an *anti-homomorphism* from the Lie algebra  $\mathfrak{g}$  to the Lie algebra  $\mathcal{X}(M)$ .

Now suppose that  $\sigma$  is a symplectic form on  $M$ . The action of  $G$  on  $M$  will be called *Hamiltonian* with respect to  $\sigma$ , if for every  $X \in \mathfrak{g}$  we have given a smooth function  $\langle X, \mu \rangle$  on  $M$  such that

$$X_M = H_{\langle X, \mu \rangle}. \quad (3.28)$$

It will furthermore be required that  $\langle X, \mu \rangle$  depends linearly on  $X \in \mathfrak{g}$  and that

$$\{\langle X, \mu \rangle, \langle Y, \mu \rangle\} = -\langle [X, Y], \mu \rangle, \quad X, Y \in \mathfrak{g}. \quad (3.29)$$

In other words, we require that  $X \mapsto \langle X, \mu \rangle$  is an anti-homomorphism of Lie algebras from  $\mathfrak{g}$  to the Poisson Lie algebra  $\mathcal{F}(M)$  of all smooth functions on  $M$ .

The condition that the infinitesimal actions are Hamiltonian implies that the one-parameter subgroups preserve the symplectic form. Therefore, if  $G$  is connected, it follows that the  $G$ -action leaves the symplectic form invariant. In other words, the action is a homomorphism from  $G$  to the group of canonical transformations in  $(M, \sigma)$ .

For every  $m \in M$ ,  $\mu(m) : X \mapsto \langle X, \mu \rangle(m)$  is a linear form on  $\mathfrak{g}$ , this defines a smooth mapping  $\mu : M \rightarrow \mathfrak{g}^*$  which is called the *momentum mapping* of the Hamiltonian action of  $G$  on  $M$ . Note that the notation has been arranged such that  $\langle X, \mu(m) \rangle = \langle X, \mu \rangle(m)$  for every  $m \in M$ .

On  $\mathfrak{g}$  we have the *adjoint action*  $(g, X) \mapsto (\text{Ad } g)(X)$  of  $G$ , and transposition leads to the action

$$(g, \xi) \mapsto ((\text{Ad } g)^*)^{-1}(\xi)$$

on the dual  $\mathfrak{g}^*$  of the Lie algebra, which is called the *co-adjoint action* of  $G$  on  $\mathfrak{g}^*$ . The infinitesimal co-adjoint action of  $X \in \mathfrak{g}$  is given by the linear mapping

$$X_{\mathfrak{g}^*} = -(\text{ad } X)^* : \mathfrak{g}^* \rightarrow \mathfrak{g}^*,$$

or, more explicitly,

$$\langle Y, X_{\mathfrak{g}^*} \xi \rangle = -\langle [X, Y], \xi \rangle, \quad \xi \in \mathfrak{g}^*, \quad X, Y \in \mathfrak{g}. \quad (3.30)$$

If in (3.30) we substitute  $\xi = \mu(m)$  and combine the resulting equation with (3.29), and use that the left hand side of (3.29) is equal to  $\mathcal{L}_{X_M} \langle Y, \mu \rangle$ , we arrive at the conclusion that

$$\mathcal{L}_{X_M} \mu = X_{\mathfrak{g}^*} \mu, \quad (3.31)$$

which means that the momentum mapping  $\mu : M \rightarrow \mathfrak{g}^*$  intertwines the infinitesimal action of  $\mathfrak{g}$  on  $M$  with the infinitesimal co-adjoint action of  $\mathfrak{g}$  on  $\mathfrak{g}^*$ . If  $G$  is connected, then this implies in turn that the momentum mapping intertwines the action of  $G$  on  $M$  with the co-adjoint action of  $G$  on  $\mathfrak{g}^*$ , in the sense that

$$g_M^* \mu = ((\text{Ad } g)^*)^{-1} \mu, \quad g \in G. \quad (3.32)$$

**Example 3.3** A very simple, but important example of a Hamiltonian group action on  $(M, \sigma)$  can be obtained as follows. Assume that  $f_i$ ,  $1 \leq i \leq k$ , are smooth functions on  $M$  which *Poisson commute*, i.e.  $\{f_i, f_j\} = 0$  for all  $1 \leq i, j \leq k$ . Let  $f : M \rightarrow \mathbf{R}^k$  be the mapping which has the  $f_i$  as components. Then for every  $c \in \mathbf{R}^k$  the level set  $M_c := \{m \in M \mid f(m) = c\}$  is invariant under the flows of the Hamiltonian vector fields  $H_{f_i}$ . Let us assume that the Hamiltonian vector fields  $H_{f_i}$  are complete, which condition in view of the above is certainly satisfied if the level set  $M_c$  is compact.

In view of (3.25), the fact that the functions  $f_i$  Poisson commute implies that the Hamiltonian vector fields  $H_{f_i}$  commute, which in turn implies that their flows commute. This implies that

$$((t_1, \dots, t_k), m) \mapsto e^{t_k H_{f_k}} \circ \dots \circ e^{t_1 H_{f_1}}(m)$$

defines an action of the additive group  $(\mathbf{R}^k, +)$  on  $M$ . This action is Hamiltonian, with  $f$  as its momentum mapping. Because the Lie algebra  $\mathbf{R}^k$  is commutative, the adjoint action is trivial, hence the co-adjoint action is trivial as well and the fact that  $f$  intertwines the action on  $M$  with the co-adjoint action reproduces the observation that the functions  $f_i$  are invariant under the action on  $M$ .

The system is called *integrable* if  $k = n$  and the mapping  $f$  has regular values. If  $c$  is a regular value of  $f$ , then  $M_c$  is an  $n$ -dimensional smooth submanifold of  $M$  on which the action is locally transitive, and therefore is transitive on each connected component  $C$  of  $M_c$ . If we choose  $m \in C$ , then the *period lattice*  $P_C$  in  $\mathbf{R}^n$  is defined as the set of all  $T \in \mathbf{R}^n$  such that  $T_M(m) = m$ .  $P_C$  does not depend on the choice of  $m \in C$  (but in general it depends sensitively on the level  $c$ ), and  $P_C$  is a discrete subgroup of  $\mathbf{R}^n$ . The mapping  $t \mapsto t_M(m)$  induces a diffeomorphism from  $\mathbf{R}^n/P_C$  onto  $C$ , which intertwines the translational action of  $\mathbf{R}^n$  on  $\mathbf{R}^n/P_C$  with the action of  $\mathbf{R}^n$  on  $C$ .

If  $C$  is compact, which certainly is the case if  $M_c$  is compact, then  $\mathbf{R}^n/P_C$  is compact, hence a *torus*, and the flow of each of the Hamiltonian vector fields  $H_{f_i}$  is quasi-periodic, meaning that by means of a diffeomorphism it can be mapped to a constant speed motion on a standard torus.  $\odot$

### 3.8 Poisson Structures

It follows from the third expression in (3.22) and from (3.11) that

$$\{f, g\}(m) = -\sigma_m^{-1}(df(m), dg(m)),$$

in which  $\pi_m := \sigma_m^{-1} : (\mathbb{T}_m M)^* \rightarrow \mathbb{T}_m M$  is regarded as an antisymmetric bilinear form on  $(\mathbb{T}_m M)^*$ , or as an element of  $\Lambda^2 \mathbb{T}_m M$ , which is also called a *two-vector* in  $\mathbb{T}_m M$ .

For any smooth manifold  $M$ , a *Poisson structure on  $M$*  is defined as a smooth two-vector field  $\pi_m \in \Lambda^2 \mathbb{T}_m M$ ,  $m \in M$ , in such a way that the corresponding *Poisson brackets*  $\{f, g\}$ , defined by

$$\{f, g\}(m) = \pi_m(df(m), dg(m)), \quad m \in M, \quad (3.33)$$

satisfy the Jacobi identity (3.26).

Viewing  $\pi_m$  as a linear mapping from  $(\mathbb{T}_m M)^*$  to  $\mathbb{T}_m M$ , we can define the Hamiltonian vector field  $H_f$  of the function  $f$  by  $H_f(m) = \pi_m df(m)$ . With this convention,  $\{f, g\} = \mathcal{L}_{H_f} g$ .

If  $\pi_m$  is surjective, then it is bijective, and we have that  $\pi_m = -\sigma_m^{-1}$  for a symplectic form on  $M$ . We therefore only get really new examples of Poisson structures if  $\pi_m$  is not surjective.

Write  $H_m = \pi_m((T_m M)^*)$ . If the rank of  $\pi_m$ , the dimension of  $H_m$ , is constant as a function of  $m \in M$ , then the  $H_m$ ,  $m \in M$ , define a smooth vector subbundle  $H$  of  $TM$ , and it follows from the Jacobi identity for the Poisson brackets that  $H$  is integrable. Furthermore, for each integral manifold  $I$  of  $H$ , the restriction of  $\{f, g\}$  to  $I$  only depends on  $f|_I$  and  $g|_I$ , and we obtain a Poisson structure on  $I$ , which turns out to be defined by a symplectic structure on  $I$ . In this way the Poisson manifold  $(M, \pi)$  can be characterized as a manifold which is foliated by symplectic leaves, where the Poisson brackets are defined by taking the Poisson brackets of the restrictions of the functions to the symplectic leaves.

A prime example is the Poisson structure in  $\mathfrak{g}^*$  which is defined by

$$\pi_\xi(X, Y) = -\langle [X, Y], \xi \rangle, \quad \xi \in \mathfrak{g}^*, \quad X, Y \in (\mathfrak{g}^*)^* = \mathfrak{g}. \quad (3.34)$$

The symplectic leaves are the co-adjoint orbits in  $\mathfrak{g}^*$ .

The formula (3.29) shows that, for a Hamiltonian action of  $G$  on the symplectic manifold  $(M, \sigma)$ , the momentum mapping  $\mu$  intertwines the Poisson structure on  $(M, \sigma)$  with the Poisson structure on the dual of the Lie algebra of  $G$ .

**Remark 3.4** The general concept of Poisson structures has been invented by Lichnerowicz [20].

However, Lie [21, Vol. 2, Kap. 8] introduced a "function group" as a fibration  $\phi$  of a symplectic manifold  $(M, \sigma)$  over a manifold  $N$  with the property that for every smooth pair of functions  $f$  and  $g$  on  $N$  the Poisson brackets  $\{f \circ \phi, g \circ \phi\}$  are constant along the fibers of  $\phi$ . This means that there is a unique Poisson structure on  $N$  such that  $\{f \circ \phi, g \circ \phi\} = \{f, g\}_N \circ \phi$  for every  $f, g \in \mathcal{F}(N)$ .

Moreover, in [21, Vol. 2, Kap. 19], Lie discussed the dual  $\mathfrak{g}^*$  of the Lie algebra of a Lie group  $G$ , and showed that the projection from  $T^*G$  onto  $\mathfrak{g}^*$  by means of the left trivialization  $T^*G = G \times \mathfrak{g}^*$  is a "function group". The Poisson structure on  $\mathfrak{g}^*$  defined by this "function group" is equal to the one in (3.34).  $\circlearrowright$

### 3.9 Exercises

**Exercise 3.1** Prove that

$$\mathcal{L}_{[u, v]}\omega = [\mathcal{L}_u, \mathcal{L}_v]\omega$$

for every  $u, v \in \mathcal{X}(M)$ ,  $\omega \in \Omega^p(M)$ ,  $p \in \mathbf{Z}_{\geq 0}$ .  $\circlearrowright$

**Exercise 3.2** Prove that in canonical coordinates the vector field  $H_f$  is linear, if and only if  $f$  is equal to a quadratic form plus a constant. Prove that if  $f$  and  $g$  are quadratic forms, then  $\{f, g\}$  is a quadratic form and we have the identity

$$H_{\{f, g\}} = H_g \circ H_f - H_f \circ H_g$$

between  $2n \times 2n$ -matrices = linear mappings from  $\mathbf{R}^{2n}$  to  $\mathbf{R}^{2n}$ . Hint: verify first that we have  $H_{\{f, g\}} = H_f \circ H_g - H_g \circ H_f$  if we view the vector fields as derivations.  $\circlearrowright$

**Exercise 3.3** Let  $p \in M$  and  $v \in \mathcal{X}(M)$ .  $p$  is called an *equilibrium point* of  $v$  if it is a fixed point for the  $v$ -flows  $e^{tv}$ ,  $t \in \mathbf{R}$ . Now let  $v = H_f$  for a smooth function  $f$  on  $M$ . Prove that the following statements are equivalent:

i)  $p$  is an equilibrium point of  $v = \mathbf{H}_f$ .

ii)  $v(p) = 0$ .

iii)  $p$  is a stationary point of the function  $f$  in the sense that  $df(p) = 0$ .

In the sequel assume that  $p$  is an equilibrium point of  $v = \mathbf{H}_f$ . Write  $A(t) := \mathbf{T}_p(e^{tv})$ , which is a linear mapping from  $E := \mathbf{T}_p M$  to itself. Write  $B := A'(0)$ , which also is a linear mapping from  $E$  to itself. Prove that  $A(t) = e^{tB}$ ,  $t \in \mathbf{R}$ , that  $A(t) \in \text{Sp}(E, \sigma_p)$ , and that  $B$  is an infinitesimally symplectic matrix.

Prove that, if we view  $B$  as a (linear) vector field on  $\mathbf{T}_p M$ , then  $B = \mathbf{H}_{f_2}$ , in which  $f_2$  denotes the quadratic term in the Taylor expansion of  $f$  at the point  $p$ , where the Taylor expansion is written down in any suitable system of local coordinates. Prove that  $f_2$  is independent of the choice of the system of local coordinates.  $\circlearrowright$

**Exercise 3.4** In (3.15), let  $\epsilon$  run over a finite-dimensional smooth manifold  $E$ . Define  $\Gamma_t : E \rightarrow \text{TX}$  by  $\Gamma_t(\epsilon) = (\gamma_\epsilon(t), \gamma'_\epsilon(t))$ , and define  $i : E \rightarrow \mathbf{R}$  by  $i(\epsilon) = I(\gamma_\epsilon)$ . Assume furthermore that, for every  $\epsilon \in E$ ,  $\gamma_\epsilon$  is a solution of the Euler-Lagrange equations  $[L] = 0$ .

Let  $\Phi_L : \text{TX} \rightarrow \text{T}^* X$  be the velocity-to-momentum mapping defined by  $L$ . Prove that

$$di = \Gamma_b^* \Phi_L^* \tau - \Gamma_a^* \Phi_L^* \tau,$$

in which  $\tau$  is the tautological one-form on  $\text{T}^* X$ . Prove that

$$\Gamma_b^* \Phi_L^* \sigma = \Gamma_a^* \Phi_L^* \sigma.$$

Now let, for every  $(x, v) \in \text{TX}$ ,  $t \mapsto \gamma_{(x,v)}(t)$  denote the solution  $\gamma$  of the Euler-Lagrange equations such that  $\gamma(0) = x$  and  $\gamma'(0) = v$ . Apply the previous equation to this family of curves in order to prove that  $\Phi_L^* \sigma$  is invariant under the Euler-Lagrange flow in  $\text{TX}$ .  $\circlearrowright$

**Exercise 3.5** With the notation of (3.13), prove that  $\{\mu_v, \mu_w\} = \mu_{[v,w]}$ .  $\circlearrowright$

**Exercise 3.6** Let  $(M, \sigma)$  be a  $2n$ -dimensional symplectic manifold and let  $\Phi : M \rightarrow \mathbf{R}^{2n}$  be a smooth mapping with coordinate functions  $x_1, \dots, x_n, \xi_1, \dots, \xi_n$ . Prove that

$$\Phi^* \left( \sum_{i=1}^n d\xi_i \wedge dx_i \right) = \sigma,$$

if and only if  $\{x_i, x_j\} = 0$ ,  $\{\xi_i, x_j\} = \delta_{ij}$ , and  $\{\xi_i, \xi_j\} = 0$ .  $\circlearrowright$

**Exercise 3.7** Let  $\beta$  be a closed two-form on the configuration space  $X$ , and let the force field  $F$  be given by  $F(x, v) = -dV(x) - \beta_x(v)$  for every  $x \in X$ ,  $v \in \mathbf{T}_x X$ . Here we identify  $\beta_x$  in the usual way with a linear mapping from  $\mathbf{T}_x X$  to  $(\mathbf{T}_x X)^*$ . The term  $-\beta_x(v)$  is called a *magnetic term* in the force field.

Prove that  $\sigma + \pi^* \beta$  is a symplectic form on  $\text{T}^* X$ . Prove that the velocity-to-momentum mapping  $\Phi_T : (x, v) \mapsto (x, \frac{\partial T(x,v)}{\partial v})$  transforms the equations of motion  $[T] = F$  into the Hamiltonian system defined by the function  $h = (T + V) \circ \Phi_T^{-1}$ , not with respect to the canonical symplectic form  $\sigma$  of  $\text{T}^* X$ , but with respect to the symplectic form  $\sigma + \pi^* \beta$ , the canonical symplectic form "shifted by the magnetic term".  $\circlearrowright$



## 4 Hamilton-Jacobi Theory

### 4.1 Lagrange Manifolds

Consider a first order partial differential equation

$$f(x, d\phi(x)) = 0, \quad (4.1)$$

in which the unknown function  $\phi$  is a smooth function defined on an open subset  $U$  of an  $n$ -dimensional smooth manifold  $X$  and  $f$  is a given smooth function on an open subset  $V$  of the cotangent bundle  $T^*X$  of  $X$ .

If we write

$$N := \{(x, \xi) \in V \mid f(x, \xi) = 0\} \quad (4.2)$$

for the zeroset of  $f$  in  $T^*X$  and

$$\Lambda = \{(x, d\phi(x)) \in T^*X \mid x \in U\} \quad (4.3)$$

for the graph of  $d\phi$  viewed as a subset of  $T^*X$ , then (4.1) is equivalent to the inclusion

$$\Lambda \subset N \quad (4.4)$$

between subsets of  $T^*X$ .

The mapping  $d\phi : x \mapsto (x, d\phi(x))$  is a smooth mapping from  $U$  to  $T^*X$ , with image equal to  $\Lambda$  and with the canonical projection  $\pi : T^*X \rightarrow X$  as a left inverse. Therefore  $d\phi$  is a smooth embedding and  $\Lambda$  is a smooth  $n$ -dimensional submanifold of  $T^*X$ . The fact that, for each  $\lambda \in \Lambda$ , the restriction to  $T_\lambda \Lambda$  of  $T_\lambda \pi$  is bijective from  $T_\lambda \Lambda$  to  $T_{\pi(\lambda)} X$  is equivalent to the condition that

$$T_\lambda \Lambda \cap \ker T_\lambda \pi = 0, \quad (4.5)$$

i.e.  $T_\lambda \Lambda$  is complementary to the tangent space  $\ker T_\lambda \pi$  of the fiber through the point  $\lambda$ .

Conversely, if  $\Lambda$  is any smooth  $n$ -dimensional submanifold of  $T^*X$  which satisfies (4.5), then  $\pi|_\Lambda$  is a local diffeomorphism from  $\Lambda$  onto an open subset  $U$  of  $X$ . If moreover  $\pi|_\Lambda$  is injective, which can be arranged by restricting to a suitable open neighborhood of any given point of  $\Lambda$ , then  $\pi|_\Lambda$  is a diffeomorphism, and its inverse  $\alpha := (\pi|_\Lambda)^{-1} : U \rightarrow T^*X$  is a smooth one-form on  $U$ . Locally the condition that  $\alpha = d\phi$  for a smooth function  $\phi$  is equivalent to the condition that  $\alpha$  is closed, i.e.  $d\alpha = 0$

It follows from (2.3) and (2.4) that

$$d\alpha = d(\alpha^* \tau) = \alpha^*(d\tau) = \alpha^* \sigma,$$

and therefore  $\alpha$  is closed if and only if  $\alpha^* \sigma = 0$ . The latter condition means that, for every  $x \in U$ ,

$$\sigma_{\alpha(x)}(T_x \alpha(u), T_x \alpha(v)) = 0, \quad u, v \in T_x X. \quad (4.6)$$

On the other hand

$$T_{(x, \alpha(x))} \Lambda = \{T_x \alpha(v) \mid v \in T_x X\},$$

and therefore (4.6) means that  $T_\lambda \Lambda$  is an isotropic linear subspace of  $T_\lambda(T^*X)$ , if we write  $\lambda = (x, \alpha(x))$ . Because  $\dim T_\lambda \Lambda = n$ , it is a Lagrange plane in  $T_\lambda(T^*X)$ .

A *Lagrange submanifold* of a  $2n$ -dimensional symplectic manifold  $(M, \sigma)$  is defined as an  $n$ -dimensional smooth submanifold  $\Lambda$  of  $M$  such that, for every  $\lambda \in \Lambda$ ,  $T_\lambda \Lambda$  is a Lagrange plane in  $T_\lambda M$ , with respect to the symplectic form  $\sigma_\lambda$ . We have just proved above that a submanifold  $\Lambda$  of  $T^*X$  is equal to  $\alpha(U)$  for a closed one-form on an open subset  $U$  of  $X$ , if and only if

- i)  $\Lambda$  is a Lagrange submanifold of  $T^*X$ ,
- ii)  $\Lambda$  is transversal to the fibers of  $T^*X$  in the sense of (4.5), and
- iii) The restriction of  $\pi$  to  $\Lambda$  is injective.

## 4.2 Lie's View on First Order PDE

It is the idea of Lie, to generalize the concept of a solution of (4.1) slightly, by first investigating what the condition means for a Lagrange submanifold  $\Lambda$  of a symplectic manifold  $(M, \sigma)$  to be contained in the given subset  $N$  of  $M$ . Thereby he dropped the conditions ii) and iii), which relate the position of  $\Lambda$  with respect to the projection  $\pi$ .

In the sequel we will assume that  $df(n) \neq 0$  for every  $n \in N$ . This implies that  $N$  is a smooth  $(2n - 1)$ -dimensional submanifold of  $M$ , and  $T_n N = \ker(df(n))$  for every  $n \in N$ . Furthermore,  $\Lambda \subset N$  implies that, for every  $\lambda \in \Lambda$ ,

$$T_\lambda \Lambda \subset T_\lambda N = \ker(df(\lambda)),$$

which is equivalent to the inclusion

$$\mathbf{R} H_f(\lambda) = \ker(df(\lambda))^{\sigma_\lambda} \subset (T_\lambda \Lambda)^{\sigma_\lambda} = T_\lambda \Lambda. \quad (4.7)$$

of the  $\sigma_\lambda$ -orthogonal complements. In the first identity in (4.7) we have used that the codimension of  $\ker(df(\lambda))$  is equal to one, and that  $\sigma_\lambda(u, H_f(\lambda)) = df(\lambda)(u) = 0$  if  $u \in \ker(df(\lambda))$ . In the third identity in (4.7) we have used that  $T_\lambda \Lambda$  is a Lagrange plane with respect to the symplectic form  $\sigma_\lambda$ .

The inclusion (4.7) means that, at every point of  $\Lambda$ , the vector field  $H_f$  is tangent tot  $\Lambda$ . This implies that  $\Lambda$  is foliated by the one-dimensional solution curves of the Hamiltonian system defined by the function  $f$ .

Clearly, every submanifold  $I$  of  $\Lambda$  is isotropic, in the sense that, for every  $i \in I$ ,  $T_i I$  is an isotropic linear subspace of  $T_i M$ . Also we have obviously that  $I \subset N$  and that the "H<sub>f</sub>-flow-out of  $I$ ", the set

$$I' := \left\{ e^{t H_f}(i) \mid (i, t) \in J \right\}$$

is contained in  $\Lambda$ , if  $J$  is a suitable open neighborhood of  $I \times \{0\}$  in  $I \times \mathbf{R}$ . If  $\dim I = n - 1$  and  $H_f(i) \notin T_i I$  for every  $i \in I$ , then the mapping

$$(i, t) \mapsto e^{t H_f}(i)$$

is a smooth immersion, hence its image is  $n$ -dimensional, and the conclusion is that  $I'$  is an open subset of  $\Lambda$ . In this sense  $\Lambda$  is locally the only Lagrange submanifold of  $M$  such that  $I \subset \Lambda \subset N$ .

Now suppose conversely that  $I$  is an  $(n - 1)$ -dimensional isotropic submanifold of  $M$ ,  $I \subset N$  and  $H_f(i) \notin T_i I$  for every  $i \in I$ . Then  $I'$  is an  $n$ -dimensional smooth submanifold of  $M$ . Furthermore  $I' \subset N$ , because  $f$  is invariant under the H<sub>f</sub>-flow, and therefore its zeroset  $N$  is invariant under the H<sub>f</sub>-flow. If  $i \in I$  then  $T_i I \subset T_i N$  implies that

$$\mathbf{R} H_f(i) = (T_i N)^{\sigma_i} \subset (T_i I)^{\sigma_i}$$

and therefore

$$\mathbf{T}_i I' = \mathbf{T}_i I + \mathbf{R} \mathbf{H}_f(i)$$

is isotropic, and hence a Lagrange plane because it is  $n$ -dimensional. If  $(i, t) \in J$  and we write  $\Phi = e^{t \mathbf{H}_f}$ , then

$$\mathbf{T}_{\Phi(i)} I' = \mathbf{T}_i \Phi (\mathbf{T}_i I')$$

is a Lagrange plane as well, because  $\mathbf{T}_i \Phi$  preserves the symplectic form. The conclusion is that  $I'$  is a Lagrange submanifold of  $M$  such that  $I \subset I' \subset N$ , and we have a local existence and uniqueness theorem for Lagrange submanifolds  $\Lambda$  of  $M$  such that  $I \subset \Lambda \subset N$ .

**Remark 4.1** More generally, let  $N$  be a smooth submanifold of  $M$  of any dimension, and suppose that the dimension of  $K_n := \mathbf{T}_n N \cap (\mathbf{T}_n N)^\sigma$  does not depend on  $n \in N$ . Because  $K_n$  is equal to the kernel of the restriction to  $\mathbf{T}_n N$  of  $\sigma_n$ , we are in the situation of Subsection (2.3) with  $\omega = \sigma|_N$ . Let  $\pi : N \rightarrow P$  be a fibration as in Subsection (2.3), and let  $\sigma_P$  be the reduced symplectic form, the symplectic form on  $P$  such that  $\sigma|_N = \pi^* \sigma_P$ . Then the Lagrange submanifolds  $\Lambda$  of  $M$  such that  $\Lambda \subset N$  are locally of the form  $\Lambda = \pi^{-1}(\Lambda_P)$ , in which  $\Lambda_P$  is an arbitrary Lagrange submanifold of  $P$  with respect to the symplectic form  $\sigma_P$ . This characterization is also due to Lie.  $\circlearrowright$

### 4.3 An Initial Value Problem

Let  $S$  be an  $(n - 1)$ -dimensional smooth submanifold of  $X$  and  $\psi$  a smooth real-valued function on  $S$ . The above leads to a local existence and uniqueness theorem for solutions  $\phi$  of (4.1) which satisfy the additional "initial condition" that

$$\phi(s) = \psi(s), \quad s \in S. \quad (4.8)$$

We will make the assumptions that  $x_0 \in S$ ,  $\xi_0 \in (\mathbf{T}_{x_0} X)^*$ ,  $f(x_0, \xi_0) = 0$ ,

$$\left. \frac{\partial f(x_0, \xi)}{\partial \xi} \right|_{\xi=\xi_0} \notin \mathbf{T}_{x_0} S. \quad (4.9)$$

and finally

$$d\psi(x_0) = \xi_0|_{\mathbf{T}_{x_0} S}. \quad (4.10)$$

Obviously the condition (4.10) is necessary if we want to have  $\xi_0 = d\phi(x_0)$  for a solution  $\phi$  of (4.1) and (4.8). The transversality condition (4.9) is the natural one in order to avoid singularities in the solution.

For every  $s \in S$ , the restriction mapping  $\xi \mapsto \xi|_{\mathbf{T}_s S}$  is a linear mapping from  $(\mathbf{T}_s X)^*$  onto  $(\mathbf{T}_s S)^*$  with a one-dimensional kernel, equal to  $(\mathbf{T}_s S)^0$ . Therefore the set

$$l_s := \left\{ \xi \in (\mathbf{T}_s X)^* \mid \xi|_{\mathbf{T}_s S} = d\psi(s) \right\}$$

is a straight line in  $(\mathbf{T}_s X)^*$ , and the condition (4.9) means that, at  $\xi_0$ ,  $l_{x_0}$  is transversal to  $N \cap (\mathbf{T}_{x_0} X)^*$ . It follows therefore from the implicit function theorem, that there is an open neighborhood  $S_0$  of  $x_0$  in  $S$  and a neighborhood  $W$  of  $(x_0, \xi_0)$  in  $\mathbf{T}^* X$ , such that for each  $s \in S_0$  there is a unique  $\xi = \xi(s) \in (\mathbf{T}_s X)^*$ , such that  $(x, \xi) \in W$  and

$$\xi|_{\mathbf{T}_s S} = d\psi(s) \quad \text{and} \quad f(s, \xi) = 0. \quad (4.11)$$

Moreover,  $s \mapsto (s, \xi(s))$  is a smooth mapping from  $S_0$  to  $T^*X$ . Because it has  $\pi$  as a left inverse, it is an embedding and the image is a smooth  $(n-1)$ -dimensional submanifold  $I$  of  $T^*X$ , which by construction is contained in  $N$ .

Below we shall prove that  $I$  is isotropic. According to Subsection 4.2, locally there is a unique Lagrange submanifold  $\Lambda$  of  $T^*X$  such that  $I \subset \Lambda \subset N$ , and

$$T_{\lambda_0} \Lambda = T_{\lambda_0} I + \mathbf{R} H_f(\lambda_0)$$

if  $\lambda_0 = (x_0, \xi_0)$ . Because  $T_{\lambda_0} \pi$  maps  $T_{\lambda_0} I$  onto  $T_{x_0} S$  and maps  $H_f(\lambda_0)$  to  $\partial f / \partial \xi(\lambda_0) \notin T_{x_0} S$ , cf. (4.9), we have that the restriction to  $T_{\lambda_0} \Lambda$  of  $T_{\lambda_0} \pi$  is surjective from  $T_{\lambda_0} \Lambda$  to  $T_{x_0} X$ , hence bijective, because both vector spaces have the same dimension  $n$ . It follows that the transversality condition (4.5) holds at  $\lambda = \lambda_0$ , and therefore  $\Lambda$  is locally equal to the graph of  $d\phi$  for a smooth function  $\phi$ , which is a solution of (4.1) because  $\Lambda \subset N$ . On the other hand  $I \subset \Lambda$  implies that for all  $s$  in a connected neighborhood  $S_0$  of  $x_0$  in  $S$  we have that the restriction to  $T_s S$  of  $d\phi(s)$  is equal to  $d\psi(s)$ , which implies that  $d(\phi|_{S_0}) = d\psi$ , or  $\phi|_{S_0} - \psi = c$  is a constant. Replacing  $\phi$  by  $\phi - c$  we arrive at the locally unique solution of the initial value problem (4.1), (4.8).

In order to prove that  $I$  is an isotropic submanifold of  $T^*X$ , we consider the submanifold  $T_S^* X$  of all  $(x, \xi) \in T^*X$  such that  $x \in S$  and  $\xi \in (T_x X)^*$ . Let  $\iota$  denote the identity as a mapping from  $T_S^* X$  to  $T^*X$ , and define the restriction mapping  $\rho : T_S^* X \rightarrow T^*S$  by

$$\rho(x, \xi) = (x, \xi|_{T_x S}), \quad x \in S, \quad \xi \in (T_x X)^*.$$

Then

$$\iota^*(\tau_{T^*X}) = \rho^*(\tau_{T^*S}) \quad \text{on} \quad T_S^* X,$$

which is a tautology if one writes out the definitions of the left and right hand side. Taking the exterior derivative of the left hand side and using that the exterior derivative commutes with pullbacks by smooth mappings, we obtain that

$$\iota^*(\sigma_{T^*X}) = \rho^*(\sigma_{T^*S}) \quad \text{on} \quad T_S^* X.$$

Because the graph of  $d\psi$  is an isotropic submanifold  $\Lambda_S$  of  $T^*S$ , it follows that  $\rho^{-1}(\Lambda_S)$  is an isotropic submanifold of  $T^*X$ , and therefore  $I = \rho^{-1}(\Lambda_S) \cap N$  is isotropic as well.

#### 4.4 Ray Bundles

It is a classical observation that the bundles of rays (= straight lines) which appear in geometrical optics, are orthogonal to some hypersurface  $S$ . An example is the bundle of rays which emanate from a given source point, these are the normals to every sphere with center at the source point. In the plane every bundle of rays = one-parameter family of straight lines is orthogonal to some curve, for this it suffices to take a solution curve of a vector field which is orthogonal to the rays. However, in higher dimensions  $n$  it is a quite special property of an  $(n-1)$ -parameter family of rays to be normal to a hypersurface  $S$ .

Let  $S$  be an oriented hypersurface, which leads to an orientation of the normals of  $S$ . Let  $\phi$  be the function which is equal to zero on  $S$  and has derivative in the direction of the oriented normals equal to 1. Then, at least where the normals define a fibration of the space,  $\phi$  is a smooth function, and the rays are orthogonal to every level hypersurface of  $\phi$ . Indeed, if  $n(s)$  denotes the normal vector to  $S$  at the point  $s \in S$ , then the level set of  $\phi$  at the level  $t$  is equal to the set of points

$$S_t = \{s + t n(s) \mid s \in S\}$$

and its tangent space consists of the points  $\delta s + t \operatorname{D}n(s) \delta s$  in which  $\delta s$  is tangent to  $S$ . It follows from  $\langle n(s), n(s) \rangle \equiv 1$  that  $\langle \operatorname{D}n(s) \delta s, n(s) \rangle = 0$ , and therefore  $\langle \delta s + t \operatorname{D}n(s) \delta s, n(s) \rangle = \langle \delta s, n(s) \rangle = 0$ .

It follows that the gradient of  $\phi$  at the point  $s + t n(s)$  is equal to  $n(s)$ , which means that the function  $\phi$  satisfies the nonlinear first order partial differential equation

$$\sum_{i=1}^n \left( \frac{\partial \phi(x)}{\partial x_i} \right)^2 = 1. \quad (4.12)$$

Conversely, if  $\phi$  satisfies (4.12), then the normals to the level hypersurfaces of  $\phi$  form an  $(n-1)$ -parameter family of straight lines.

Hamilton [10] discovered this bijective correspondence between the ray bundles in geometrical optics and real-valued function  $\phi$  which satisfy a partial differential equation of the form (4.12), where the rays are the lines which are orthogonal to the level surfaces of  $\phi$ . He called  $\phi$  the *characteristic function of the ray bundle*, and the partial differential equation (4.12) is sometimes called the *eikonal equation* of geometrical optics.

The equation (4.12) is of the form (4.1), if we take

$$f(x, \xi) = \sum_{i=1}^n (\xi_i)^2 - 1. \quad (4.13)$$

The corresponding Hamiltonian system is  $dx/dt = 2\xi$ ,  $d\xi/dt = 0$ , we obtain that the velocity vector  $dx(t)/dt$  does not depend on  $t$ . Moreover, it is pointing in the direction of  $\xi = \operatorname{grad} \phi$ , which means that the projections  $x(t)$  to the position space  $X = \mathbf{R}^n$  are straight lines orthogonal to the level hypersurfaces of  $\phi$ , they form the ray bundle corresponding to solution  $\phi$  of the eikonal equation.

In the construction of the characteristic function, it is essential that the rays define a fibration of  $X$ , which is the case as long as the Lagrange manifold  $\Lambda$ , which is supposed to be the graph of  $d\phi$ , is transversal to the fibers, cf. (4.5). However, it is a very common phenomenon that at some points the rays start criss-crossing, which correspond to points where the transversality condition (4.5) no longer holds. At such a point the density of the rays becomes infinite, and for this reason such a point is called a *caustic point*, a point "where the light burns". At caustic points the characteristic function is no longer smooth, and at points through which more than one ray passes the function  $\phi$  becomes multi-valued.

The good news is that, even if such singularities in the ray bundle and its characteristic function occur, *the Lagrange manifold  $\Lambda$  remains a smoothly immersed submanifold of  $T^*X$* . Therefore, in order to include ray bundles with caustics, we propose the following definition:

*A ray bundle consists of the projections to the base manifold  $X$  of the  $H_f$ -solution curves in a smoothly immersed Lagrange submanifold of  $T^*X$ , which is contained in the zeroset of  $f$ .*

This makes Lie's point of view, of allowing *any* smoothly immersed Lagrange submanifold of  $T^*X$  which is contained in the zeroset of  $f$  as a solution of (4.1), not just a matter of abstract generalization, but very relevant from the point of view of practical applications.

Ray propagation in inhomogenous media, where the local speed of propagation  $c(x)$  depends smoothly on the position  $x \in X$ , is described by the above theory in which the function  $f$  in (4.13) is replaced by

$$f(x, \xi) = c(x)^2 \sum_{i=1}^n (\xi_i)^2 - 1.$$

Ray bundles defined by other types of functions  $f$  on  $T^*X$  also have applications.

## 4.5 High Frequency Waves and Fourier Integral Operators

Huygens [15] could not convince the physicists of the wave nature of light, but Young and especially Fresnel [8] did, with their beautiful quantitative analysis of interference patterns, which have no decent explanation in a particle model.

Let us consider waves which are solutions of the standard *wave equation*

$$\square := \frac{\partial^2 u}{\partial t^2} - \Delta u := \frac{\partial^2 u}{\partial t^2} - \sum_{j=1}^n \frac{\partial^2 u}{\partial x_j^2}. \quad (4.14)$$

Let us see what happens if we apply the wave operator  $\square$  to a *simple progressing wave*

$$u(x, t) = e^{i\omega(t-\phi(x))} a(x), \quad (4.15)$$

in which  $\phi(x)$  and  $a(x)$  are real valued functions of  $x$ , called the *phase function* and the *amplitude function* of the simple progressing wave, where the level surfaces  $\phi(x) = t$  in the position space, the  $x$ -space, are the *wave fronts*.  $\omega$  is a *frequency* variable, and we are in particular interested in the asymptotic behaviour as  $\omega \rightarrow \infty$ . (Due to the small wave length and the very high speed of propagation, visible light has an extremely high frequency.)

We have

$$(\square u)(x, t) = e^{i\omega(t-\phi(x))} [\omega^2 u_2(x, t) + i\omega u_1(x, t) + u_0(x, t)],$$

in which

$$u_2(x, t) = \left( \sum_{j=1}^n \left( \frac{\partial \phi(x)}{\partial x_j} \right)^2 - 1 \right) a(x),$$

$$u_1(x, t) = 2 \sum_{j=1}^n \frac{\partial \phi(x)}{\partial x_j} \frac{\partial a(x)}{\partial x_j} + (\Delta \phi)(x) a(x)$$

and  $u_0(x, t) = -\Delta a(x)$ .

If  $a(x) \neq 0$ , then  $(\square u)(x, t)$  grows quadratically as a function of  $\omega$  for  $\omega \rightarrow \infty$ , unless the phase function  $\phi(x)$  satisfies the eikonal equation (4.12).

Assuming that  $\phi$  satisfies the eikonal equation, the leading term in  $\square u$  grows linearly with  $\omega$ , unless the amplitude satisfies the homogeneous linear first order partial differential equation

$$u_1(x, t) = 2 \sum_{j=1}^n \frac{\partial \phi(x)}{\partial x_j} \frac{\partial a(x)}{\partial x_j} + (\Delta \phi)(x) a(x) = 0, \quad (4.16)$$

called the *transport equation* for the amplitude.

In order to analyse (4.16), we consider the solutions  $x(s)$  of the system of ordinary differential equations

$$\frac{dx_j}{ds} = 2 \frac{\partial \phi(x)}{\partial x_j}, \quad 1 \leq j \leq n. \quad (4.17)$$

The solution curves of (4.17) are orthogonal to the wave fronts  $\phi(x) = \text{constant}$ , and therefore are equal to the rays of the ray bundle of which  $\phi$  is the characteristic function. The transport equation (4.16) then is equivalent to

$$\frac{da(x(s))}{ds} + (\Delta\phi)(x(s)) a(x(s)) = 0,$$

a homogeneous first order linear ordinary differential equation for the function  $s \mapsto a(x(s))$ . This implies that we can prescribe  $a(x)$  freely on an  $(n-1)$ -dimensional manifold  $S$  which is transversal to the rays. If we choose the "initial amplitude"  $a|_S$  equal to zero outside a small neighborhood of a given point  $x_0$  then the solution  $a(x)$  will be equal to zero outside a narrow tube along the ray through the point  $x_0$ . In this sense the waves  $u$  of the form (4.15) which are *asymptotic solutions* of (4.14) in the sense that  $\square u$  remains bounded as  $\omega \rightarrow \infty$ , will *propagate along the rays*. In this sense geometrical optics is the high frequency limit of wave optics.

The procedure can be refined by replacing the amplitude function  $a(x)$  in (4.15) by an asymptotic expansion of the form

$$a(x, \omega) \sim \sum_{k=0}^{\infty} a_j(x) \omega^{-k}, \quad \omega \rightarrow \infty$$

in negative powers of the frequency  $\omega$ . One may verify that by successively solving inhomogenous linear ordinary differential equations along the rays for the  $a_k(x)$ ,  $k \geq 1$ , one can arrange that, for any  $K$ ,  $\square u = \mathcal{O}(\omega^{-K})$  as  $\omega \rightarrow \infty$ .

As observed in Subsection 4.4, the construction of the phase function  $\phi(x)$ , and therefore of the simple progressing wave (4.15), breaks down at caustic points, and the amplitude  $a(x)$  becomes infinite if one approaches such a point. It turns out that near such points one can still obtain asymptotic solutions of the wave equation by replacing the simple progressing wave (4.15) by a "continuous superposition" of such waves, an *oscillatory integral* of the form

$$u(x, t) \sim \sum_{k=0}^{\infty} \omega^{N/2-k} \int_{\mathbf{R}^N} e^{i\omega(t-\phi(x,\theta))} a_k(x, \theta) d\theta, \quad \omega \rightarrow \infty. \quad (4.18)$$

Here, in order to avoid any problems with the convergence of the integral over the auxiliary  $\theta$ -variables, it is assumed that  $a_k(x, \theta) = 0$  for all  $\theta$  outside a compact subset. Maslov [23] showed that for quite general linear partial differential equations  $Pu = 0$  one can construct oscillatory integrals (4.18) which are global asymptotic solutions, i.e. also in neighborhoods of caustic points. These oscillatory integrals correspond to Lagrange submanifolds  $\Lambda$  of  $T^*X$  which are contained in the zeroset of a certain function  $p$  on  $T^*X$  which is called the *principal symbol* of the linear differential operator  $P$ . Cf. [6] for a survey.

If one also integrates over the frequency variable  $\omega$ , then one obtains distributions, which are called *Fourier integral distributions*, invented by Hörmander [14]. The singularities of a Fourier integral distribution, i.e. its behaviour near points where it is not equal to a smooth function, have a very precise description in terms of the corresponding Lagrange submanifold of  $T^*X$ .

A simple example is Dirac's delta function situated at a point  $x \in X$ . It is a Fourier integral distribution and the Lagrange manifold corresponding to it is the fiber  $(T_x X)^*$  of the cotangent bundle  $T^*X$  over the point  $x$ . This is an extreme example of a Lagrange manifold which does not satisfy the transversality condition (4.5).

Linear integral operators from  $\mathcal{F}(Y)$  to  $\mathcal{F}(X)$  which have a distribution kernel on  $X \times Y$  which is a Fourier integral distribution are called *Fourier integral operators*. These form a very wide class

of operators, which include all the propagation operators = "Green functions" of linear partial differential equations  $Pu = 0$  of wave type. The aforementioned description of the singularities of Fourier integral distributions leads to very detailed descriptions of the propagation of singularities of the solutions of the partial differential equation  $Pu = 0$ , or of the corresponding inhomogeneous equation  $Pu = f$ , cf. [5].

## 4.6 Some History

As already mentioned, Hamilton [10] found the description of ray bundles in geometrical optics in terms of their characteristic functions, together with the nonlinear partial differential equation (4.12) satisfied by the characteristic function.

However, he did not ask the question how to solve a general nonlinear partial differential equation of the form (4.1). — it was Jacobi [16], [17] who observed that the solution of such a partial differential equation can be reduced to the solution of a *Hamiltonian* system of ordinary differential equations. Since then the theory is called "Hamilton-Jacobi theory".

Lie developed the idea of viewing the solution as a Lagrange submanifold of the cotangent bundle in a series of articles in 1872–78, and stressed the point that it is based on the fact that the flow of the Hamiltonian system of the function  $f$  leaves the canonical two-form  $\sigma$  of the cotangent bundle invariant. For him the use of the group of transformations which leave  $f$  and  $\sigma$  invariant was analogous to the use of the Galois group in the solution of polynomial equations. Engel introduced the name "Verein" = "club" for any submanifold of the cotangent bundle. Lie followed this, but later authors like Élie Cartan didn't. An accessible account of Lie's ideas on first order partial differential equations is the book of Engel and Faber [7].

The fact that Lagrange manifolds are fundamental in so many situations made Weinstein [28] talk about the *Symplectic Creed*: "Everything is a Lagrange manifold". More recently so-called *special Lagrange manifolds*, invented by Harvey and Lawson [12], made their appearance in mirror symmetry.

## 4.7 Exercises

**Exercise 4.1** In the notation of Example 3.3 on integrable systems, prove that the level sets  $M_c$  are Lagrange submanifolds of  $M$ . If  $M = T^*X$  and the connected component  $C$  of  $M_c$  is compact, then  $C$  can only be equal to  $d\phi$  for some smooth function  $\phi$  on  $X$ , if  $X$  is diffeomorphic to a torus.

◊

**Exercise 4.2** Prove that, for any given  $x \in X$ , the fiber  $(T_x X)_j$  over  $x$  is a Lagrange submanifold of  $T^*X$ . Define

$$I = \{\xi \in (T_x X)^* \mid f(x, \xi) = 0\}.$$

Suppose that  $\partial f(x, \xi)/\partial \xi \neq 0$  for every  $\xi \in I$ . Prove that  $N$  is an  $(n - 1)$ -dimensional isotropic submanifold of  $T^*X$ , that  $(i, t) \mapsto e^{tH_f}(i)$  defines a smooth immersion from an open subset  $J$  of  $I \times \mathbf{R}$  to  $T^*X$ , and that the image is a smoothly immersed Lagrange submanifold  $\Lambda$  of  $T^*X$  such that  $\Lambda \subset N$ , where  $N$  denotes the zero set of  $f$  in  $T^*X$ . Verify that the projections in  $X$  of the

$H_f$ -solution curves in  $\Lambda$  all pass through the given point  $x$ . This is called the *ray bundle emanating from  $x$* , for a general function  $f$  on  $T^*X$ . ◊



## References

- [1] V.I. Arnol'd: Characteristic class entering in quantization condition. *Func. Anal. Appl.* **1** (1967) 1–13.
- [2] N. Burgoyne and R. Cushman: Conjugacy classes in linear groups. *J. Algebra* **44** (1977) 339–362.
- [3] E.A. Coddington and N. Levinson: *Theory of Ordinary Differential Equations*. McGraw-Hill Book Company, Inc., New York, etc., 1955.
- [4] R. Cushman and J.J. Duistermaat: The behavior of the index of a periodic linear hamiltonian system under iteration. *Advances in Math.* **23** (1977) 1–21.
- [5] J.J. Duistermaat and L. Hörmander: Fourier integral operators II. *Acta Math.* **128** (1972) 183–269.
- [6] J.J. Duistermaat: Oscillatory integrals, Lagrange immersions and unfoldings of singularities. *Comm. Pure Appl. Math.* **27** (1974) 207–281.
- [7] F. Engel und K. Faber: *Die Liesche Theorie der partiellen Differentialgleichungen Erster Ordnung*. Teubner, Berlin, 1935.
- [8] A. Fresnel: Mémoire sur la diffraction de la lumière, couronné par l'Académie des Sciences en 1819. pp. 247–382 in vol. 1 of: *Œuvres Complètes d'Augustin Fresnel*. Paris, 1866. Johnson Reprint Corporation, New York, 1965.
- [9] P. Griffiths and J. Harris: *Principles of Algebraic Geometry*. John Wiley and Sons, New York, etc., 1978.
- [10] W.R. Hamilton: Essay on the theory of systems of rays, with three supplements. *Trans. Royal Irish Acad.* **15** (1828) 69–174, **16** (1830 and 1831) 4–62 and 85–92, **17** (1837) 1–144.
- [11] W.R. Hamilton: On a general method in dynamics. *Phil. Trans. Royal Soc. London* (1834) 247–308, (1835) 95–144.
- [12] R. Harvey and H.B. Lawson, Jr.: Calibrated geometries. *Acta Math.* **148** (1982) 47–157.
- [13] L. Hörmander: *An Introduction to Complex Analysis in Several Variables*. North-Holland Publ. Co., Amsterdam, 1973.
- [14] L. Hörmander: Fourier integral operators I. *Acta Math.* **127** (1971) 79–183.
- [15] C. Huygens: *Traité de la Lumière*. Van der Aa, Leyden, 1690.
- [16] C.G.J. Jacobi: Über die Reduction der Integration der partiellen Differentialgleichungen erster ordnung zwischen irgend einer Zahl Variabeln auf die Integration eines einzigen Systemes gewöhnlicher Differentialgelichungen. *Crelle's Journal für die reine und angewandte Mathematik* **1** (8137) 97–162 = *Gesammelte Werke*, 4. Band, 1886. Reprint Chelsea Publ. Cy., New York, 1969.

- [17] C.G.J. Jacobi: *Vorlesungen über Dynamik*. Gehalten an der Universität Königsberg im Wintersemester 1842–43 und nach einem von C.W. Borchardt ausgearbeiteten Hefte. Verlag G. Reimer, Berlin, 1881. Reprint Chelsea Publ. Cy., New York, 1969.
- [18] C.G.J. Jacobi: Nova methodus, equationes differentiales partiales primi orins inter numerum variabilium quemcuque propositas integrandi. *Crelle journal für die reine und angewandte Mathematik* **60** (1862) 1–181. Also in: *Gesammelte Werke*, V. Band, 1–189., Berlin 1890. Reprint Chelsea Publ. Cy., New York, 1969.
- [19] J.-L. Lagrange: *Mécanique Analytique*. First print 1788 in Paris. Reprint by A. Blanchard, Paris, 1965.
- [20] A. Lichnerowicz: Les variétés de Poisson et leurs algèbres de Lie associées. *J. Differential Geometry* **12** (1977) 253–300.
- [21] S. Lie: *Theorie der Transformationsgruppen* I–III. Unter Mitwirkung von Prof.dr. F. Engel. Teubner, Leipzig, 1888, 1890, 1893.
- [22] B. Malgrange: Sur l'intégrabilité des structures presque-complexes. pp. 289-296 in: *Symposia Mathematica* II, INDAM, Rome, 1968.
- [23] V.P. Maslov: *Théorie des Perturbations et Méthodes Asymptotiques*. Dunod, Gauthier-Villars, Paris, 1972. This is a French translation of the book published in 1965 by the University of Moscow, with some additions.
- [24] J. Moser: On the volume elements on a manifold. *Trans. Amer. Math. Soc.* **120** (1965) 286–294.
- [25] A. Newlander and L. Nirenberg: Complex analytic coordinates in an almost complex manifold. *Annals of Math.* **65** (1957) 391–404.
- [26] Julius Plücker: *Neue Geometrie des Raumes*, gegründet auf die Betrachtung der gerade Linie als Raumelement. Teubner Verlag, Leipzig, 1868/89.
- [27] A. Weinstein: Symplectic manifolds and their Lagrangian submanifolds. *Advances in Mathematics* **6** (1971) 329–346.
- [28] A. Weinstein: Symplectic geometry. *Bull. Amer. Math. Soc.* **5** (1981) 1–13.
- [29] Hermann Weyl: *The Classical Groups*, their invariants and representations. Princeton University Press, Princeton, N.J., 1939, 1946.
- [30] John Williamson: On the algebraic problem concerning normal forms of linear dynamical systems. *Amer. J. Math.* **58** (1936) 141–163.