# Appendix 1

# OVERVIEW OF SYMPLECTIC GEOMETRY

Symplectic geometry is the language underlying the theory of Hamiltonian systems. This appendix is a short review of the main concepts, especially as they apply to Hamiltonian systems and symplectic maps in cotangent bundles. These spaces are natural when considering mechanical systems, where the base, or *configuration space* describes the position and the momentum belongs to the fiber of the cotangent bundle of the configuration space. In our optic of symplectic twist maps, one important concept studied in this chapter is that of exact symplectic map. Theorem 59.7 proves that Hamiltonian systems give rise to exact symplectic maps.

We assume here some familiarity with the notions of manifold, vector bundle and differential form. The reader who is uncomfortable with these concepts should consult Guillemin & Pollack (1974) , Spivak (1970) or Arnold (1978). For more on symplectic geometry and Hamiltonian systems, see Arnold (1978), Weinstein (1979), Abraham & Marsden (1985) or McDuff & Salamon (1996).

## 54. Symplectic Vector Spaces

In this section, we review some essentials of the linear theory of symplectic vector spaces and transformations. They will be our tools in understanding the infinitesimal behavior of symplectic maps and Hamiltonian systems in cotangent bundles. A *symplectic form* $\Omega$ on a real vector space $V$ is a bilinear form $\Omega$ which is skew symmetric and nondegenerate:

$$\Omega(a\boldsymbol{v}+b\boldsymbol{v}',\boldsymbol{w})=a\Omega(\boldsymbol{v},\boldsymbol{w})+b\Omega(\boldsymbol{v}',\boldsymbol{w}),\ \ (\boldsymbol{v},\boldsymbol{v}',\boldsymbol{w}\in V,\ a,b\in\mathbb{R}).$$

$$\Omega(\boldsymbol{v},\boldsymbol{w})=-\Omega(\boldsymbol{w},\boldsymbol{v})$$

$$\boldsymbol{v}\neq 0\Rightarrow\exists\boldsymbol{w}\text{ such that }\ \Omega(\boldsymbol{v},\boldsymbol{w})\neq 0$$

A *symplectic vector space* is a vector space $V$ together with a symplectic form.

**Example 54.1** The determinant in $\mathbb{R}^2$ is a symplectic form. More generally, the *canonical symplectic form* on $\mathbb{R}^{2n}$, is given by:

$$\Omega_0(\boldsymbol{v},\boldsymbol{w})=\langle J\boldsymbol{v},\boldsymbol{w}\rangle,\quad J=\begin{pmatrix}0 & -Id\\ Id & 0\end{pmatrix}\tag{54.1}$$

where the brackets $\langle\ ,\ \rangle$ denote the usual dot product. We will see in the next theorem that all symplectic vector spaces "look" like this. In particular, their dimension is always even. Usually, one writes:

$$\Omega_0=d\boldsymbol{q}\wedge d\boldsymbol{p}=\sum_{k=1}^{n}dq_k\wedge dp_k$$

where it is understood that $dq_k, dp_k$ are elements of the dual basis for the coordinates $(q_1,\dots,q_n,p_1,\dots,p_n)$ of $\mathbb{R}^{2n}$. One can view $dq_k\wedge dp_k(\boldsymbol{v},\boldsymbol{w})$ as the determinant of the projections of $\boldsymbol{v}$ and $\boldsymbol{w}$ on the plane of coordinates $(q_k,p_k)$ (see Exercise 54.5). The symplectic space $(\mathbb{R}^{2n},\Omega_0)$ can also be interpreted as $\mathbb{R}^n\oplus(\mathbb{R}^n)^*$, equipped with the canonical symplectic form:

$$\Omega_0(\boldsymbol{a}\oplus\boldsymbol{b},\boldsymbol{c}\oplus\boldsymbol{d})=\boldsymbol{d}(\boldsymbol{a})-\boldsymbol{b}(\boldsymbol{c}).$$

It is often convenient to view a bilinear form as a matrix. To do this, fix a basis $(\boldsymbol{e}_1,\dots,\boldsymbol{e}_n)$ of $V$, and set:

$$A^{\Omega}_{ij}=\Omega(\boldsymbol{e}_i,\boldsymbol{e}_j)$$

Equivalently, if $\langle\ ,\ \rangle$ is the dot product associated with the basis $(\boldsymbol{e}_1,\dots,\boldsymbol{e}_n)$, then $A^{\Omega}$ is the matrix satisfying:

$$\Omega(\boldsymbol{v},\boldsymbol{w})=\langle A^{\Omega}\boldsymbol{v},\boldsymbol{w}\rangle.$$

With any bilinear form $\Omega$ on a vector space comes a notion of *orthogonal subspace* $W^{\perp}$ to a given subspace (or vector) $W$ :

$$W^{\perp}=\{\boldsymbol{v}\in V\mid\Omega(\boldsymbol{v},\boldsymbol{w})=0,\forall\boldsymbol{w}\in W\}$$

In the case of symplectic forms, the analogy with the usual notion of orthogonality can be quite misleading, as a subspace and its orthogonal will often intersect. We now show that all symplectic vector spaces are isomorphic to the canonical $(\mathbb{R}^{2n}, \Omega_0)$.

**Theorem 54.2 (Linear Darboux)** *If $(V, \Omega)$ is a symplectic space, one can find a basis for $V$ in which the matrix $A^\Omega$ of $\Omega$ is given by $A^\Omega = J = \begin{pmatrix} 0 & -Id \\ Id & 0 \end{pmatrix}$.*

Hence, the isomorphism that sends each vector in $V$ to its coordinate vector in the basis given by the theorem will be an isomorphism between $(V, \Omega)$ and $(\mathbb{R}^{2n}, \Omega_0)$. In classical notation, the coordinates in the Darboux coordinates are denoted by[20]

$$(\boldsymbol{q}, \boldsymbol{p}) = (q_1, \ldots, q_n, p_1, \ldots, p_n).$$

Note in particular that *a symplectic space always has even dimension.*

*Proof of Theorem 54.2.* Since $\Omega$ is nondegenerate, given any $\boldsymbol{v} \neq 0 \in V$, we can find a vector $\boldsymbol{w} \in W$ such that $\Omega(\boldsymbol{v}, \boldsymbol{w}) = -1$. In particular, the plane $P$ spanned by $\boldsymbol{v}$ and $\boldsymbol{w}$ is a symplectic plane and the bilinear form induced by $\Omega$ on $P$ with this basis has matrix:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{54.2}$$

Since $\Omega$ is nondegenerate on $P$, we must have $P^\perp \cap P = \{0\}$. Furthermore $V = P + P^\perp$, since if $\boldsymbol{u} \in V$,

$$\boldsymbol{u} - \Omega(\boldsymbol{u}, \boldsymbol{v})\boldsymbol{w} + \Omega(\boldsymbol{u}, \boldsymbol{w})\boldsymbol{v} \in P^\perp.$$

$\Omega$ must be nondegenerate on the subspace $P^\perp$ of dimension $dimV - 2$, so we can proceed by induction, and decompose $P^\perp$ into $\Omega$–orthogonal planes on which the matrix of $\Omega$ is as in (54.2). A permutation of the vectors of the basis we have found gives $A^\Omega = J$. $\square$

**Exercise 54.3** Show that the linear transformation whose matrix is $J$ in the canonical basis is orthogonal (i.e belongs to $O(2n)$), that it satisfies $J^2 = -Id$ ($J$ is then called a complex structure) as well as

---

[20] In the literature, one also sees frequently $(\boldsymbol{p}, \boldsymbol{q})$, with $-J = \begin{pmatrix} 0 & Id \\ -Id & 0 \end{pmatrix}$ as canonical matrix.

$$\Omega_0(J\boldsymbol{v}, J\boldsymbol{w}) = \Omega_0(\boldsymbol{v}, \boldsymbol{w})$$

(that is, $J$ is *symplectic*, see section 56).

**Exercise 54.4** Show that a one dimensional vector subspace in a symplectic vector space is included in its own orthogonal subspace.

**Exercise 54.5** Show that in a Darboux basis for a symplectic plane,

$$\Omega(\boldsymbol{v}, \boldsymbol{w}) = det(\boldsymbol{v}, \boldsymbol{w}).$$

Hence the form $\sum_1^n dq_k \wedge dp_k$ applied to 2 vectors can be seen as the sum of minor determinants of these two vectors (see Arnold (1978)).

**Exercise 54.6** Prove that a general skew symmetric (not necessarily nondegenerate) form $\Omega$ can be put, by an appropriate change of variable, into the "normal form":

$$A^{\Omega} = \begin{pmatrix} 0_k & -Id_k & \\ Id_k & 0_k & \\ & & 0_l \end{pmatrix}$$

where the integers $k, l$ do not depend on the basis chosen.

## 55. Subspaces of a Symplectic Vector Space

Let $V$ be a symplectic vector space of dimension $2n$, $W \subset V$ a subspace, and $\Omega_W$ the symplectic form restricted to $W$. The previous exercise shows that we can find a basis for $W$ in which :

$$A^{\Omega_W} = \begin{pmatrix} 0_k & -Id_k & \\ Id_k & 0_k & \\ & & 0_l \end{pmatrix} \quad \text{with} \quad dimW = 2k + l$$

Furthermore, even though there may be many bases in which $\Omega_W$ has a matrix of this form, $k$ and $l$ are independent of the choice of such a basis. In other words, $(W, \Omega_W)$ is determined up to isomorphism by its dimension and by $k$. We will say that $W$ is:

- *null* or *isotropic* if $k = 0$ (and $l = dimW$),
- *coisotropic* if $k + l = n$.
- *Lagrangian* if $k = 0$ and $l = n$ (i.e. W is isotropic and coisotropic).
- *symplectic* if $l = 0$ and $k \neq 0$.

The *rank* of $W$ is the integer $2k$.

The next theorem tells us that the qualitatively different subspaces of a symplectic space can be represented by coordinate subspaces in some Darboux coordinates.

**Theorem 55.1** *A subspace $W$ of rank $2k$ and dimension $2k+l$ in a symplectic space can be represented, in appropriate Darboux coordinates, by the subspace of coordinates:*

$$(q_1, \dots, q_{k+l}, p_1, \dots, p_k).$$

In particular, in some well chosen bases, an isotropic space is made entirely of $q$'s and a coisotropic one must have at least $n$ $q$'s (the role of $p$'s and $q$'s can be reversed, of course) and a symplectic space has the same number of $q$'s and $p$'s.

*Proof*. From the definition of the rank of $W$, there is a subspace $U$ of $W$ of dimension $2k$ which is symplectic, on which we can put Darboux coordinates. $U^\perp \cap W$, the null space of $\Omega_W$, is in the subspace $U^\perp$, which is symplectic (see Exercise 55.4 ). The next lemma shows that we can complete any basis of $U^\perp \cap W$ with coordinates that we denote by $(q_{k+1}, \dots, q_{k+l})$ into a symplectic basis of $U^\perp$, with coordinates $(q_{k+1}, \dots, q_{k+l}, p_{k+1}, \dots, p_{k+l})$. The union of this basis and the one for $U$ (of coordinates say $(q_1, \dots, q_k, p_1, \dots p_k)$) is a symplectic basis, in which $W$ can be expressed as advertised. □

**Lemma 55.2** *Let $U$ be a null (i.e. isotropic) subspace of a symplectic space $V$ Then one can complete any basis of $U$ into a symplectic basis of $V$.*

*Proof*. Without loss of generality, $V$ is $\mathbb{R}^{2n}$ with its standard dot product and canonical symplectic form. Choose an orthonormal basis $(u_1, \dots, u_l)$ for $U$ (in the sense of the dot product). Using (54.1) and the results of Exercise 54.3, the reader can easily check that $JU$ is orthogonal to $U$ and that $(u_1, \dots, u_l, Ju_1, \dots, Ju_l)$ is a symplectic basis for $E = U \oplus JU$. From Exercise 55.4, $E^\perp \oplus E = V$ and $E^\perp$ is symplectic. We can complete the symplectic basis of $E$ by any symplectic basis of $E^\perp$ and get a symplectic basis for $V$. □

As a simple consequence of Theorem 55.1, we get:

**Corollary 55.3** *If $U$ is an isotropic subspace of a symplectic space $V$, one can find a coisotropic $W$ such that $V = U \oplus W$. One can also find a Lagrangian subspace in which $U$ is included.*

This applies in particular to Lagrangian subspaces: given any Lagrangian subspace $L$, we can find another one $L'$ such that $V = L \oplus L'$. In the normal coordinates of the theorem, $L$ could be the $\boldsymbol{q}$ coordinate subspace and $L'$ the $\boldsymbol{p}$ coordinate subspace.

**Exercise 55.4** Let $W$ be a subspace of a symplectic space $V$. Show that: $W$ is symplectic $\iff W \oplus W^{\perp} = V \iff W^{\perp}$ is symplectic (*Hint.* see the proof of the Linear Darboux theorem).

**Exercise 55.5** Show that:
$W$ isotropic $\iff W \subset W^{\perp}$.
$W$ coisotropic $\iff W^{\perp} \subset W$.
$W$ is Lagrangian $\iff W$ is a maximal isotropic subspace, or minimal coisotropic subspace (for the inclusion).

Note that the above equivalence are the definition most often seen in the literature.

**Exercise 55.6** This exercise shows how symmetric matrices can be used to locally parameterize the space of Lagrangian planes. Suppose you are given a basis $\boldsymbol{v}_1, \dots \boldsymbol{v}_n$ for a Lagrangian subspace $L$ of $\mathbb{R}^{2n}$. In the canonical coordinates $(\boldsymbol{q}, \boldsymbol{p})$, write $\boldsymbol{v}_k = (\boldsymbol{x}_k, \boldsymbol{y}_k)$. Let $X$ and $Y$ be the $n \times n$ matrices whose columns are the $x_k$'s and $y_k$'s respectively. Suppose that $L$ is a graph over the $\boldsymbol{q}$–plane.

(a) Show that $X$ is invertible and that the column vectors of the $2n \times n$ matrix $\begin{pmatrix} I \\ YX^{-1} \end{pmatrix}$ form a basis for $L$.
(b) Show that the matrix $YX^{-1}$ is symmetric.
(c) Deduce from this that the (Grassmanian) space of Lagrangian subspaces of $\mathbb{R}^{2n}$ has dimension $n(n+1)/2$.

# 56. Symplectic Linear Maps

**The Symplectic Group.** The Linear Darboux Theorem tells us that, up to changes of coordinates, all symplectic vector spaces are identical to $(\mathbb{R}^{2n}, \Omega_0)$. Therefore, as we define and study the transformations that preserve the symplectic form on a vector space, we need only consider the case $(\mathbb{R}^{2n}, \Omega_0)$.

**Definition 56.1** A *symplectic linear map* $\Phi$ of $(\mathbb{R}^{2n}, \Omega_0)$ is a 1 to 1 linear map which leaves invariant the symplectic form:

$$\Phi^* \Omega_0 = \Omega_0, \quad \text{where } \Phi^* \Omega_0(\boldsymbol{v}, \boldsymbol{w}) \stackrel{\text{def}}{=} \Omega_0(\Phi \boldsymbol{v}, \Phi \boldsymbol{w}). \tag{56.1}$$

The group formed by symplectic linear maps is called the *symplectic group* and is denoted by $Sp(2n;\mathbb{R})$, or in short $Sp(2n)$. Because of (56.1) and (54.1), this group is naturally identified with the group of $2n \times 2n$ real matrices $\Phi$ that satisfy:

$$\Phi^t J \Phi = J \tag{56.2}$$

**Examples 56.2**

(a) The group Sp(2) is the group of $2 \times 2$ matrices of determinant 1.
(b) The transformation $F(\boldsymbol{q},\boldsymbol{p}) = (\boldsymbol{q}+\boldsymbol{p},\boldsymbol{p})$, with matrix $\begin{pmatrix} Id & Id \\ 0 & Id \end{pmatrix}$ is symplectic in $\mathbb{R}^{2n}$, and so is any with matrix $\begin{pmatrix} Id & A \\ 0 & Id \end{pmatrix}$, where $A^t = A$. These maps are called *completely integrable* as they preserve the $n$ dimensional foliation of (affine) Lagrangian planes $\{\boldsymbol{p} = constant\}$.
(c) The examples of symplectic linear map this book is most concerned with are the differentials of symplectic twist maps and of the time 1 map of Hamiltonian flows.

**Spectral Properties of Symplectic Maps.**

**Theorem 56.3** *Symplectic linear maps have determinant 1. If $\lambda$ is an eigenvalue of a symplectic linear map, so is $\lambda^{-1}$, and they appear with the same multiplicity. If $\lambda$ is a complex eigenvalue, then so are $\lambda^{-1}, \overline{\lambda}, \overline{\lambda}^{-1}$, all with the same multiplicity.*

*Proof.* Let $\Phi$ be a symplectic map. It is not hard to see that :

$$dq_1 \wedge \ldots \wedge dq_n \wedge dp_1 \wedge \ldots \wedge dp_n = \frac{(-1)^{[n/2]}}{n!} \Omega_0 \wedge \ldots \wedge \Omega_0$$

where $[n/2]$ is the integer part of $n/2$. Since $\Phi$ preserves the right hand side of this equation, it must preserves the left hand side, i.e., the volume. Hence $\det \ \Phi = 1$. The rest of the theorem is a consequence of the fact that the characteristic polynomial $C(\lambda)$ of a symplectic transformation $\Phi$ has real coefficients and that, by (56.2), $\Phi^t$ is similar to $\Phi^{-1}$:

$$\Phi^t = J\Phi^{-1}J^{-1}.$$

$\square$

**Dynamical Type of the Origin as Fixed Point.** The origin is a fixed point for any linear map. Its dynamical properties ($i.e.$ how orbits under the iteration of the map of points close to the origin behave) are entirely given by its spectrum. In particular the *stability* ($i.e.$ whether points near 0 stay in a neighborhood of 0) of the fixed point is function of the spectrum: if any eigenvalue is greater than 1 in modulus, the fixed point is unstable. Conversely, if all the eigenvalues are within the unit disk, the fixed point is stable. In symplectic maps, by Theorem 56.3, stability can only occur when all the eigenvalues are on the unit circle. In this case the fixed point is called *elliptic* if all its eigenvalues are distinct from $\pm 1$ or *parabolic* if all its eigenvalues are $\pm 1$. On the other extreme, the origin is a *hyperbolic fixed point* for a symplectic linear map when no eigenvalue is on the unit circle. In this case, we can't have stability since, by Theorem 56.3 necessarily half of the eigenvalues have modulus greater than 1. Hence, in this case, the *stable and unstable manifold* (the $n$–dimensional union of eigensubspaces with eigenvalues larger (resp. smaller) than 1 in absolute value) are each $n$ dimensional. These manifolds are also Lagrangian (see Proposition 36.1 and Exercise 36.3). Note that hyperbolicity can come in different flavors: Inversion hyperbolic (when all eigenvalues are negative) or spiral hyperbolic, when some of the eigenvalues are not real. Finally, a general fixed point exhibits a combination of elliptic, parabolic and hyperbolic behaviors, each in different even dimensional eigenspaces.

**Exercise 56.4** (a) Show that if a $2n\times 2n$ matrix $\Phi$ is given by its $n\times n$ block representation:

$$\Phi = \begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{c} & \boldsymbol{d} \end{pmatrix},$$

then $\Phi$ is symplectic if and only if $\boldsymbol{a}\boldsymbol{b}^t = \boldsymbol{b}\boldsymbol{a}^t, \quad \boldsymbol{c}\boldsymbol{d}^t = \boldsymbol{d}\boldsymbol{c}^t, \quad \boldsymbol{a}\boldsymbol{d}^t - \boldsymbol{b}\boldsymbol{c}^t = Id_n.$
(b) Show that

$$\Phi^{-1} = \begin{pmatrix} \boldsymbol{d}^t & -\boldsymbol{b}^t \\ -\boldsymbol{c}^t & \boldsymbol{a}^t \end{pmatrix}.$$

In particular, if $\Phi$ is symplectic, so are $\Phi^{-1}$ and $\Phi^t$ (this can also be shown directly from (56.2) ).

**Exercise 56.5** The groups of $2n \times 2n$ real matrices $Gl(n, \mathbb{C})$ and $O(2n)$ are defined by:

$$\Phi \in Gl(n, \mathbb{C}) \Leftrightarrow \Phi J = J\Phi; \quad \Phi \in O(2n) \Leftrightarrow \Phi^t \Phi = Id$$

Show that if $\Phi$ is in any two of the groups $Sp(2n), O(2n), Gl(n, \mathbb{C})$, it is in the third. Show that, in this case, we can write:

$$\Phi = \begin{pmatrix} \boldsymbol{a} & -\boldsymbol{b} \\ \boldsymbol{b} & \boldsymbol{a} \end{pmatrix} \quad \text{with} \quad \begin{cases} \boldsymbol{a}^t\boldsymbol{b} = \boldsymbol{b}^t\boldsymbol{a} \\ \boldsymbol{a}^t\boldsymbol{a} + \boldsymbol{b}^t\boldsymbol{b} = Id \end{cases}$$

that is, the complex matrix $\boldsymbol{a} + i\boldsymbol{b}$ is in the *unitary group* $U(n)$.

**Exercise 56.6** (a) Show that, when $\pm 1$ is an eigenvalue of $\Phi \in Sp(n)$, it must appear with even multiplicity.
(b) Show that if $\lambda, \lambda'$ are eigenvalues of $\Phi$ with eigenvectors $\boldsymbol{v}, \boldsymbol{v}'$ and $\lambda\lambda' \neq 1$ then $\Omega_0(\boldsymbol{v}, \boldsymbol{v}') = 0$.
(c) Deduce from (b) that, if $\Phi$ is hyperbolic, its (un)stable manifold is Lagrangian.

**Exercise 56.7** (a) Show that any nonsingular, real matrix $\Phi$ has the *polar decomposition*: $\Phi = PO$ where $P = (\Phi\Phi^t)^{1/2}$ is symmetric positive definite, and $O = \Phi P^{-1}$ is orthogonal (*Hint.* to make sense of $(\Phi\Phi^t)^{1/2}$, note that $\Phi\Phi^t$ is symmetric, positive, and hence diagonalizable with positive diagonal terms).
(b) Show that if $\Phi$ is symplectic, then $P$ and $O$ are also symplectic.(*Hint.* Prove it for $P$ by decomposing $\mathbb{R}^{2n}$ into eigenspaces for $\Phi\Phi^t$. Notice, in particular, that $O \in U(n)$, by Exercise 56.5.
(c) Show more generally that $(\Phi\Phi^t)^\alpha$ is symplectic for all real $\alpha$, and deduce from this that $U(n)$ is a deformation retract of $Sp(2n)$.

## 57. Symplectic Manifolds

Let $N$ be a differentiable manifold. A *symplectic structure* on $N$ is a family of symplectic forms on the tangent spaces of $N$ which depends smoothly on the base point and has a certain nondegeneracy condition. More precisely, a symplectic structure is given by a closed nondegenerate differential 2–form $\Omega$:

$$d\Omega = 0 \text{ and, for all } z \in M, \boldsymbol{v} \neq 0 \in T_zM, \exists \boldsymbol{w} \in T_zM \text{ such that } \Omega(\boldsymbol{v}, \boldsymbol{w}) \neq 0.$$

$\Omega$ is called a *symplectic form* and $(M, \Omega)$ a *symplectic manifold*. A *symplectic map* or *symplectomorphism* between two symplectic manifolds $(N_1, \Omega_1)$ and $(N_2, \Omega_2)$ is a differentiable map $F : N_1 \to N_2$ such that:

$$F^*\Omega_2 = \Omega_1.$$

In other words, the tangent space at each point of a symplectic manifold is a symplectic vector space, and the differential of a symplectic map at a point is a symplectic linear map between symplectic vector spaces.

**Example 57.1** (a) Once again , the canonical example of symplectic manifold is given by $(\mathbb{R}^{2n}, \Omega_0)$, where $\mathbb{R}^{2n}$ is thought of as a manifold. The tangent space at a point is identified with $\mathbb{R}^{2n}$ itself, and the form $\Omega_0$ is a constant differential form on this manifold.
(b) Any surface with its area form is a symplectic manifold. Symplectic maps in dimension $2$ are just area preserving maps.
(c) Kähler manifolds (see McDuff & Salamon (1996)) are symplectic.
(d) Cotangent bundles are non compact symplectic manifolds (see Section $58$) and time 1 maps of Hamiltonian vector fields on them are symplectic maps.

The fundamental theorem by Darboux (of which we have proven the linear version) says that locally, all symplectic manifolds are isomorphic to $(\mathbb{R}^{2n}, \Omega_0)$. See Arnold (1978), Weinstein (1979) or McDuff & Salamon (1996) for a proof of this.

**Theorem 57.2 (Darboux)** *Let $(N, \Omega)$ be a symplectic manifold. Around each point of $N$, one can find a coordinate chart $(\boldsymbol{q}, \boldsymbol{p})$ such that :*

$$\Omega = \sum_1^n dq_k \wedge dp_k := d\boldsymbol{q} \wedge d\boldsymbol{p}.$$

Hence all $2n$–dimensional symplectic manifolds are locally symplectomorphic. This is in sharp contrast with Riemannian geometry where, for example, the curvature is an obstruction for two manifolds to be locally isometric.

**Submanifolds Of Symplectic Manifolds.** Submanifolds of a symplectic manifold can inherit the qualitative features of their tangent spaces: A submanifold $Z \subset (N, \Omega)$ is *(co)isotropic* if each of its tangent spaces is (co)isotropic in the symplectic tangent space of $N$. Hence a *Lagrangian submanifold* is an isotropic submanifold of dimension $n = \frac{1}{2} dim N$. Any curve on a surface is a Lagrangian submanifold. The 0-section and the fiber of the cotangent bundle of a manifold are Lagrangian submanifolds, and so is the graph of any closed differential form (see next Section $58$.C).

**Exercise 57.3** Show the following:
(a) Any symplectic manifold has even dimension.

(b) If $(N, \Omega)$ is a $2n$ dimensional symplectic manifold, then $\Omega^n$ is a volume form.
(c) A symplectomorphism is a volume preserving diffeomorphism.

**Exercise 57.4** Let $(N, \Omega)$ be a symplectic manifold and $F : N \to N$ a symplectomorphism. Show that the set *graph* $F$ is a Lagrangian submanifold of $(N \times N, \Omega \oplus (-\Omega))$

# 58. Cotangent Bundles

## A. Some Definitions

Let $M$ be a differentiable manifold of dimension $n$. Its *cotangent bundle* $T^*M \xrightarrow{\pi} M$ is the fiber bundle whose fiber $T^*_{\boldsymbol{q}}M$ at a point $\boldsymbol{q}$ of M is the dual to the fiber $T_{\boldsymbol{q}}M$ of the tangent bundle. The elements of $T^*_{\boldsymbol{q}}M$ are *cotangent vectors* or linear 1–forms, based at $\boldsymbol{q}$. Given local coordinates $(q_1, \ldots, q_n)$ in a chart of $M$ , one usually denotes a tangent vector $\boldsymbol{v}$ by:

$$\boldsymbol{v} = \sum_1^n v_k \frac{\partial}{\partial q_k}$$

where $\frac{\partial}{\partial q_k}$ denotes the tangent vector to the $k$ th coordinate line at the point $\boldsymbol{q}$ considered. A cotangent vector $\boldsymbol{p}$ at the point $\boldsymbol{q}$ takes the form:

$$\boldsymbol{p} = \sum_1^n p_k dq_k \tag{58.1}$$

Where $dq_k$ denotes the 1–form dual to $\frac{\partial}{\partial q_k}$:

$$dq_j\left(\frac{\partial}{\partial q_k}\right) = \delta_{jk} \overset{\text{def}}{=} \begin{cases} 0 \text{ if } j \neq k, \\ 1 \text{ if } j = k \end{cases} .$$

Once the system of coordinates $\boldsymbol{q} = (q_1, \ldots, q_n)$ is chosen, the coordinates $\boldsymbol{p} = (p_1, \ldots, p_n)$ for $T^*_q M$ as defined by (58.1) are uniquely determined, and we call them the *conjugate coordinates*. We will also refer to the pair $(\boldsymbol{q}, \boldsymbol{p})$ as a chart of conjugate coordinates. The cotangent bundle $T^*M$ as a smooth union of the fibers $T^*_q M$ is a differentiable manifold of dimension $2n$, with local coordinates $(\boldsymbol{q}, \boldsymbol{p})$ as presented above. More precisely, if $\boldsymbol{Q} \xrightarrow{\Psi} \boldsymbol{q}$ is a coordinate change between two charts $U$ and $V$ of $M$, above which $T^*M$ is trivial, then :

$$\Psi^*(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{Q}, \boldsymbol{P}) = \left(\Psi^{-1}(\boldsymbol{q}), D\Psi^t_{\boldsymbol{q}}(\boldsymbol{p})\right)$$

is a change of coordinates between the corresponding charts $V \times \mathbb{R}^n$ and $U \times \mathbb{R}^n$ of $T^*M$. This law of change of coordinates is what distinguishes tangent vectors from cotangent vectors. More generally, given a (local) diffeomorphism $F : M \to N$ between two manifolds $M$ and $N$, there is (locally) an induced *pull–back map*: $F^* : T^*N \to T^*M$ which can be written $F^*(\boldsymbol{q}, \boldsymbol{p}) = (F^{-1}(\boldsymbol{q}), DF^t_{\boldsymbol{q}}(\boldsymbol{p}))$ in conjugate coordinates.

**Example 58.1** (a)$\mathbb{R}^{2n} \cong \mathbb{R}^n \oplus (\mathbb{R}^n)^*$ can be seen as the cotangent bundle of the manifold $\mathbb{R}^n$: this bundle is trivial, as any bundle over a contractible manifold.
(b)The cotangent bundle of $\mathbb{T}^n$ is $\mathbb{T}^n \times \mathbb{R}^n$. That $T^*\mathbb{T}^n$ is trivial is a consequence of the fact that $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$, where $\mathbb{Z}^n$ acts as a group of translations on $\mathbb{R}^n$, whose differentials are the $Id$. See the following exercise.

**Exercise 58.2** More generally, if $M \cong \mathbb{R}^n/\Gamma$ where $\Gamma$ is a group of diffeomorphisms of $\mathbb{R}^n$ acting *properly discontinuously* (i.e. around each point $q$ of $M$ there is a neighborhood $U(q)$ such that $U \cap (\Gamma \backslash Id)(U) = \emptyset$), then

$$T^*M \cong \mathbb{R}^{2n}/\Gamma^*$$

where $\Gamma^*$ is the set of pullback diffeomorphisms of $\mathbb{R}^{2n}$ of the form $\gamma^*$, where $\gamma \in \Gamma$.

## B. Cotangent Bundles as Symplectic Manifold

We now show that there is a natural symplectic structure on $T^*M$. We first construct a canonical differential 1–form called the *Liouville form* which, as we will prove, has the following expression in any set of conjugate coordinates:

$$\lambda = \sum_1^n p_k dq_k = \boldsymbol{p} d\boldsymbol{q}.$$

We then obtain a symplectic form by differentiating $\lambda$:

$$\Omega = -d\lambda, \quad \Omega = d\boldsymbol{q} \wedge d\boldsymbol{p},$$

the latter holding in any conjugate coordinate system. We first present a coordinate free construction of $\lambda$. To define a 1–form on $T^*M$, it suffices to determine how it acts on any given tangent vector $v$ in a fiber $T_\alpha(T^*M)$ of the tangent space of $T^*M$. Since the base point $\alpha$ is in $T^*M$, it is a linear 1–form on $T^*_{\boldsymbol{q}}M$ for $\boldsymbol{q} = \pi(\alpha)$. Let $\pi : T^*M \to M$ be the canonical projection. The derivative $\pi_* : T(T^*M) \to TM$ takes a vector $\boldsymbol{v}$ to the vector $\pi_*\boldsymbol{v}$ in $T_{\boldsymbol{q}}M$. We can evaluate the 1–form $\alpha$ on that vector, and define:

$$\lambda(\boldsymbol{v}) = \alpha(\pi_*\boldsymbol{v})$$

We now compute $\lambda$ in a local, conjugate coordinate chart $(\boldsymbol{q}, \boldsymbol{p})$ of $T^*M$. Write:

$$\alpha = \sum \alpha_k dq_k \text{ and } \boldsymbol{v} = \boldsymbol{v}_{\boldsymbol{q}} \frac{\partial}{\partial \boldsymbol{q}} + \boldsymbol{v}_{\boldsymbol{p}} \frac{\partial}{\partial \boldsymbol{p}}.$$

Then $\pi_*(\boldsymbol{v}) = \boldsymbol{v}_{\boldsymbol{q}} \frac{\partial}{\partial \boldsymbol{q}}$ and $\alpha(\pi_*\boldsymbol{v}) = \sum \alpha_k v_{q_k}$ which exactly says that $\lambda = \boldsymbol{p}d\boldsymbol{q}$. $\square$

**Exact Symplectic Maps on Cotangent Bundles.** That the symplectic form $\Omega$ is *exact* ($i.e.$ the differential of 1-form, here $\lambda$) on a cotangent bundle enables us to single out the important class of exact symplectic maps. Indeed, if a map $F : T^*M \to T^*M$ is symplectic in $T^*M$ then the form $F^*\lambda - \lambda$ is closed: $d(F^*\lambda - \lambda) = F^*d\lambda - d\lambda = -(F^*\Omega - \Omega) = 0$. This justifies the following

**Definition 58.3** A map $F : T^*M \to T^*M$ is *exact symplectic* if the 1 form $F^*\lambda - \lambda$ is exact:

$$F^*\lambda - \lambda = dS$$

for some real valued function $S$ on $T^*M$.

We will see in Section 59 that time $t$ maps of Hamiltonian flows are exact symplectic. So are most of the maps in this book. Note that in $\mathbb{R}^{2n}$, since any closed form is exact, symplectic and exact symplectic are two equivalent properties. On the other hand, the map $(x, y) \to (x, y + a),\ a \neq 0$, of the cylinder is an example of a map which is symplectic but not exact symplectic.

**Remark 58.4** 1) The term $exact$ diffeomorphism, or even $exact\ symplectic$ diffeomorphism is sometimes used to denote the time 1 map of a (time dependent) Hamiltonian system. But it can be shown that the map $(\boldsymbol{q}, \boldsymbol{p}) \mapsto (\boldsymbol{q} + A\boldsymbol{p}, \boldsymbol{p}),\ A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ is exact symplectic according to our definition, but not isotopic to $Id$ (true more generally whenever $A$ is not homotopic (cannot be deformed) to $I$ on $\mathbb{T}^2$, see Exercise 23.4). Hence these maps cannot be time–1 maps of Hamiltonians.

2) Cotangent bundles are just one example, albeit the most important one, of *exact symplectic manifolds*: symplectic manifolds whose symplectic form is exact. Many facts that are true for cotangent bundles also hold for exact symplectic manifolds.

**Exercise 58.5** Show that the set of exact symplectic maps forms a group under composition. In particular, show that if $G^*\lambda - \lambda = S_G$ and $F^*\lambda - \lambda = S_F$ then

$$(F \circ G)^*\lambda - \lambda = d[(S_F \circ G) + S_G].$$

**Exercise 58.6** Show that a map $F$ of $T^*M$ is exact symplectic if and only if :

$$\int_{F\gamma} \boldsymbol{p} d\boldsymbol{q} = \int_{\gamma} \boldsymbol{p} d\boldsymbol{q}$$

for all differentiable closed curve $\gamma$.

## C. Notable Lagrangian Submanifolds of Cotangent Bundles

It is not hard to see that the fibers of $T^*M$ are Lagrangian submanifolds: in coordinates they are given by $\{\boldsymbol{q} = \boldsymbol{q}_0\}$ and hence their tangent space is of the form $\{\boldsymbol{q} = 0\}$. Likewise, the zero section $0^*_M$ of $T^*M$ is Lagrangian. Another class of example is of importance to us in Chapter 10. Consider a function $W : M \to \mathbb{R}$. Its differential $dW$ can be seen as a section of $T^*M$, i.e. a map $M \to T^*M$ whose image $dW(M)$ can be written as $\{(\boldsymbol{q}, dW(\boldsymbol{q})) \mid \boldsymbol{q} \in M\}$. Hence, a basis for the tangent space of $dW(M)$ at a point $(\boldsymbol{q}, dW(\boldsymbol{q}))$ is given by:

$$\boldsymbol{v}_k = \frac{\partial}{\partial q_k} + \sum_{j=1}^{n} \frac{\partial^2 W(\boldsymbol{q})}{\partial q_j \partial q_k} \frac{\partial}{\partial p_j}$$

It is not hard to see that:

$$\Omega(\boldsymbol{v}_k, \boldsymbol{v}_l) = \frac{\partial^2 W(\boldsymbol{q})}{\partial q_k \partial q_l} - \frac{\partial^2 W(\boldsymbol{q})}{\partial q_l \partial q_k} = 0,$$

so that $dW(M)$ *is a Lagrangian submanifold of* $T^*M$. We can generalize this argument somewhat. Any 1–form $\alpha$ can be seen as a map from $M$ to $T^*M$, so we can ask the question: for what $\alpha$ is $\alpha(M)$ a Lagrangian manifold ? To answer this question, one can check (Exercise 58.7) the following formula:

$$\alpha^*\lambda = \alpha. \tag{58.2}$$

where $\lambda$ is the Liouville form (the reader has to get used to the fact that we see $\alpha$ either as a form or a map, at our convenience. When seen as a map, $\alpha$ is actually an embedding of $M$ into $T^*M$). The manifold $\alpha(M)$ is Lagrangian exactly when:

$$0 = \alpha^* \Omega = \alpha^*(-d\lambda) = -d(\alpha^* \lambda) = -d\alpha,$$

that is, exactly when $\alpha$ is a closed form. In particular, if the form $\alpha$ is exact with $\alpha = dW$, this gives another proof that $dW(M)$ is Lagrangian. $W$ is the simplest instance of *generating function* for the Lagrangian manifold $\alpha(M) = dW(M)$ (*generating phase* or *generating phase function* is also used). We expend on this important notion of symplectic topology in Chapter 10.

**Exercise 58.7** Verify Formula (58.2) using local coordinates.

# 59. Hamiltonian Systems

## A. Lagrangian Systems Versus Hamiltonian Systems

**Euler-Lagrange Equations.** A lot of mechanical systems can be put in terms of a variational problem. In these systems, under the *principle of least action,* trajectories are critical points of an *action functional* of the form:

$$A(\gamma) = \int_{t_0}^{t_1} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) dt,$$

with boundary condition $\gamma(t_0) = \boldsymbol{q}_0, \gamma(t_1) = \boldsymbol{q}_1$ ($\boldsymbol{q}(t)$ is the parameterization of $\gamma$ in the above integral). The function $L$ is twice differentiable in each variables, say (absolute continuity is enough). It is called the *Lagrangian function* of the system. As this is a somewhat heuristic discussion, we will not specify here the functional space to which $\gamma$ belongs. In concrete cases (say $\gamma \in C^1([t_0, t_1]), \mathbb{R}^n)$ or $C^1([t_0, t_1], M)$, or some Sobolev space of parameterized curves), the following can be made quite rigorous.

To compute the differential of $A$, one applies a small variation $\delta\gamma$ to $\gamma$, with $\delta\gamma(t_0) = \delta\gamma(t_1) = 0$. Then:

$$\delta A(\gamma) = \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial \boldsymbol{q}}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \delta \boldsymbol{q} + \frac{\partial L}{\partial \dot{\boldsymbol{q}}}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \delta \dot{\boldsymbol{q}} \right) dt.$$

Performing an integration by parts on the second term of this integral, we get:

$$\delta A(\gamma) = \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial \boldsymbol{q}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\boldsymbol{q}}} \right) \delta \boldsymbol{q} dt$$

For $\gamma$ to be a critical point, the above integral must be 0. Since this should be true for any variation $\delta\gamma$, we must have:

$$\frac{\partial L}{\partial \boldsymbol{q}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\boldsymbol{q}}} = 0, \tag{59.1}$$

which is a second order differential equation in $\boldsymbol{q}$ called the *Euler-Lagrange* equations. (The plural to "equations" just refers to the fact that the dimension is usually greater than 1). As an example, a large number of mechanical systems have a Lagrangian function of the form:

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = \frac{1}{2} \left\| \dot{\boldsymbol{q}} \right\|^2 - V_t(\boldsymbol{q}).$$

("Kinetic - potential". The time dependence of $V$ usually refers to some forcing) where $V : \mathbb{R}^n \to \mathbb{R}$. The Euler–Lagrange equations for such a system are:

$$\ddot{\boldsymbol{q}} + \nabla V_t(\boldsymbol{q}) = 0, \tag{59.2}$$

which, in mechanical cases, can be seen as the expression of Newton's Law: acceleration is proportional to the the force.

**From Lagrange to Hamilton.** To solve the O.D.E. (59.2) , one usually proceeds by introducing $\boldsymbol{p} = \dot{\boldsymbol{q}}$ to get a system of first order ODE's:

$$\begin{aligned} \dot{\boldsymbol{q}} &= \boldsymbol{p} \\ \dot{\boldsymbol{p}} &= -\nabla V(\boldsymbol{q}). \end{aligned}$$

As we will see presently, we have just put the Lagrangian problem into a Hamiltonian form. In general, if the following *Legendre Condition:*

$$\det \frac{\partial^2 L}{\partial \dot{\boldsymbol{q}}^2} \neq 0, \tag{59.3}$$

is satisfied, we can introduce

$$\boldsymbol{p} = \frac{\partial L}{\partial \dot{\boldsymbol{q}}}$$

to transform the Euler–Lagrange equations (59.1) into a system of first order O.D.E.'s: because of the nondegeneracy condition (59.3), the implicit function theorem implies that, locally, we can make a change of variables :

$$\mathcal{L} : (\boldsymbol{q}, \dot{\boldsymbol{q}}) \to (\boldsymbol{q}, \boldsymbol{p} = \frac{\partial L}{\partial \dot{\boldsymbol{q}}}) \tag{59.4}$$

This is, when $\boldsymbol{q}$ is seen as a point on a manifold $M$, a local diffeomorphism between $T_{\boldsymbol{q}}M$ and $T^*_{\boldsymbol{q}}M$ (see Exercise 59.2 ). This change of variables is called the *Legendre transformation*.[21] Define the *Hamiltonian function* by:

$$H(\boldsymbol{q}, \boldsymbol{p}, t) = \boldsymbol{p}\dot{\boldsymbol{q}} - L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t),$$

Where it is understood that $\dot{\boldsymbol{q}} = \dot{\boldsymbol{q}} \circ \mathcal{L}^{-1}(\boldsymbol{q}, \boldsymbol{p})$. We can compute:

$$\begin{aligned}
\frac{\partial H}{\partial \boldsymbol{q}} &= \boldsymbol{p}\frac{\partial \dot{\boldsymbol{q}}}{\partial \boldsymbol{q}} - \frac{\partial L}{\partial \boldsymbol{q}} - \frac{\partial L}{\partial \dot{\boldsymbol{q}}}\frac{\partial \dot{\boldsymbol{q}}}{\partial \boldsymbol{q}} = -\frac{\partial L}{\partial \boldsymbol{q}}, \\
\frac{\partial H}{\partial \boldsymbol{p}} &= \dot{\boldsymbol{q}} + \boldsymbol{p}\frac{\partial \dot{\boldsymbol{q}}}{\partial \boldsymbol{p}} - \frac{\partial L}{\partial \dot{\boldsymbol{q}}}\frac{\partial \dot{\boldsymbol{q}}}{\partial \boldsymbol{p}} = \dot{\boldsymbol{q}}.
\end{aligned}$$

But the Euler-Lagrange equations imply that:

$$\dot{\boldsymbol{p}} = \frac{d}{dt}\frac{\partial L}{\partial \dot{\boldsymbol{q}}} = \frac{\partial L}{\partial \boldsymbol{q}} = -\frac{\partial H}{\partial \boldsymbol{q}}.$$

Combining this with the previous formula yields *Hamilton's equations*:

$$\begin{aligned}
\dot{\boldsymbol{q}} &= H_{\boldsymbol{p}} \\
\dot{\boldsymbol{p}} &= -H_{\boldsymbol{q}}.
\end{aligned} \tag{59.5}$$

**Remark 59.1** (a) The Legendre transformation is *involutive*: it is its own inverse, in the following sense. The map $(\boldsymbol{q}, \dot{\boldsymbol{q}}) \to (\boldsymbol{q}, \frac{\partial L}{\partial \dot{q}} = \boldsymbol{p})$ has inverse:

$$(\boldsymbol{q}, \boldsymbol{p}) \to (\boldsymbol{q}, \frac{\partial H}{\partial \boldsymbol{p}} = \dot{\boldsymbol{q}})$$

and $L$ is the Legendre transformed of $H$ in the sense that:

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = \boldsymbol{p}\dot{\boldsymbol{q}} - H(\boldsymbol{q}, \boldsymbol{p}, t)$$

where $\boldsymbol{p} = \boldsymbol{p}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ is given implicitly by $\frac{\partial H}{\partial \boldsymbol{p}} = \dot{\boldsymbol{q}}$ (to see that this is all legal, see Exercise 59.3which shows that if $L$ satisfies the Legendre condition, so does $H$).

[21] In the classical literature the term Legendre transformation refers to the complete process of changing the Lagrangian $L$ into the Hamiltonian $H$ as shown in this section, and $H$ is then called the Legendre transformation of $L$. It is grammatically less awkward to call $H$ the Legendre *transformed* of $L$, which we do in this book.

(b) In the new coordinates, the action functional becomes:

$$A(\gamma) = \int_\gamma \boldsymbol{p} d\boldsymbol{q} - H dt$$

where $\gamma$ is seen as a curve $(\boldsymbol{q}(t), \boldsymbol{p}(t), t)$ in the extended phase space $\mathbb{R}^{2n} \times \mathbb{R}$, or $T^*M \times \mathbb{R}$.

**Symplectic formulation of Hamilton's Equations.** Hamilton's equations have a natural expression in the symplectic setting. We assume for now that $\boldsymbol{q}$ is in $\mathbb{R}^n$. Using the notation $H_t(\boldsymbol{q}, \boldsymbol{p}) = H(\boldsymbol{q}, \boldsymbol{p}, t)$, we can rewrite (59.5) as

$$\dot{z} = -J\nabla H_t(z) \stackrel{def}{:=} X_H(z, t).$$

where $\nabla H_t = \binom{H_q}{H_p}$ is the gradient of $H_t$ with respect to the scalar product on $\mathbb{R}^{2n}$:

$$\langle \nabla H_t, \boldsymbol{v} \rangle = dH_t(\boldsymbol{v})$$

Likewise, $X_H$, which we call the *Hamiltonian vector field* can be seen as the *symplectic gradient* of $H_t$:

$$\Omega_0(X_H, \boldsymbol{v}) = \langle -J^2 \nabla H_t, \boldsymbol{v} \rangle = \langle \nabla H_t, \boldsymbol{v} \rangle = dH_t(\boldsymbol{v}).$$

This can also be written using the contraction operator on differential forms:

$$i_{X_H} \Omega = dH_t$$

**Exercise 59.2** Show that, if $\phi : U \to V$ is a change of coordinate charts in a manifold $M$, then $\frac{\partial L}{\partial \dot{q}}$ changes according to $\phi^* : V \to U$. Hence $\frac{\partial L}{\partial \dot{q}}$ is a covector.

**Exercise 59.3** (a) Compute the Legendre transformed of $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = \frac{1}{2}\langle A\dot{\boldsymbol{q}}, \dot{\boldsymbol{q}} \rangle - V(\boldsymbol{q})$.
(b) Show that, in general, if $H$ is the Legendre transformed of $L$, then

$$L_{\dot{q}\dot{q}} H_{pp} = Id.$$

## B. Hamiltonian Systems on Symplectic Manifolds

Motivated by the last expression that we found for the Hamiltonian vector field in $\mathbb{R}^{2n}$, we extend the definition to symplectic manifolds:

**Definition 59.4** Let $(N, \Omega)$ be a symplectic manifold and $H(z,t) = H_t(z)$ be a $C^k$ real valued function on $N \times \mathbb{R}$. The *Hamiltonian vector field* associated with $H$ is the (time dependent) vector field $X_H$ defined by:

$$\Omega(X_H, \boldsymbol{v}) = dH_t(\boldsymbol{v}), \quad \forall \boldsymbol{v} \in TM.$$

Equivalently:

$$i_{X_H}\Omega = dH_t$$

The (possibly time dependent) O.D.E.:

$$\dot{\boldsymbol{z}} = X_H(\boldsymbol{z}, t) \tag{59.6}$$

is called *Hamilton's equations.*

In local Darboux coordinate charts (*eg.* in conjugate coordinates chart of a cotangent bundles), these equations take the form of (59.5) (see Exercise 59.5). If $H$ is time independent, then (59.6) generates a (local) flow on $N$. If $H$ is time dependent, then $X_H$ generates a (local) flow in the space $N \times \mathbb{R}$, called the *extended phase space* in mechanics. Specifically, one solves the following time independent system on $N \times \mathbb{R}$:

$$\begin{aligned}\dot{\boldsymbol{z}} &= X_H(\boldsymbol{z}, s)\\ \dot{s} &= 1\end{aligned}$$

which generates a flow $\phi^t$ in $N \times \mathbb{R}$ satisfying:

$$\phi^t(\boldsymbol{z}, s) = (h_s^{t+s}(\boldsymbol{z}), s+t),$$

where $h_s^t$ is a family of $C^{k-1}$ diffeomorphisms of $N$, $C^{k-1}$ smooth in $s$ and $t$. This is a general procedure for time dependent vector field. The diffeomorphism $h_s^t$ is called a *Hamiltonian map* and, for each fixed $s$ the curve $t \to h_s^t$ is a *Hamiltonian isotopy* (an *isotopy* is a smoothly varying 1–parameter family of diffeomorphisms). Another way of describing $h_s^t(\boldsymbol{z})$ is by saying that it is the unique solution $\boldsymbol{z}(t)$ of Hamilton's equation with initial condition $\boldsymbol{z}(s) = \boldsymbol{z}$. In practice, one often fixes $s = 0$ and denotes $h_0^t$ by $h^t$. The following exercise shows the one to one correspondence between time dependent vector fields and isotopies. It also shows that, even though the time $0$ of a solution flow to a time dependent vector field is the Identity, the maps $h^t$ do not in general form a 1-parameter group.

**Exercise 59.5** Show that the equation (59.6) takes the form of (59.5) in local coordinates.

**Exercise 59.6** Let $X_t$ be a vector field (not necessarily Hamiltonian) on a manifold $N$. Let $h_s^t$ be the solution flow to the O.D.E. $\dot{\boldsymbol{z}} = X_t(\boldsymbol{z}), \dot{s} = 1$. Prove that:
(i) $h_s^s = Id, \ \forall s \in \mathbb{R}$,
(ii)$h_s^{t'} = h_t^{t'} \circ h_s^t$, so that in particular $h_s^t = h^t \circ (h^s)^{-1}$. Compute $(h_s^t)^{-1}$ and deduce that, in general, $(h_0^t)^{-1}$ may not be equal to $(h_0^r)$ for any $r$: the maps $h^t$ do not form a group.
(iii) Conversely, given any (sufficiently smooth) isotopy $g^t$ in $N$, with $g^0 = Id$, show that the time dependent vector field:

$$\dot{g}^t(\boldsymbol{z}) = \frac{d}{du}\bigg|_{u=0} g^{t+u} \circ (g^t)^{-1}(\boldsymbol{z}) = \frac{dg^t}{dt}\left((g^t)^{-1}(\boldsymbol{z})\right)$$

has solution $h_s^t = g^t \circ (g^s)^{-1}$.

## C. Invariants of the Hamiltonian Flow

**Conservation of Energy.** If $G$ is a function on a differentiable manifold $N$, and $X$ is a vector field, we recall that the *Lie derivative* of $G$ along $X$ is:

$$L_X G(z) = \frac{d}{dt}\bigg|_{t=0} G(\phi^t(z)) = dG(X(z))$$

where $\phi^t$ is the flow solution for $X$ .

**Theorem 59.7** *Let $H$ be a time independent Hamiltonian function on $(M, \Omega)$ . Then $H$ is constant under the Hamiltonian flow it generates:*

$$L_{X_H} H = 0.$$

*Proof.* $L_{X_H} H = dH(X_H) = \Omega(X_H, X_H) = 0$ □

**Remark 59.8** $L_{X_H}G = \Omega(X_G, X_H) = -L_{X_G}H$ is also denoted by $\{G, H\}$ and it is called the *Poisson bracket* of $H$ and $G$. Hence, the poisson bracket measures how far the function $G$ (resp. $H$) is from being constant along the flow of $X_H$ (resp.$X_G$). When $\{H, G\} = 0$, one says that $G$ (resp. $H$) is a *first integral* of the Hamiltonian flow of $H$ (resp. $G$), or that the functions $H$ and $G$ are *in involution*. One can show (see *eg.* Arnold (1978), Abraham & Marsden (1985)) that the set of Hamiltonian vector fields form a Lie sub–algebra of the Lie algebra of vector fields on a manifold, in the sense that:

$$X_{\{H,G\}} = [X_H, X_G].$$

In particular, the poisson bracket of two functions measures how far from commuting their Hamiltonian flows are.

**Hamiltonian Flows are Exact Symplectic.** We extend the notion of Lie derivative to differential forms. If $X$ is any (possibly time dependent) vector field, and $\alpha$ a differential form, we define *the Lie derivative of $\alpha$ in the direction of $X$* by:

$$L_X\alpha = \left.\frac{d}{dt}\right|_{t=0} (g^t)^*\alpha.$$

where $g^t$ is the flow generated by $X$. At time $t \neq 0$,

$$\frac{d}{dt}(g^t)^*\alpha = (g^t)^* L_X\alpha. \tag{59.7}$$

The isotopy $g^t$ preserves the form $\alpha$ whenever the Lie derivative is zero:

$$(g^t)^*\alpha = \alpha,\ \forall t \Longleftrightarrow L_X\alpha = 0.$$

We have the important *homotopy formula* (see *eg.* Weinstein (1979) or McDuff & Salamon (1996)):

$$L_X\alpha = i_X d\alpha + d(i_X\alpha) \tag{59.8}$$

and, at time $t \neq 0$,

$$\frac{d}{dt}(g^t)^*\alpha = (g^t)^*\left(i_X d\alpha + d(i_X\alpha)\right)$$

A *symplectic isotopy* $g^t$ on $(M, \Omega)$ is an isotopy such that $g^t$ is a symplectic map for all $t$. By the homotopy formula (and the fact that a symplectic form is closed), this can be reworded:

$$g^t \text{ is a symplectic isotopy} \Longleftrightarrow L_X\Omega = 0 \Longleftrightarrow d(i_X\Omega) = 0 \tag{59.9}$$

where $X$ is the (time dependent) vector field $\frac{dg^t}{dt}\left((g^t)^{-1}(\boldsymbol{z})\right)$ (see Exercise 59.6). The following theorem characterizes Hamiltonian isotopies, at least in cotangent bundles (or in any exact symplectic manifold, i.e. one whose symplectic form is exact)

**Theorem 59.9** *(a) On any symplectic manifold, Hamiltonian isotopies are symplectic.*
*(b) On a cotangent bundle $T^*M$, a Hamiltonian isotopy with Hamiltonian $H(\boldsymbol{z},t)$ is also exact symplectic:*

$$h^{t*}\lambda - \lambda = h^{t*}\boldsymbol{p}d\boldsymbol{q} - \boldsymbol{p}d\boldsymbol{q} = dS_t, \quad \text{with } S_t = \int_\gamma \boldsymbol{p}d\boldsymbol{q} - Hd\tau$$

*where $\gamma$ is the curve $(h^\tau(\boldsymbol{z}),\tau)$, $\tau \in [0,t]$ solution of Hamilton's equations in the extended phase space $T^*M \times \mathbb{R}$, and $\boldsymbol{z}$ is the point at which the form is evaluated.*
*(c) Conversely, if an isotopy $g^t$ is exact symplectic then it is Hamiltonian, with the Hamiltonian function given by:*

$$H_t = i_{X_t}\boldsymbol{p}d\boldsymbol{q} - ((g^t)^{-1})^*\frac{d}{dt}(S_t).$$

*where $X_t(\boldsymbol{z}) = \frac{dg^t}{dt}((g^t)^{-1}(\boldsymbol{z}))$.*

*Proof*. The first assertion (a) is an immediate consequence of (59.9) : if $h^t$ is a Hamiltonian isotopy then $i(\dot{h}_t)\Omega = dH_t$ is exact, and therefore closed. In cotangent bundles, it is also a consequence of the second assertion. Now, look for $\frac{d}{dt}(S_t)$ in (b):

$$\frac{d}{dt}h_t^*\lambda = h_t^*\left(i_{X_H}d\lambda + d(i_{X_H}\lambda)\right) = h_t^*d(-H_t + i_{X_H}\lambda) = dh_t^*(-H_t + i_{X_H}\lambda) \tag{59.10}$$

From this we get:

$$h_t^*\lambda - \lambda = d\int_0^t h_\tau^*(-H_\tau + i_{X_H}\lambda)d\tau \overset{\text{def}}{=} dS_t \tag{59.11}$$

that is, $h^t$ is exact symplectic. We leave it to the reader to rewrite the integral as the one advertised in the theorem. This finishes the proof of (b). To prove the converse (c), let $g^t$ be an exact symplectic isotopy:

$$(g^t)^*\boldsymbol{p}d\boldsymbol{q} - \boldsymbol{p}d\boldsymbol{q} = dS_t$$

for some $S_t$ differentiable in all of $(\boldsymbol{q}, \boldsymbol{p}, t)$. We claim that the (time dependent ) vector field $X_t(\boldsymbol{z})$ whose flow is $g^t$, is Hamiltonian. To see this, we differentiate (59.11) :

$$\frac{d}{dt}(dS_t) = \frac{d}{dt}(g^t)^*\boldsymbol{p}d\boldsymbol{q} = (g^t)^* L_{X_t}\boldsymbol{p}d\boldsymbol{q} = (g^t)^* \left(i_{X_t}d(\boldsymbol{p}d\boldsymbol{q}) + d(i_{X_t}\boldsymbol{p}d\boldsymbol{q})\right),$$

from which we get

$$i_{X_t}d\boldsymbol{q} \wedge d\boldsymbol{p} = d\left(i_{X_t}\boldsymbol{p}d\boldsymbol{q} - ((g^t)^{-1})^*\frac{d}{dt}(S_t)\right) = dH_t,$$

which exactly means that $X_t$ is Hamiltonian with $H_t$ as Hamiltonian function. □

**Integral Invariant of Poincaré-Cartan.** A less formal proof of (b) in the above theorem yields extra information. We follow Chapter 9 in Arnold (1978). We first prove that the vector field $(X_H, 1)$ in $T^*M \times \mathbb{R}$ generates the kernel of the form $d(\boldsymbol{p}d\boldsymbol{q} - Hdt) = d\boldsymbol{p} \wedge \boldsymbol{dq} - H_{\boldsymbol{q}}d\boldsymbol{q} \wedge dt - H_{\boldsymbol{p}}d\boldsymbol{p} \wedge dt$. The matrix of this form in the (Darboux) coordinate $(\boldsymbol{q}, \boldsymbol{p}, t)$ is:

$$A = \begin{pmatrix} 0 & -Id & H_{\boldsymbol{q}} \\ Id & 0 & H_{\boldsymbol{p}} \\ -H_{\boldsymbol{q}} & -H_{\boldsymbol{p}} & 0 \end{pmatrix},$$

where $H_{\boldsymbol{q}}, H_{\boldsymbol{p}}$ are column vectors on the right and row vectors on the bottom of the matrix. Since the upper left $2n \times 2n$ matrix is the nonsingular matrix $J$, $A$ is of rank (at least) $2n$. It is easy to see that the Hamiltonian vector field $(H_{\boldsymbol{p}}, -H_{\boldsymbol{q}}, 1) = (X_H, 1)$ generates its kernel. Now, take a closed curve $\gamma$ in $T^*M \times \mathbb{R}$. The image under the Hamiltonian flow of $\gamma$ forms an embedded tube in $T^*M \times \mathbb{R}$. Since the tangent space to this tube at any of its point $\boldsymbol{z}$ contains the vector $(X_H(\boldsymbol{z}), 1)$, the form $d(\boldsymbol{p}d\boldsymbol{q} - Hdt)$ restricted to this tube is null. As a result, because of Stokes' theorem, if $\gamma_1$ and $\gamma_2$ in $T^*M \times \mathbb{R}$ encircle the same tube of orbits of the extended flow, we must have:

$$\int_{\gamma_1} \boldsymbol{p}d\boldsymbol{q} - Hdt = \int_{\gamma_2} \boldsymbol{p}d\boldsymbol{q} - Hdt \tag{59.12}$$

since $\gamma_1 - \gamma_2$ is the boundary of a region of the tube. The form $\boldsymbol{p}d\boldsymbol{q} - Hdt$ is called the *integral invariant of Poincaré–Cartan.* As a particular case, if $\gamma_1$ is of the form $(\gamma, t_1)$ and $\gamma_2 = (h_{t_1}^{t_2}\gamma, t_2)$, the form $Hdt$ is null on these curves and hence Equation (59.12) reads:

$$\int_{\gamma_1} \boldsymbol{p}d\boldsymbol{q} = \int_{h_{t_1}^{t_2}\gamma} \boldsymbol{p}d\boldsymbol{q} \tag{59.13}$$

This last equation implies the statement (b) in Theorem 59.9: it proves that the function

$$S_t = \int_{\boldsymbol{z}_0}^{\boldsymbol{z}} h^{t*}\boldsymbol{p}d\boldsymbol{q} - \boldsymbol{p}d\boldsymbol{q} \tag{59.14}$$

is well defined, i.e. the integral does not depend on the path chosen between $z_0$ and $z$. This proves in turn that $h^t$ is exact symplectic. □

**Conservation of Hamilton's Vector Field Under Symplectic Maps.** The property that Hamilton's equations are preserved under symplectic maps characterizes these maps, which are for this reason called *canonical transformations* in some of the classical literature. Even though we do not need this theorem in this book, we include it here since it explains why symplectic geometry came to exist.

**Theorem 59.10** *Let $F : (M, \omega_M) \to (N, \omega_N)$ be a diffeomorphism. Then $F$ is symplectic if and only if for all function $H : N \to \mathbb{R}$,*

$$F_* X_{H\circ F} = X_H. \tag{59.15}$$

*In this case, $F$ conjugates the Hamiltonian flows $h^t$ of $H$ and $g^t$ of $H \circ F$:*

$$g^t = F^{-1} \circ h^t \circ F.$$

*This holds also when $H$ is time dependent.*

*Proof.* Reminding the reader that by definition $F_*X(F(z)) = DF_z X(z)$ for any vector field $X$, we also use the notation $F^*Y$ to mean $(F^{-1})_*Y$. It is not hard to check that the following formula holds:

$$F^* i_X \alpha = i_{F^*X} F^* \alpha \tag{59.16}$$

for any vector field $X$ and differential form $\alpha$. Coming back to our statement, we have on one hand:

$$F^* i_{X_H} \omega_N = F^* dH = dH \circ F,$$

and on the other hand,

$$F^* i_{X_H} \omega_N = i_{F^* X_H} F^* \omega_N = i_{F^* X_H} \omega_M$$

because of (59.16) and the fact that $F$ is symplectic. This proves (59.15) . Conversely, if (59.15) holds for any $H$, the same kind of computation shows that,

$$i_{X_{H\circ F}} F^* \omega_N = i_{X_{H\circ F}} \omega_M$$

and since any tangent vector at a point of $M$ is of the form $X_{H\circ F}$ for some $H$, we must have $F^*\omega_N = \omega_M$, i.e. $F$ is symplectic. The conjugacy statement, a general fact about O.D.E.'s, is left to the reader, as well as checking that everything still works with time dependent systems. □

**Exercise 59.11** The Lie derivative of a function can be defined, in the obvious way, along any differentiable isotopy. What fails in Theorem 59.7 when $H$ is time dependent?

**Exercise 59.12** Show that in Darboux coordinates:

$$\{H, G\} = \frac{\partial H}{\partial \boldsymbol{q}} \frac{\partial G}{\partial \boldsymbol{p}} - \frac{\partial H}{\partial \boldsymbol{p}} \frac{\partial G}{\partial \boldsymbol{q}}.$$

**Exercise 59.13** Prove that the function $S_t$ defined in (59.14) satisfies:

$$S_t(\boldsymbol{z}) = \int_\gamma \boldsymbol{p} d\boldsymbol{q} - H dt \quad + C(\boldsymbol{z}_0, t),$$

for some $C$, and $\gamma$ as in Theorem 59.12. (*Hint.* Apply Stokes on the appropriate surface.)

**Exercise 59.14** Prove that $h_s^t$ is exact symplectic (i.e. even for $s \neq 0$), where $h_s^t(\boldsymbol{z})$ is, as in subsection B, the solution of Hamilton's equation such that $\boldsymbol{z}(s) = \boldsymbol{z}$.

**Exercise 59.15** Let $H$ be autonomous, or of period $\tau$. Show that $X_H(\boldsymbol{z})$ is preserved by $Dh^\tau(\boldsymbol{z})$, *i.e.* $X_H$ is an eigenvector of $Dh^\tau$ with eigenvalue 1.