BASIC SYMPLECTIC GEOMETRY

NOTATION.

$\mathbb{R}^n = n$-dimensional Euclidean space

x, y denote typical points in $\mathbb{R}^n$: $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$

$\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$

$z = (x, \xi), w = (y, \eta)$ denote typical points in $\mathbb{R}^n \times \mathbb{R}^n$:
$z = (x_1, \dots, x_n, \xi_1, \dots, \xi_n)$, $w = (y_1, \dots, y_n, \eta_1, \dots, \eta_n)$

$\mathbb{C} =$ complex plane

$\mathbb{C}^n =$ n-dimensional complex space

$\langle x, y\rangle = \sum_{i=1}^n x_i \bar{y}_i =$ inner product on $\mathbb{C}^n$
$\mathbb{M}^{m\times n} = m \times n$-matrices

$A^T =$ transpose of the matrix A

I denotes both the identity matrix and the identity mapping.

$J = \begin{pmatrix} O & I \\ -I & O \end{pmatrix}$

$\sigma(z, w) = \langle Jz, w\rangle =$ symplectic inner product

If $\varphi : \mathbb{R}^n \to \mathbb{R}$, then we write

$$\partial\varphi := (\varphi_{x_1}, \dots, \varphi_{x_n}) = \text{gradient},$$

and

$$\partial^2\varphi := \begin{pmatrix} \varphi_{x_1x_1} & \cdots & \varphi_{x_1x_n} \\ & \ddots & \\ \varphi_{x_nx_1} & \cdots & \varphi_{x_nx_n} \end{pmatrix} = \text{Hessian matrix}$$

If φ depends on both the variables $x, y \in \mathbb{R}^n$, we put

$$\partial_x^2\varphi := \begin{pmatrix} \varphi_{x_1x_1} & \cdots & \varphi_{x_1x_n} \\ & \ddots & \\ \varphi_{x_nx_1} & \cdots & \varphi_{x_nx_n} \end{pmatrix}$$

and

$$\partial^2_{x,y}\varphi := \begin{pmatrix} \varphi_{x_1y_1} & \cdots & \varphi_{x_1y_n} \\ & \ddots & \\ \varphi_{x_ny_1} & \cdots & \varphi_{x_ny_n} \end{pmatrix}.$$

- Jacobians: Let

$$x \mapsto y = y(x)$$

be a diffeomorphism, $y = (y^1, \ldots, y^n)$. The Jacobian matrix is

$$\partial y = \partial_x y := \begin{pmatrix} \frac{\partial y^1}{\partial x_1} & \cdots & \frac{\partial y^1}{\partial x_n} \\ & \ddots & \\ \frac{\partial y^n}{\partial x_1} & \cdots & \frac{\partial y^n}{\partial x_n} \end{pmatrix}_{n\times n}.$$

1. DIFFERENTIAL FORMS

In this section we provide a minimalist review of differential forms on $\mathbb{R}^N$. For more a detailed and fully rigorous description of differential forms on manifolds we refer to [W, Chapter 2].

NOTATION.

(i) If $x = (x_1, \dots, x_n)$, $\xi = (\xi_1, \dots, \xi_n)$, then $dx_j, d\xi_j \in (\mathbb{R}^{2n})^*$ satisfy

$$\begin{aligned} dx_j(u) &= dx_j(x,\xi) = x_j \\ d\xi_j(u) &= d\xi_j(x,\xi) = \xi_j. \end{aligned}$$

(ii) If $\alpha, \beta \in (\mathbb{R}^{2n})^*$, then

$$(\alpha \wedge \beta)(u,v) := \alpha(u)\beta(v) - \alpha(v)\beta(u)$$

for $u, v \in \mathbb{R}^{2n}$. More generally, for $\alpha_j \in (\mathbb{R}^{2n})^*$, $j = 1, \cdots, m \leq 2n$, and $u = (u_1, \cdots, u_m)$, an m-tuple of $u_k \in \mathbb{R}^{2n}$,

$$(\alpha_1 \wedge \cdots \wedge \alpha_m)(u) = \det([\alpha_j(u_k)]_{1\leq j,k\leq 2n}). \tag{1.1}$$

(iii) If $f : \mathbb{R}^n \to \mathbb{R}$, the *differential* of f, is the 1-form

$$df = \sum_{j=1}^{n} \frac{\partial f}{\partial x_i} dx_i.$$

(iv) An m-form on $\mathbb{R}^n$ is given by

$$w = \sum_{i_1<i_2<\cdots<i_m} f_{i_1\cdots i_m}(x) dx_{i_1} \wedge \cdots dx_{i_m}, \quad f_{i_1\cdots i_m} \in C^\infty(\mathbb{R}^n).$$

Its action at x on m-tuples of vectors is given using (ii).

(v) The differential of m-form is defined by induction using (iii) and $d(fg) = df \wedge g + f dg$, where f is a function and g is an $(m-1)$-form. It satisfies $d^2 = 0$.

THEOREM 1.1 (**Alternative definition of $\boldsymbol{d}$).** *Suppose w is a differential 2-form, and $u \in C^\infty(\mathbb{R}^n, \mathbb{R}^3)$, $u = (u_1, u_3, u_3)$ is a 3-tuple of vectorfields. Then*

$$\begin{aligned} dw(u) = {} & u_1\left(w(u_2,u_3)\right) + u_2\left(w(u_3,u_1)\right) + u_3\left(w(u_1,u_2)\right) \\ & - w([u_1,u_2],u_3) - w([u_2,u_3],u_1) - w([u_3,u_1],u_2). \end{aligned} \tag{1.2}$$

1. Both sides of (1.2) are linear in w and trilinear in u.

2. When, say, u_1 is multiplied by $f \in C^\infty(\mathbb{R}^n)$, then

$$dw(fu_1, u_2, u_3) = fdw(u),$$

and the right hand side of (1.2) is equal to

$$\begin{aligned} &fu_1\left(w(u_2, u_3)\right) + u_2\left(fw(u_3, u_1)\right) + u_3\left(fw(u_1, u_2)\right) \\ &- w([fu_1, u_2], u_3) - w([u_2, u_3], fu_1) - w([u_3, fu_1], u_2), \end{aligned}$$

and this is equal to the right hand side of (1.2) multiplied by f. In fact,

$$[fu_1, u_2] = f[u_1, u_2] - (u_2 f)u_1, \quad [u_3, fu_1] = f[u_3, u_1] + (u_3 f)u_1,$$

and

$$\begin{aligned} u_2\left(fw(u_3, u_1)\right) &= fu_2\left(w(u_3, u_1)\right) + (u_2 f)w(u_3, u_1), \\ u_3\left(fw(u_1, u_2)\right) &= fu_3\left(w(u_1, u_2)\right) + (u_3 f)w(u_1, u_2). \end{aligned}$$

3. Hence we only need to check this identity for u constant and for $w = w_1 dw_2 \wedge dw_3$, where $w_1 \in C^\infty$, and w_2, w_3 are coordinate functions (that is are among $x_1, \cdots x_n$). Then

$$dw(u) = \det\left([u_j w_i]_{1 \leq i,j \leq n}\right),$$

and the right hand side of (1.2) is given by (remember that now $u_j w_i$, $i = 2, 3$ are constants) by the expansion of this determinant with respect to the first row, $(u_1 w_1, u_2 w_1, u_3 w_1)$. $\square$

DEFINITION. If η is a differential m-form and V a vector field, then the *contraction of* η *by* V, denoted

$$V \lrcorner \eta,$$

is the $(m-1)$-form defined by

$$(V \lrcorner \eta)(u) = \eta(V, u),$$

where u is an $(m-1)$-tuple of vectorfields. We use the consistent convention that for 0-forms, that is for functions, $V \lrcorner f = 0$.

We note the following property of contraction which can be deduced from (1.1): if v is a k-form and w is an m-form, then

$$V \lrcorner (v \wedge w) = (V \lrcorner v) \wedge w + (-1)^k v \wedge (V \lrcorner w). \tag{1.3}$$

DEFINITIONS. Let $\kappa : \mathbb{R}^n \to \mathbb{R}^n$ be a smooth mapping.

(i) If V is a vector field on $\mathbb{R}^n$, the *push-forward* is

$$\kappa_* V = \partial\kappa(V).$$

(ii) If η is a 1-form on $\mathbb{R}^n$, the *pull-back* is

$$(\kappa^*\eta)(u) = \eta(\kappa_* u).$$

THEOREM 1.2 (**Differentials and pull-backs).** *Let w be a differential m-form. We have*

$$d(\kappa^* w) = \kappa^*(dw). \tag{1.4}$$

Proof. 1. We first prove this for functions: $d(\kappa^* f) = d(\kappa(f)) = \sum_{j=1}^n \frac{\partial y_i}{\partial x_j}\frac{\partial f}{\partial y_i} dx_j$. Furthermore,

$$\kappa^*(df) = \kappa^*\left(\sum_{i=1}^n \frac{\partial f}{\partial y_i} dy_i\right) = \sum_{i=1}^n \frac{\partial f}{\partial y_i}\kappa^*(dy_i).$$

2. The proof now follows by induction on the order of the differential form: any m-form can be written as a linear combination of forms $f dg$ where f is a function, and g is $(m-1)$-form. □

DEFINITION. If V is a vector field generating the flow φ_t, then the *Lie derivative* of w is

$$\mathcal{L}_V w := \frac{d}{dt}((\varphi_t)^* w)|_{t=0}.$$

Here w denotes a function, a vector field or a form. We recall that φ_t is generated by a time independent vectorfield, V, $\varphi_t = \exp(tV)$, means that

$$(d/dt)\varphi_t(m) = V(\varphi_t(m)), \quad \varphi_0(m) = m.$$

EXAMPLE S. (i) If f is a function,

$$\mathcal{L}_V f = V(f).$$

(ii) If W is a vector field

$$\mathcal{L}_V W = [V, W].$$

Since for differential forms, w, $d(\varphi_t)^* w = \varphi_t^*(dw)$, we see that $\mathcal{L}_V$ commutes with d:

$$d(\mathcal{L}_V w) = \mathcal{L}_V(dw). \tag{1.5}$$

We also note that $\mathcal{L}_V$ is a derivation: for a function $f \in C^\infty$ and a differential form w,

$$\mathcal{L}_V(fw) = (\mathcal{L}_V f)w + f\mathcal{L}_V w. \tag{1.6}$$

THEOREM 1.3 (Cartan's formula). *If w is a differential form,*

$$\mathcal{L}_V w = d(V \lrcorner w) + (V \lrcorner dw). \tag{1.7}$$

Proof. 1. We proceed by induction on the order of differential forms. For 0-forms, that is for functions, we have

$$\mathcal{L}_V f = Vf = V \lrcorner df = d(V \lrcorner f) + (V \lrcorner df),$$

since by our convention $V \lrcorner f = 0$.

2. Any m-form is a linear combinations of forms $f dg$ where f is a function and g in an $(m-1)$-form. Then, using (1.5), (1.6), $d^2 = 0$, and the induction hypothesis,

$$\begin{aligned}\mathcal{L}_V(f dg) &= (\mathcal{L}_V f)dg + f\mathcal{L}_V dg \\ &= (Vf)dg + fd(\mathcal{L}_V g) \\ &= (Vf)dg + fd(d(V \lrcorner g) + V \lrcorner dg) \\ &= (Vf)dg + fd(V \lrcorner dg).\end{aligned} \tag{1.8}$$

3. The right hand side of (1.7) for $w = fdg$ is equal to

$$\begin{aligned}&d(V \lrcorner (f dg)) + V \lrcorner (d(f dg)) = \\ &\qquad f(V \lrcorner dg) + df \wedge (V \lrcorner dg) + V \lrcorner (df \wedge dg).\end{aligned} \tag{1.9}$$

Now we can use (1.3) with $w = df$, $v = dg$, $k = 1$, to obtain

$$V \lrcorner (df \wedge dg) = (Vf)dg - df \wedge (V \lrcorner dg).$$

Inserting this in (1.9) and the comparison with (1.8) gives (1.7) for $w = f dg$ and hence for all differential m-forms. □

THEOREM 1.4 (Poincaré's Lemma). *If α is a k-form defined in the open ball $U = B^0(0, R)$ and if*

$$d\alpha = 0,$$

then there exists a $(k-1)$ form ω in U such that

$$d\omega = \alpha.$$

Proof. 1. Let $\Omega^k(U)$ denote the space of k-forms on U. We will build a linear mapping

$$H : \Omega^k(U) \to \Omega^{k-1}(U)$$

such that

$$d \circ H + H \circ d = I. \tag{1.10}$$

Then

$$d(H\alpha) + Hd\alpha = \alpha$$

and so $d\omega = \alpha$ for $\omega := H\alpha$.

2. Define $A : \Omega^k(U) \to \Omega^k(U)$ by

$$A(f\,dx_{i_1} \wedge \cdots \wedge dx_{i_k}) = \left(\int_0^1 t^{k-1} f(tp)\, dt \right) dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

Set

$$X := \langle x, \partial_x \rangle = \sum_{j=1}^n x_j \frac{\partial}{\partial x_j}.$$

We claim

$$A\mathcal{L}_X = I \quad \text{on } \Omega^k(U). \tag{1.11}$$

and

$$d \circ A = A \circ d. \tag{1.12}$$

Assuming these assertions, define

$$H := A \circ X \lrcorner .$$

By Cartan's formula, Theorem 1.3,

$$\mathcal{L}_X = d \circ (X \lrcorner) + X \lrcorner \circ d.$$

Thus

$$\begin{aligned} I = A\mathcal{L}_X &= A \circ d \circ (X \lrcorner) + A \circ X \lrcorner \circ d \\ &= d(A \circ X \lrcorner) + (A \circ X \lrcorner) \circ d \\ &= d \circ H + H \circ d; \end{aligned}$$

and this proves (1.10).

3. To prove (1.11), we compute

$$
\begin{aligned}
&A\mathcal{L}_X(f dx_{i_1} \wedge \cdots \wedge dx_{i_2}) \\
&\qquad = A\left[\left(kf + \sum_{j=1}^{n} x_j \frac{\partial f}{\partial x_j}\right)(dx_{i_1} \wedge \cdots \wedge dx_{i_k})\right] \\
&\qquad = \int_0^1 kt^{k-1} f(tp) + \sum_{j=1}^{n} t^{k-1} x_j \frac{\partial f}{\partial x_j}(tp) \\
&\qquad\qquad\qquad dt dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\
&\qquad = \int_0^1 \frac{d}{dt}(t^k f(tp))\, dt dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\
&\qquad = f dx_{i_1} \wedge \cdots \wedge dx_{i_k}.
\end{aligned}
$$

4. To verify (1.12), note

$$
\begin{aligned}
&A \circ d(f dx_{i_1} \wedge \cdots \wedge dx_{i_k}) \\
&\qquad = A\left(\sum_{j=1}^{n} \frac{\partial f}{\partial x_j} dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}\right) \\
&\qquad = \left(\int_0^1 t^{k-1} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(tp) dx_j dt\right) dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\
&\qquad = d\left(\left(\int_0^1 t^{k-1} f(tp) dt\right) dx_{i_1} \wedge \cdots \wedge dx_{i_k}\right) \\
&\qquad = d \circ A(f dx_{i_1} \wedge \cdots \wedge dx_{i_k}).
\end{aligned}
$$

$\square$

2. FLOWS

Let $V : \mathbb{R}^N \to \mathbb{R}^N$ denote a smooth vector field. Fix a point $z \in \mathbb{R}^N$ and solve the ODE

$$
\begin{cases} \dot{z}(t) = V(z(t)) & (t \in \mathbb{R}) \\ z(0) = z. \end{cases} \tag{2.1}
$$

We assume that the solution of the flow (2.1) exists and is unique for all times $t \in \mathbb{R}$.

NOTATION. We define

$$\varphi_t z := z(t)$$

and sometimes also write

$$\varphi_t =: \exp(tV).$$

We call $\{\varphi_t\}_{t\in\mathbb{R}}$ the *exponential map.*

The following lemma records some standard assertions from theory of ordinary differential equations:

LEMMA 2.1 (Properties of flow map).
(i) $\varphi_0 z = z$.
(ii) $\varphi_{t+s} = \varphi_t \circ \varphi_s$ *for all* $s, t \in \mathbb{R}$.
(iii) *For each time* $t \in \mathbb{R}$*, the mapping* $\varphi_t : \mathbb{R}^N \to \mathbb{R}^N$ *is a diffeomorphism, with*

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

3. SYMPLECTIC STRUCTURE ON $\mathbb{R}^{2n}$

We henceforth specialize to the even-dimensional space $\mathbb{R}^N = \mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$.

NOTATION. We refine our previous notation and henceforth denote an element of $\mathbb{R}^{2n}$ as

$$z = (x, \xi),$$

and interpret $x \in \mathbb{R}^n$ as denoting *position*, $\xi \in \mathbb{R}^n$ as *momentum.* Alternatively, we can think of ξ as belonging to $T_x^*\mathbb{R}^n$, the cotangent space of $\mathbb{R}^n$ at x. We will likewise write

$$w = (y, \eta)$$

for another typical point of $\mathbb{R}^{2n}$.

We let $\langle\cdot,\cdot\rangle$ denote the usual inner product on $\mathbb{R}^n$, and then define this pairing on $\mathbb{R}^{2n}$:

DEFINITION. Given $z = (x, \xi)$, $w = (y, \eta)$ on $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$, define their *symplectic product*

$$\sigma(z, w) := \langle \xi, y\rangle - \langle x, \eta\rangle. \tag{3.1}$$

Note that

$$\sigma(z, w) = \langle Jz, w\rangle \tag{3.2}$$

for the $2n \times 2n$ matrix

$$J := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}. \tag{3.3}$$

Observe

$$J^2 = -I, \quad J^T = -J.$$

LEMMA 3.1 (Properties of $\boldsymbol{\sigma}$). *The bilinear form σ is antisymmetric:*

$$\sigma(z, w) = -\sigma(w, z)$$

and nondegenerate:

$$\text{if } \sigma(z, w) = 0 \text{ for all } w, \text{ then } z = 0.$$

These assertions are straightforward to check.

NOTATION. We now bring in the terminology of differential forms, reviewed in Section 1. Using the notation discussed above, we introduce for $x = (x_1, \dots, x_n)$ and $\xi = (\xi_1, \dots, \xi_n)$ the 1-forms dx_j and $d\xi_j$ for $j = 1, \dots, n$. We then can write

$$\sigma = d\xi \wedge dx = \sum_{j=1}^{n} d\xi_j \wedge dx_j. \tag{3.4}$$

Observe also

$$\sigma = d\omega \quad \text{for} \quad \omega := \xi dx = \sum_{j=1}^{n} \xi_j dx_j. \tag{3.5}$$

It follows that

$$d\sigma = 0. \tag{3.6}$$

4. CHANGING VARIABLES.

Suppose next that $U, V \subseteq \mathbb{R}^{2n}$ are open sets and

$$\kappa : U \to V$$

is a smooth mapping. We will write

$$\kappa(x, \xi) = (y, \eta) = (y(x, \xi), \eta(x, \xi)).$$

DEFINITION. We call κ a *symplectic* mapping, or a *symplectomorphism*, provided

$$\kappa^* \sigma = \sigma. \tag{4.1}$$

Here the *pull-back* $\kappa^* \sigma$ of the symplectic product σ is defined by

$$(\kappa^* \sigma)(z, w) := \sigma(\kappa_*(z), \kappa_*(w)),$$

κ_* denoting the *push-forward* of vectors: see Section 1.

NOTATION. We will usually write (4.1) in the more suggestive notation

$$d\eta \wedge dy = d\xi \wedge dx. \tag{4.2}$$

EXAMPLE 1: Linear symplectic mappings. Suppose $\kappa : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is linear:

$$\kappa(x,\xi) = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ \xi \end{pmatrix} = (Ax + B\xi, Cx + D\xi) = (y,\eta),$$

where A, B, C, D are $n \times n$ matrices.

THEOREM 4.1 (Symplectic matrices). *The linear mapping κ is symplectic if and only if the matrix*

$$M := \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

satisfies

$$M^T J M = J. \tag{4.3}$$

DEFINITION. We call a $2n \times 2n$ matrix M *symplectic* if (4.3) holds.

In particular the linear mapping $(x,\xi) \mapsto (\xi, -x)$ determined by J is symplectic.

Proof of Theorem 4.1. Let us compute

$$\begin{aligned} d\eta \wedge dy &= (Cdx + Dd\xi) \wedge (Adx + Bd\xi) \\ &= A^T C dx \wedge dx + B^T D d\xi \wedge d\xi + (A^T D - C^T B) d\xi \wedge dx \\ &= d\xi \wedge dx \end{aligned}$$

if and only if

$$A^T C \text{ and } B^T D \text{ are symmetric, } A^T D - C^T B = I. \tag{4.4}$$

Then

$$\begin{aligned} M^T J M &= \begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix} \begin{pmatrix} O & I \\ -I & O \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \\ &= \begin{pmatrix} A^T C - C^T A & A^T D - C^T B \\ B^T C - D^T A & B^T D - D^T B \end{pmatrix} \\ &= J \end{aligned}$$

if and only if (4.4) holds. $\square$

EXAMPLE 2: Nonlinear symplectic mappings. Assume next that $\kappa : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is nonlinear:

$$\kappa(x, \xi) = (y, \eta)$$

for smooth functions $y = y(x, \xi), \eta = \eta(x, \xi)$. Its linearization is the $2n \times 2n$ matrix

$$\partial\kappa = \partial_{x,\xi}\kappa = \begin{pmatrix} \partial_x y & \partial_\xi y \\ \partial_x \eta & \partial_\xi \eta \end{pmatrix}.$$

THEOREM 4.2 (Symplectic transformations). *The mapping κ is symplectic if and only if the matrix $\partial\kappa$ is symplectic at each point.*

Proof. We have

$$d\eta \wedge dy = (C dx + D d\xi) \wedge (A dx + B d\xi)$$

for

$$A := \partial_x y, B := \partial_\xi y, C := \partial_x \eta, D := \partial_\xi \eta.$$

Consequently, as in the previous proof, we have $d\eta \wedge dy = d\xi \wedge dx$ if and only if (4.4) is valid, which in turn is so if and only if $\partial\kappa$ is a symplectic matrix. $\square$

EXAMPLE 3: Lifting diffeomorphisms. Let

$$\gamma : \mathbb{R}^n \to \mathbb{R}^n$$

be a diffeomorphism on $\mathbb{R}^n$, with nondegenerate Jacobian matrix $\partial_x \gamma$. We propose to extend γ to a symplectomorphism

$$\kappa : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$$

having the form

$$\kappa(x, \xi) = (\gamma(x), \eta(x, \xi)) = (y, \eta), \tag{4.5}$$

by lifting γ to the momentum variables.

THEOREM 4.3 (Extending to a symplectic mapping). *The transformation* (4.5) *is symplectic for*

$$\eta(x, \xi) := \left[\partial_x \gamma(x)^{-1}\right]^T \xi. \tag{4.6}$$

Proof. As the statement of the theorem suggests it is easier to look for ξ as a function of x and η. We compute

$$dy = A\ dx, \quad d\xi = E\ dx + F\ d\eta,$$

for

$$A := \partial_x y, \quad E := \partial_x \xi, \quad F := \partial_\eta \xi.$$

Therefore

$$d\eta \wedge dy = d\eta \wedge (A\ dx)$$

and

$$d\xi \wedge dx = (Edx \wedge Fd\eta) \wedge dx = Edx \wedge dx + d\eta \wedge F^T dx.$$

We would like to construct $\xi = \xi(x, \eta)$ so that

$$A = F^T \quad \text{and} \quad E \text{ is symmetric},$$

the latter condition implying that $Edx \wedge dx = 0$. To do so, let us define

$$\xi(x, \eta) := (\partial_x \gamma)^T \eta.$$

Then clearly $F^T = A$, and $E = E^T = ((\gamma_{x_i x_j}))$, as required. □

EXAMPLE 4: Generating functions. Our last example demonstrates that we can, locally at least, build a symplectic transformation from a real-valued *generating function.*

Suppose $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\varphi = \varphi(x, y)$, is smooth. Assume also that

$$\det(\partial^2_{xy} \varphi(x_0, y_0)) \neq 0. \tag{4.7}$$

Define

$$\xi = \partial_x \varphi, \ \ \eta = -\partial_y \varphi, \tag{4.8}$$

and observe that the Implicit Function Theorem implies (y, η) is a smooth function of (x, ξ) near $(x_0, \partial_x \varphi(x_0, y_0))$.

THEOREM 4.4 (Generating functions and symplectic maps). *The mapping γ defined by*

$$(x, \partial_x \varphi(x, y)) \mapsto (y, -\partial_y \varphi(x, y)) \tag{4.9}$$

is a symplectomorphism near (x_0, ξ_0).

Proof. We compute

$$\begin{aligned} d\eta \wedge dy &= d(-\partial_y \varphi) \wedge dy \\ &= [(-\partial^2_y \varphi dy) \wedge dy] + [(-\partial^2_{xy} \varphi dx) \wedge dy] \\ &= -(\partial^2_{xy} \varphi) dx \wedge dy, \end{aligned}$$

since $\partial_y^2\varphi$ is symmetric. Likewise,

$$\begin{aligned} d\xi \wedge dx &= d(\partial_x\varphi)\wedge dx \\ &= [(\partial_x^2\varphi\ dx)\wedge dx] + [(\partial_{xy}^2\varphi\ dy)\wedge dx] \\ &= -(\partial_{xy}^2\varphi)dx\wedge dy = d\eta\wedge dy. \end{aligned}$$

□

TERMINOLOGY. In Greek, the word "symplectic" means "intertwined", This is consistent with Example 4, since the generating function $\varphi = \varphi(x, y)$ is a function of a mixture of half of the original variables (x, ξ) and half of the new variables (y, η). "Symplectic" can also be interpreted as "complex", mathematical usage due to Hermann Weyl who renamed "line complex group" the "symplectic group": see Cannas da Silva [CdS].

APPLICATION: Lagrangian submanifolds. A *Lagrangian submanifold* Λ is an n-dimensional submanifold of $\mathbb{R}^{2n}$ for which

$$\sigma|_\Lambda = 0.$$

Then

$$d\omega|_\Lambda = \sigma|_\Lambda = 0;$$

and so according to Poincaré's Theorem 1.4, we locally have

$$\omega|_\Lambda = d\varphi,$$

for some smooth function φ on Λ. □

5. HAMILTONIAN VECTOR FIELDS

DEFINITION. Given $f \in C^\infty(\mathbb{R}^{2n})$, we define the corresponding *Hamiltonian vector field* by requiring

$$\sigma(z, H_f) = df(z) \qquad \text{for all } z = (x, \xi). \tag{5.1}$$

This is well defined, since σ is nondegenerate. We can write explicitly that

$$H_f = \langle \partial_\xi f, \partial_x\rangle - \langle \partial_x f, \partial_\xi\rangle = \sum_{j=1}^n f_{\xi_j}\partial_{x_j} - \sum_{j=1}^n f_{x_j}\partial_{\xi_j}. \tag{5.2}$$

Another way to write the definition of H_f is by using the contraction $\lrcorner$ defined in Section 1:

$$df = -(H_f \lrcorner \sigma)\,, \tag{5.3}$$

which follows directly from the definition: we calculate for each z that

$$(H_f \lrcorner \sigma)(z) = \sigma(H_f, z) = -\sigma(z, H_f) = -df(z).$$

$\square$

DEFINITION. If $f, g \in C^\infty(\mathbb{R}^{2n})$, we define their *Poisson bracket*

$$\{f, g\} := H_f g = \sigma(\partial f, \partial g). \tag{5.4}$$

That is,

$$\{f, g\} = \langle \partial_\xi f, \partial_x g\rangle - \langle \partial_x f, \partial_\xi g\rangle = \sum_{j=1}^{n} \left(f_{\xi_j} g_{x_j} - f_{x_j} g_{\xi_j}\right) . \tag{5.5}$$

LEMMA 5.1 (Brackets, commutators).

(i) *We have* Jacobi's identity

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0 \tag{5.6}$$

for all functions $f, g, h \in C^\infty(\mathbb{R}^{2n})$.

(ii) *Furthermore,*

$$H_{\{f,g\}} = [H_f, H_g]. \tag{5.7}$$

Proof. 1. A direct calculation verifies assertion (i). For an alternative proof, showing that the essential property is $d\sigma = 0$, note that Lemma 1.1 provides the identity

$$\begin{aligned} 0 &= d\sigma(H_f, H_g, H_h) \\ &= H_f \sigma(H_g, H_h) + H_g \sigma(H_h, H_f) + H_h \sigma(H_f, H_g) \\ &\quad - \sigma([H_f, H_g], H_h) - \sigma([H_g, H_h], H_f) - \sigma([H_h, H_f], H_g) . \end{aligned} \tag{5.8}$$

Now (5.4) implies

$$H_f \sigma(H_g, H_h) = \{f, \{g, h\}\}$$

and

$$\begin{aligned} \sigma([H_f, H_g], H_h) &= [H_f, H_g]h = H_f H_g h - H_g H_f h \\ &= \{f, \{g, h\}\} - \{g, \{f, h\}\}. \end{aligned}$$

Similar identities hold for other terms. Substituting into (5.8) gives Jacobi's identity.

2. We observe that

$$H_{\{f,g\}} h = [H_f, H_g] h$$

is a rewriting of (5.6). □

THEOREM 5.2 (Jacobi's Theorem). *If κ is a symplectomorphism, then*

$$H_f = \kappa_*(H_{\kappa^* f}). \tag{5.9}$$

In other words, the pull-back of a Hamiltonian vector field generated by f,

$$\kappa^* H_f := (\kappa^{-1})_* H_f, \tag{5.10}$$

is the Hamiltonian vector field generated by the pull-back of f.

Proof. Using the notation of (5.10),

$$\begin{aligned} \kappa^*(H_f) \lrcorner \sigma &= \kappa^*(H_f) \lrcorner \kappa^* \sigma = \kappa^*(H_f \lrcorner \sigma) \\ &= -\kappa^*(df) = -d(\kappa^* f) \\ &= H_{\kappa^* f} \lrcorner \sigma. \end{aligned}$$

Since σ is nondegenerate, (5.9) follows. □

EXAMPLE. Define $\kappa = J$, so that $\kappa(x, \xi) = (\xi, -x)$; and recall κ is a symplectomorphism. We have $\kappa^* f(x, \xi) = f(\xi, -x)$, and therefore

$$H_{\kappa^* f} = \langle \partial_x f(\xi, -x), \partial_x \rangle + \langle \partial_\xi f(\xi, -x), \partial_\xi \rangle.$$

Then

$$\kappa^* H_f = \langle \partial_\xi f(\xi, -x), \partial_\xi \rangle - \langle \partial_x f(\xi, -x), \partial_{-x} \rangle = H_{\kappa^* f}.$$

□

THEOREM 5.3 (Hamiltonian flows as symplectomorphisms). *If f is smooth, then for each time t, the mapping*

$$(x, \xi) \mapsto \varphi_t(x, \xi) = \exp(tH_f)$$

is a symplectomorphism.

Proof. According to Cartan's formula (Theorem 1.3), we have

$$\frac{d}{dt}((\varphi_t)^* \sigma) = \mathcal{L}_{H_f} \sigma = d(H_f \lrcorner \sigma) + (H_f \lrcorner d\sigma).$$

Since $d\sigma = 0$, it follows that

$$\frac{d}{dt}((\varphi_t)^* \sigma) = d(-df) = -d^2 f = 0.$$

Thus $(\varphi_t)^* \sigma = \sigma$ for all times t. □

The next result shows that locally all nondegenerate closed two forms are equivalent to the standard symplectic form on $\mathbb{R}^{2n}$, σ.

THEOREM 5.4 (Darboux's Theorem). *Let U be a neighborhood of (x_0, ξ_0) and suppose η is a nondegenerate 2-form defined on U, satisfying*

$$d\eta = 0.$$

Then near (x_0, ξ_0) there exists a diffeomorphism κ such that

$$\kappa^* \eta = \sigma. \tag{5.11}$$

A symplectic structure is the existence of a form η. Darboux's theorem states that all symplectic structures are identical locally, in the sense that all are equivalent to that given by σ. This is dramatic contrast to Riemannian geometry: there are no local invariants in symplectic geometry.

Proof. 1. Let us assume $(x_0, \xi_0) = (0, 0)$. We first find a linear mapping L so that

$$L^* \eta(0, 0) = \sigma(0, 0).$$

This means that we find a basis $\{e_k, f_k\}_{k=1}^n$ of $\mathbb{R}^{2n}$ such that

$$\begin{cases} \eta(f_l, e_k) = \delta_{kl} \\ \eta(e_k, e_l) = 0 \\ \eta(f_k, f_l) = 0 \end{cases}$$

for all $1 \leq k, l \leq n$. Then if $u = \sum_{i=1}^n x_i e_i + \xi_i f_i$, $v = \sum_{j=1}^n y_j e_j + \eta_j f_j$, we have

$$\begin{aligned} \eta(u, v) &= \sum_{i,j=1}^n x_i y_j \eta(e_i, e_j) + \xi_i \eta_j \eta(f_i, f_j) + x_i \eta_j \sigma(e_i, f_j) + \xi_i y_j \sigma(f_i, e_j) \\ &= \langle \xi, y \rangle - \langle x, \eta \rangle = \sigma((x, \xi), (y, \eta)). \end{aligned}$$

We leave finding L as a linear algebra exercise.

2. Next, define $\eta_t := t\eta + (1 - t)\sigma$ for $0 \leq t \leq 1$. Our intention is to find κ_t so that $\kappa_t^* \eta_t = \sigma$ near $(0, 0)$; then $\kappa := \kappa_1$ solves our problem. We will construct κ_t by solving the flow

$$\begin{cases} \dot{z}(t) = V_t(z(t)) \qquad (0 \leq t \leq 1) \\ z(0) = z, \end{cases} \tag{5.12}$$

and setting $\kappa_t := \varphi_t$.

For this to work, we must design the vector fields V_t in (5.12) so that

$$\frac{d}{dt}(\kappa_t^* \eta_t) = 0\,.$$

Let us therefore calculate

$$\begin{aligned}\frac{d}{dt}(\kappa_t^* \eta_t) &= \kappa_t^* \left(\frac{d}{dt}\eta_t\right) + \kappa_t^* \mathcal{L}_{V_t} \eta_t \\ &= \kappa_t^* \left[(\eta - \sigma) + d(V_t \lrcorner \eta_t) + V_t \lrcorner d\eta_t\right],\end{aligned}$$

where we used Cartan's formula, Theorem 1.3. Note that $d\eta_t = td\eta + (1-t)d\sigma$. Hence $(d/dt)(\kappa_t^* \eta_t) = 0$ provided

$$(\eta - \sigma) + d(V_t \lrcorner \eta_t) = 0. \tag{5.13}$$

According to Poincaré's Theorem 1.4, we can write

$$\eta - \sigma = d\alpha \quad \text{near } (0,0).$$

So (5.13) will hold, provided

$$V_t \lrcorner \eta_t = -\alpha \qquad (0 \le t \le 1). \tag{5.14}$$

Since $\eta = \sigma$ at $(0,0)$, $\eta_t = \sigma$ at $(0,0)$. In particular, η_t is nondegenerate for $0 \le t \le 1$ in a neighbourhood of $(0,0)$, and hence we can solve (5.13) for the vector field V_t. □

References

[CdS] A. Cannas da Silva, *Lectures on Symplectic Geometry,* Lecture Notes in Mathematics **1764**, 2001.

[W] F.W. Warner, *Foundations of differentiable manifolds and Lie groups*, GMT 94, Springer Verlag, 1983.