

Weak symplectic forms and differential calculus in Banach spaces

Jordan Bell

May 16, 2015

1 Introduction

There are scarcely any decent expositions of infinite dimensional symplectic vector spaces. One good basic exposition is by Marsden and Ratiu.¹ The Darboux theorem for a real reflexive Banach space is proved in Lang and probably in fewer other places than one might guess.² (Other references.³)

2 Bilinear forms

Let E be a real Banach space. For a bilinear form $B : E \times E \rightarrow \mathbb{R}$, define

$$\|B\| = \sup_{\|e\| \leq 1, \|f\| \leq 1} |B(e, f)|.$$

One proves that B is continuous if and only if $\|B\| < \infty$. Namely, a bilinear form is continuous if and only if it is bounded.

If $B : E \times E \rightarrow \mathbb{R}$ is a continuous bilinear form, we define $B^\flat : E \rightarrow E^*$ by

$$B^\flat(e)(f) = B(e, f), \quad e \in E, f \in E;$$

indeed, for $e \in E$, $\|B^\flat(e)f\| = \|B(e, f)\| \leq \|B\| \|e\| \|f\|$, showing that $\|B^\flat(e)\| \leq \|B\| \|e\|$, so that $B^\flat(e)$ is continuous $E \rightarrow \mathbb{R}$. Moreover, it is apparent that B^\flat

¹Jerrold E. Marsden and Tudor S. Ratiu, *Introduction to Mechanics and Symmetry*, second ed., Chapter 2.

²Serge Lang, *Differential and Riemannian Manifolds*, p. 150, Theorem 8.1; Mircea Puta, *Hamiltonian Mechanical Systems and Geometric Quantization*, p. 12, Theorem 1.3.1.

³Andreas Kriegl and Peter W. Michor, *The Convenient Setting of Global Analysis*, p. 522, §48; Peter W. Michor, *Some geometric evolution equations arising as geodesic equations on groups of diffeomorphisms including the Hamiltonian approach*, pp. 133–215, in Antonio Bove, Ferruccio Colombini, and Daniele Del Santo (eds.), *Phase Space Analysis of Partial Differential Equations*; K.-H. Need, H. Sahlmann, and T. Thiemann, *Weak Poisson Structures on Infinite Dimensional Manifolds and Hamiltonian Actions*, pp. 105–135, in Vladimir Dobrev (ed.), *Lie Theory and Its Applications in Physics*; Tudor S. Ratiu, *Coadjoint Orbits and the Beginnings of a Geometric Representation Theory*, pp. 417–457, in Karl-Hermann Neeb and Arturo Pianzola (eds.), *Developments and Trends in Infinite-Dimensional Lie Theory*.

is linear, and

$$\begin{aligned}\|B^b\| &= \sup_{\|e\| \leq 1} \|B^b(e)\| \\ &= \sup_{\|e\| \leq 1} \sup_{\|f\| \leq 1} |B^b(e)(f)| \\ &= M,\end{aligned}$$

so $B^b : E \rightarrow E^*$ is continuous.

We call a continuous bilinear form $B : E \times E \rightarrow F$ **weakly nondegenerate** if $B^b : E \rightarrow E^*$ is one-to-one. Since B^b is linear, this is equivalent to the statement that $B^b(e) = 0$ implies that $e = 0$, which is equivalent to the statement that if $B(e, f) = 0$ for all f then $e = 0$.

An **isomorphism of Banach spaces** is a linear isomorphism $T : E \rightarrow F$ that is continuous such that $T^{-1}F \rightarrow E$ is continuous. Equivalently, to say that $T : E \rightarrow F$ is an isomorphism of Banach spaces means that $T : E \rightarrow F$ is a bijective bounded linear map such that $T^{-1} : F \rightarrow E$ is a bounded linear map. It follows from the open mapping theorem that if $T : E \rightarrow F$ is an onto bounded linear isomorphism, hence is an isomorphism of Banach spaces.

We say that a continuous bilinear form $B : E \times E \rightarrow \mathbb{R}$ is **strongly nondegenerate** if $B^b : E \rightarrow E^*$ is an isomorphism of Banach spaces.

For a real vector space V and a bilinear form $B : V \times V \rightarrow \mathbb{R}$, we say that B is **alternating** if $B(v, v) = 0$ for all $v \in V$. We say that B is **skew-symmetric** if $B(u, v) = -B(v, u)$ for all $u, v \in V$. It is straightforward to check that B is alternating if and only if B is skew-symmetric.

For Banach spaces E_1, \dots, E_p and F , let $\mathcal{L}(E_1, \dots, E_p; F)$ denote the set of continuous multilinear maps $E_1 \times \dots \times E_p \rightarrow F$. For a multilinear map $T : E_1 \times \dots \times E_p \rightarrow F$ to be continuous it is equivalent that

$$\|T\| = \sup_{\|e_1\| \leq 1, \dots, \|e_p\| \leq 1} \|T(e_1, \dots, e_p)\| < \infty,$$

namely that it is bounded with the operator norm. With this norm, $\mathcal{L}(E_1, \dots, E_p; F)$ is a Banach space.⁴ We write

$$\mathcal{L}_p(E; F) = \mathcal{L}(E_1, \dots, E_p; F).$$

For Banach spaces E and F , we denote by $\text{GL}(E; F)$ the set of isomorphisms $E \rightarrow F$. One proves that $\text{GL}(E; F)$ is an open set in the Banach space $\mathcal{L}(E; F)$ and that with the subspace topology, $u \mapsto u^{-1}$ is continuous $\text{GL}(E; F) \rightarrow \text{GL}(F; E)$.⁵

For Banach spaces E, F, G , define

$$\phi : \mathcal{L}(E, F; G) \rightarrow \mathcal{L}(E; \mathcal{L}(F, G))$$

by $\phi(f)(x)(y) = f(x, y)$ for $f \in \mathcal{L}(E, F; G)$, $x \in E$, and $y \in F$. One proves that ϕ is an isometric isomorphism.⁶

⁴Henri Cartan, *Differential Calculus*, p. 22, Theorem 1.8.1.

⁵Henri Cartan, *Differential Calculus*, p. 20, Theorem 1.7.3.

⁶Henri Cartan, *Differential Calculus*, p. 23, §1.9.

3 Differentiable functions

Let E and F be Banach spaces and let U be a nonempty open subset of E . For $a \in U$, a function $f : U \rightarrow F$ is said to be **differentiable at a** if (i) f is continuous at a and (ii) there is a linear mapping $g : E \rightarrow F$ such that

$$\|f(x) - f(a) - (g(x) - g(a))\|_F = o(\|x - a\|_E),$$

as $x \rightarrow a$ in E . We prove that there is at most one such linear mapping g and write $f'(a) = g$, and call $f'(a)$ the **derivative of f at a** . We also prove that if f is differentiable at a then $f'(a) : E \rightarrow F$ is continuous at a and therefore, being linear, is continuous on E , namely $f'(a) \in \mathcal{L}(E; F)$.⁷

If $f : U \rightarrow F$ is differentiable at each $a \in U$, we say that f is **differentiable on U** . We call $f' : U \rightarrow \mathcal{L}(E; F)$ the **derivative of f** . We also write $Df = f'$.

We say that $f : U \rightarrow F$ is C^1 , also called **continuously differentiable**, if (i) f is differentiable on U and (ii) $f' : U \rightarrow \mathcal{L}(E; F)$ is continuous.

Let E, F, G be Banach spaces, let U be an open subset of E , let V be an open subset of F , and let $f : U \rightarrow F$ and $g : V \rightarrow G$ be continuous. Suppose that $a \in U$ and that $f(a) \in V$. We define $g \circ f : f^{-1}(V) \rightarrow G$ on $f^{-1}(V)$. One proves that if f is differentiable at a and g is differentiable at $f(a)$, then $h = g \circ f : f^{-1}(V) \rightarrow G$ is differentiable at a and satisfies⁸

$$h'(a) = g'(f(a)) \circ f'(a).$$

For Banach spaces E and F , let $\phi : \text{GL}(E; F) \rightarrow \mathcal{L}(F; E)$ be defined by $\phi(u) = u^{-1}$. $\text{GL}(E; F)$ is an open subset of the Banach space $\mathcal{L}(E; F)$ and ϕ is continuous. It is proved that ϕ is continuously differentiable, and that for $u \in \text{GL}(E; F)$, the derivative of ϕ at u ,

$$\phi'(u) \in \mathcal{L}(\mathcal{L}(E; F); \mathcal{L}(F; E)),$$

satisfies⁹

$$\phi'(u)(h) = -u^{-1} \circ h \circ u^{-1}, \quad h \in \mathcal{L}(E; F).$$

4 Symplectic forms

A **weak symplectic form** on a Banach space E is a continuous bilinear form $\Omega : E \times E \rightarrow \mathbb{R}$ that is weakly nondegenerate and alternating.

A **strong symplectic form** on a Banach space E is a continuous bilinear form $\Omega : E \times E \rightarrow \mathbb{R}$ that is strongly nondegenerate and alternating. If Ω is a strong symplectic form on a Banach space E , we define $\Omega^\# : E^* \rightarrow E$ by $\Omega^\# = (\Omega^\flat)^{-1}$, which is an isomorphism of Banach spaces.

⁷Henri Cartan, *Differential Calculus*, p. 25.

⁸Henri Cartan, *Differential Calculus*, p. 27, Theorem 2.2.1.

⁹Henri Cartan, *Differential Calculus*, p. 31, Theorem 2.4.4.

5 Hamiltonian functions

Let E be a real Banach space E , let $\mathcal{D}(A)$ be a linear subspace of E , and let $A : \mathcal{D}(A) \rightarrow E$ be a linear map, called an **operator in E** . Write $\mathcal{R}(A) = A\mathcal{D}(A)$. For a weak symplectic form ω on E , we say that A is **ω -skew** if

$$\omega(Ae, f) = -\omega(e, Af), \quad e, f \in \mathcal{D}(A).$$

If $\mathcal{R}(A) \subset \mathcal{D}(A)$ and $A^2 = -I$, then for $e, f \in \mathcal{D}(A)$ we have $\omega(Ae, Af) = -\omega(e, A^2f) = -\omega(e, -f) = \omega(e, f)$.

For an ω -skew operator A in E , we define $H : \mathcal{D}(A) \rightarrow \mathbb{R}$, called the **Hamiltonian function of A** ,¹⁰ by

$$H(u) = \frac{1}{2}\omega(Au, u), \quad u \in \mathcal{D}(A).$$

For a linear operator A in E , we define

$$\mathcal{G}(A) = \{(u, Au) : u \in \mathcal{D}(A)\}.$$

$\mathcal{G}(A)$ is a linear subspace of $E \times E$. We say that A is **closed** if $\mathcal{G}(A)$ is a closed subset of $E \times E$. One proves that a linear operator A in E is closed if and only if the linear space $\mathcal{D}(A)$ with the norm

$$\|e\|_A = \|e\| + \|Ae\|, \quad e \in \mathcal{D}(A)$$

is a Banach space.

For $T \in \mathcal{L}(E)$, we define $T^*\omega : E \times E \rightarrow \mathbb{R}$ by

$$(T^*\omega)(e, f) = \omega(Te, Tf), \quad (e, f) \in E \times E;$$

$T^*\omega$ is called the **pullback of ω by T** . It is apparent that $T^*\omega$ is bilinear. We have

$$\begin{aligned} \|T^*\omega\| &= \sup_{\|e\| \leq 1, \|f\| \leq 1} |\omega(Te, Tf)| \\ &\leq \sup_{\|e\| \leq 1, \|f\| \leq 1} \|\omega\| \|Te\| \|Tf\| \\ &\leq \sup_{\|e\| \leq 1, \|f\| \leq 1} \|\omega\| \|T\| \|e\| \|T\| \|f\| \\ &= \|\omega\| \|T\|^2, \end{aligned}$$

showing that $T^*\omega$ is continuous. For $e \in E$, because ω is alternating we have

$$(T^*\omega)(e, e) = \omega(Te, Te) = 0,$$

i.e. $T^*\omega$ is alternating. For $e \in E$, suppose that $(T^*\omega)(e, f) = 0$ for all $f \in E$. That is, $\omega(Te, Tf) = 0$ for all $f \in E$, and thus, to establish that $T^*\omega$ is weakly

¹⁰See Jerrold E. Marsden and Thomas J. R. Hughes, *Mathematical Foundations of Elasticity*, p. 253, §5.1.

nondegenerate it suffices that T be onto. In the case that $T^*\omega = \omega$, we say that $T \in \mathcal{L}(E)$ is a **canonical transformation**.

Suppose that A is a closed ω -skew operator in E , with Hamiltonian function $H : \mathcal{D}(A) \rightarrow \mathbb{R}$. $\mathcal{D}(A)$ is a Banach space with the norm $\|e\|_A = \|e\| + \|Ae\|$. For $u \in \mathcal{D}(A)$ and $v \in \mathcal{D}(A)$, using the fact that A is ω -skew we check that

$$H(v) - H(u) - \omega(Au, v - u) = \frac{1}{2}\omega(A(v - u), v - u),$$

hence

$$|H(v) - H(u) - \omega(Au, v - u)| \leq \frac{1}{2} \|\omega\| \|A(v - u)\| \|v - u\| \leq \frac{1}{2} \|\omega\| \|v - u\|_A^2.$$

This shows that H is differentiable on the Banach space $\mathcal{D}(A)$, with derivative $H' : \mathcal{D}(A) \rightarrow \mathcal{D}(A)^*$ defined by¹¹

$$H'(u)(e) = \omega(Au, e), \quad u \in \mathcal{D}(A), \quad e \in \mathcal{D}(A).$$

Moreover, for $u, v \in \mathcal{D}(A)$ we have

$$\begin{aligned} \|H'(v) - H'(u)\| &= \sup_{\|e\|_A \leq 1} |H'(v)(e) - H'(u)(e)| \\ &= \sup_{\|e\|_A \leq 1} |\omega(Av, e) - \omega(Au, e)| \\ &= \sup_{\|e\|_A \leq 1} |\omega(A(v - u), e)| \\ &\leq \sup_{\|e\|_A \leq 1} \|\omega\| \|A(v - u)\| \|e\| \\ &\leq \|\omega\| \|A(v - u)\| \\ &\leq \|\omega\| \|v - u\|_A, \end{aligned}$$

showing that $H' : \mathcal{D}(A) \rightarrow \mathcal{D}(A)^*$ is continuous, namely that H is C^1 . (We also write $DH = H'$.)

Suppose that A is a closed operator in E and that $H : \mathcal{D}(A) \rightarrow \mathbb{R}$ is some function such that $H'(u)e = \omega(Au, e)$ for all $u \in \mathcal{D}(A)$ and $e \in \mathcal{D}(A)$. On the one hand, because H' is continuous and linear, the second derivative $D^2H : \mathcal{D}(A) \rightarrow \mathcal{L}(\mathcal{D}(A), \mathcal{D}(A)^*)$ is

$$(D^2H)(u)(e)(f) = H'(e)(f) = \omega(Ae, f), \quad u, e, f \in \mathcal{D}(A).$$

On the other hand, because D^2H is continuous, for each $u \in \mathcal{D}(A)$, the bilinear form $(D^2H)(u) : \mathcal{D}(A) \times \mathcal{D}(A) \rightarrow \mathbb{R}$ is symmetric.¹² That is, $(D^2H)(u)(e)(f) = (D^2H)(u)(f)(e)$, which by the above means

$$\omega(Ae, f) = \omega(Af, e), \quad e, f \in \mathcal{D}(A),$$

¹¹cf. Jerrold E. Marsden and Thomas J. R. Hughes, *Mathematical Foundations of Elasticity*, p. 254, Proposition 2.2.

¹²Serge Lang, *Real and Functional Analysis*, third ed., p. 344, Theorem 5.3.

showing that A is ω -skew. Let $G : \mathcal{D}(A) \rightarrow \mathbb{R}$ be the Hamiltonian function of A , i.e.

$$G(u) = \frac{1}{2}\omega(Au, u), \quad u \in \mathcal{D}(A).$$

What we established earlier tells us that

$$G'(u)(e) = \omega(Au, e), \quad u \in \mathcal{D}(A), \quad e \in \mathcal{D}(A).$$

Then we have that for $G' = H'$. Let $K = G - H$, which is C^1 with $K' = 0$. The **mean value theorem**¹³ tells us that for any $x, y \in \mathcal{D}(A)$,

$$K(x + y) - K(x) = \int_0^1 K'(x + ty)(y)dt = 0,$$

and thus $K(u) = K(0) = C$ for all $u \in \mathcal{D}(A)$. Therefore, $G = H + C$.

6 Semigroups

Let E be a real Banach space, let ω be a weak symplectic form on E , and let A be a closed densely defined ω -skew linear operator in E . Suppose that A is the infinitesimal generator of a strongly continuous one-parameter semigroup $\{U_t : t \geq 0\}$, where $U_t \in \mathcal{L}(E)$ for each t , and let H be the Hamiltonian function of A .¹⁴

Theorem 1. For each $t \geq 0$, U_t is a canonical transformation.

For each $t \geq 0$ and for each $x \in \mathcal{D}(A)$,

$$H(U_t x) = H(x).$$

Proof. For $u, v \in \mathcal{D}(A)$ and $t \geq 0$, using the chain rule and the fact that ω is a bilinear form,¹⁵

$$\frac{d}{dt}\omega(U_t u, U_t v) = \omega\left(\frac{d}{dt}U_t u, U_t v\right) + \omega\left(U_t u, \frac{d}{dt}U_t v\right).$$

Because A is the infinitesimal generator of $\{U_t : t \geq 0\}$, it follows that $\frac{d}{dt}(U_t w) = AU_t w$ for each $w \in \mathcal{D}(A)$. Using this and the fact that A is ω -skew,

$$\begin{aligned} \frac{d}{dt}\omega(U_t u, U_t v) &= \omega(AU_t u, U_t v) + \omega(U_t u, AU_t v) \\ &= -\omega(U_t u, AU_t v) + \omega(U_t u, AU_t v) \\ &= 0. \end{aligned}$$

¹³Serge Lang, *Real and Functional Analysis*, third ed., p. 341, Theorem 4.2.

¹⁴Jerrold E. Marsden and Thomas J. R. Hughes, *Mathematical Foundations of Elasticity*, p. 256, Proposition 2.6.

¹⁵Henri Cartan, *Differential Calculus*, p. 30, Theorem 2.4.3.

This implies that $\omega(U_t u, U_t v) = \omega(U_0 u, U_0 v) = \omega(u, v)$ for all $t \geq 0$, which means that U_t is a canonical transformation for each $t \geq 0$.

For any $t \geq 0$ and $x \in \mathcal{D}(A)$, $AU_t x = U_t A x$. (The infinitesimal generator of a one-parameter semigroup commutes with each element of the semigroup.) Then, using the fact that U_t is a canonical transformation,

$$\begin{aligned} H(U_t x) &= \frac{1}{2} \omega(A(U_t x), U_t x) \\ &= \frac{1}{2} \omega(U_t A x, U_t x) \\ &= \frac{1}{2} \omega(A x, x) \\ &= H(x). \end{aligned}$$

□

Suppose that there is some $c > 0$ such that $H(u) \geq c \|u\|_A^2$ for all $u \in \mathcal{D}(A)$, namely that H is **coercive** on the Banach space $\mathcal{D}(A)$. Let $t \geq 0$ and let $u \in \mathcal{D}(A)$. Then $U_t u \in \mathcal{D}(A)$, so using the hypothesis and Theorem 1,

$$\|U_t u\|_A^2 \leq \frac{1}{c} H(U_t u) = \frac{1}{c} H(u) = \frac{1}{2c} \omega(Au, u) \leq \frac{1}{2c} \|\omega\| \|Au\| \|u\| \leq \frac{\|\omega\|}{2c} \|u\|_A^2.$$

Therefore, for each $t \geq 0$ and $u \in \mathcal{D}(A)$,

$$\|U_t u\|_A \leq \sqrt{\frac{\|\omega\|}{2c}} \|u\|_A.$$

7 Hilbert spaces

For a real vector space V , a **complex structure on V** is a linear map $J : V \rightarrow V$ such that $J^2 = -I$. For $v \in V$, define $iv = Jv \in V$, for which on the one hand,

$$\begin{aligned} (\alpha + i\beta)(\gamma + i\delta)v &= (\alpha + i\beta)(\gamma v + \delta Jv) \\ &= \alpha\gamma v + \alpha\delta Jv + J(\beta\gamma v) + J(\beta\delta Jv) \\ &= \alpha\gamma v + (\alpha\delta + \beta\gamma)Jv + \beta\delta J^2 v \\ &= (\alpha\gamma - \beta\delta)v + (\alpha\delta + \beta\gamma)Jv, \end{aligned}$$

and on the other hand,

$$(\alpha + i\beta)(\gamma + i\delta)v = (\alpha\gamma - \beta\delta + (\alpha\delta + \beta\gamma)i)v.$$

It follows that V with $iv = Jv$ is a complex vector space. We emphasize that the complex vector space V contains the same elements as the real vector space V . The following theorem connects symplectic forms, real inner products, and complex inner products.¹⁶ By a complex inner product on a complex vector

¹⁶Paul R. Chernoff and Jerrold E. Marsden, *Properties of Infinite Dimensional Hamiltonian Systems*, p. 6, Theorem 2.

space W , we mean a function $h : W \times W \rightarrow \mathbb{C}$ that is conjugate symmetric, complex linear in the first argument, $h(w, w) \geq 0$ for all $w \in W$, and $h(w, w) = 0$ implies $w = 0$.

Theorem 2. Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle : H \times H \rightarrow \mathbb{R}$ and let ω be a weak symplectic form on H . Then there is a complex structure $J : H \rightarrow H$ and a real inner product s on H such that

$$s(x, y) = -\omega(Jx, y), \quad x, y \in H$$

is a real inner product on the real vector space H , and

$$h(x, y) = s(x, y) - i\omega(x, y), \quad x, y \in H$$

is a complex inner product on H with the complex structure J .

Furthermore, the following are equivalent:

1. The norm induced by h is equivalent with the norm induced by $\langle \cdot, \cdot \rangle$.
2. The norm induced by s is equivalent with the norm induced by $\langle \cdot, \cdot \rangle$.
3. ω is a strong symplectic form on the real Hilbert space H .

Proof. By the Riesz representation theorem,¹⁷ because ω is a bounded bilinear form there is a unique $A \in \mathcal{L}(H)$ such that

$$\omega(x, y) = \langle Ax, y \rangle, \quad x, y \in H. \quad (1)$$

Because ω is skew-symmetric,

$$\langle Ax, y \rangle = \omega(x, y) = -\omega(y, x) = -\langle Ay, x \rangle = \langle (-A)y, x \rangle.$$

On the other hand, because $\langle \cdot, \cdot \rangle$ is a real inner product, $\langle Ax, y \rangle = \langle x, A^*y \rangle = \langle A^*y, x \rangle$. Therefore $A^* = -A$.

$A^*A = (-A)A = -A^2$ and $AA^* = A(-A) = -A^2$, so A is normal. Therefore A has a **polar decomposition**:¹⁸ there is a unitary $U \in \mathcal{L}(H)$ and some $P \in \mathcal{L}(H)$ with $P \geq 0$, such that

$$A = UP,$$

and such that A, U, P commute; a fortiori, P is self-adjoint. If $Ax = 0$, then $\omega(x, y) = \langle Ax, y \rangle = \langle 0, y \rangle = 0$ for all $y \in H$, and because ω is weakly nondegenerate this implies that $x = 0$, hence A is one-to-one, which implies that P is one-to-one (this implication does not use that U is unitary). We have

$$A^* = (UP)^* = P^*U^* = PU^*, \quad A^* = -A = -UP = -PU,$$

hence

$$PU^* = P(-U).$$

¹⁷Walter Rudin, *Functional Analysis*, second ed., p. 310, Theorem 12.8.

¹⁸Walter Rudin, *Functional Analysis*, second ed., p. 332, Theorem 12.35.

Because P is one-to-one, this yields $U^* = -U$. But U is unitary, i.e. $U^*U = I$ and $UU^* = I$. Therefore $(-U)U = I$, i.e. $-U^2 = I$. This means that U is a complex structure on the real Hilbert space H . We write $J = U$.

The complex structure J satisfies, for $x, y \in H$,

$$\omega(Jx, Jy) = \langle AJx, Jy \rangle = \langle JAx, Jy \rangle = \langle Ax, J^*Jy \rangle = \langle Ax, y \rangle = \omega(x, y),$$

showing that J is a canonical transformation.

$s : H \times H \rightarrow \mathbb{R}$ is defined, for $x, y \in H$, by

$$s(x, y) = -\omega(Jx, y) = -\langle AJx, y \rangle = \langle (-J)Ax, y \rangle = \langle J^{-1}Ax, y \rangle = \langle Px, y \rangle.$$

It is apparent that s is bilinear. Because P is self-adjoint and $\langle \cdot, \cdot \rangle$ is symmetric,

$$s(x, y) = \langle Px, y \rangle = \langle x, Py \rangle = \langle Py, x \rangle = s(y, x),$$

showing that s is symmetric. Because $P \geq 0$, for any $x \in H$ we have $s(x, x) = \langle Px, x \rangle \geq 0$, namely s is positive. Also because $P \geq 0$, there is a unique $S \in \mathcal{L}(H)$, $S \geq 0$, satisfying $S^2 = P$.¹⁹ If $s(x, x) = 0$, we get

$$0 = \langle Px, x \rangle = \langle S^2x, x \rangle = \langle Sx, Sx \rangle = \|Sx\|^2,$$

hence $Sx = 0$ and so $Px = 0$, and because P is one-to-one, $x = 0$. Therefore s is positive definite, and thus is a real inner product on H .

$h : H \times H \rightarrow \mathbb{C}$ is defined, for $x, y \in H$, by

$$h(x, y) = s(x, y) - i\omega(x, y) = \langle Px, y \rangle - i\omega(x, y) = \langle Px, y \rangle - i\langle Ax, y \rangle.$$

For $x_1, x_2, y \in H$,

$$h(x_1 + x_2, y) = h(x_1, y) + h(x_2, y).$$

For $\alpha + i\beta \in \mathbb{C}$,

$$\begin{aligned} h((\alpha + i\beta)x, y) &= h(\alpha x + \beta Jx, y) \\ &= h(\alpha x, y) + \beta h(Jx, y) \\ &= \alpha h(x, y) + \beta \langle PJx, y \rangle - i\beta \langle AJx, y \rangle \\ &= \alpha h(x, y) + \beta \langle Ax, y \rangle - i\beta \langle A(-J^{-1})x, y \rangle \\ &= \alpha h(x, y) + \beta \omega(x, y) + i\beta \langle Px, y \rangle \\ &= \alpha h(x, y) + \beta \omega(x, y) + i\beta s(x, y) \\ &= \alpha h(x, y) + i\beta (s(x, y) - i\omega(x, y)) \\ &= \alpha h(x, y) + i\beta h(x, y) \\ &= (\alpha + i\beta)h(x, y). \end{aligned}$$

Therefore h is complex linear in its first argument. Because s is symmetric and ω is skew-symmetric, $h(x, y) = s(x, y) - i\omega(x, y)$ satisfies

$$h(y, x) = s(y, x) - i\omega(y, x) = s(x, y) + i\omega(x, y) = \overline{h(x, y)},$$

¹⁹Walter Rudin, *Functional Analysis*, second ed., p. 331, Theorem 12.33.

showing that h is conjugate symmetric. For $x \in H$,

$$h(x, x) = s(x, x) - i\omega(x, x) = s(x, x) \geq 0.$$

If $h(x, x) = 0$, then $s(x, x) = 0$, which implies that $x = 0$. Therefore h is a complex inner product on H with the complex structure J .

Suppose that ω is a strong symplectic form on the real Hilbert space H . That is, $\omega^\flat : H \rightarrow H^*$ is an isomorphism of Banach spaces. We shall show that A , from (1), is onto. For $y \in H$, define $\lambda : H \rightarrow \mathbb{R}$ by $\lambda(x) = \langle x, y \rangle$. Then $\lambda \in H^*$, so there is some $v \in H$ for which $\omega^\flat(v) = \lambda$. That is, $\omega(v, x) = \lambda(x) = \langle x, y \rangle = \langle y, x \rangle$ for all $x \in H$. But $\omega(v, x) = \langle Av, x \rangle$, so $\langle Av, x \rangle = \langle y, x \rangle$ for all $x \in H$, which implies that $Av = y$, and thus shows that A is onto, and hence invertible in $\mathcal{L}(H)$. Because $A = UP$ and A, U are invertible in $\mathcal{L}(H)$, P is invertible in $\mathcal{L}(H)$. Therefore $S, P = S^2, S \geq 0$, is invertible in $\mathcal{L}(H)$, whence

$$\begin{aligned} \|x\|^2 &= \|S^{-1}Sx\|^2 \\ &\leq \|S^{-1}\|^2 \|Sx\|^2 \\ &= \|S^{-1}\|^2 \langle Sx, Sx \rangle \\ &= \|S^{-1}\|^2 \langle Px, x \rangle \\ &= \|S^{-1}\|^2 s(x, x) \\ &= \|S^{-1}\|^2 \|x\|_s^2, \end{aligned}$$

and on the other hand

$$\|x\|_s^2 = s(x, x) = \langle Px, x \rangle \leq \|Px\| \|x\| \leq \|P\| \|x\|^2 = \|S\|^2 \|x\|^2.$$

so

$$\|x\| \leq \|S^{-1}\| \|x\|_s, \quad \|x\|_s \leq \|S\| \|x\|.$$

Namely this establishes that the norms $\|x\|^2 = \langle x, x \rangle$ and $\|x\|_s^2 = s(x, x)$ are equivalent. \square

8 Hamiltonian vector fields

Let E be a real Banach space and let $k \geq 1$; if we do not specify k we merely suppose that it is ≥ 1 . A C^k **vector field on U** , where U an open subset of E , is a C^k function $v : U \rightarrow E$.

Let v be a C^k , $k \geq 1$, vector field on E . For $x \in E$, an **integral curve of v through x** is a differentiable function $\phi : J \rightarrow E$, where J is some open interval in \mathbb{R} containing 0, that satisfies

$$\phi'(t) = (v \circ \phi)(t), \quad t \in J, \quad \phi(0) = x.$$

If $\psi : I \rightarrow E$ and $\phi : J \rightarrow E$ are integral curves of v through x , it is proved that for $t \in I \cap J$, $\psi(t) = \phi(t)$.²⁰ An integral curve of v through x , $\phi : J \rightarrow E$, is

²⁰Rodney Coleman, *Calculus on Normed Vector Spaces*, p. 194, Proposition 9.3.

said to be **maximal** if there is no integral curve of v through x whose domain strictly includes J . If $X : E \rightarrow E$ is a C^1 vector field, for each $x \in E$ it is proved that there is a unique maximal integral curve of v through x , denoted $\phi_x : J_x \rightarrow E$.²¹ A vector field $v : E \rightarrow E$ is called **complete** when $J_x = \mathbb{R}$ for each $x \in E$. For a vector field $v : E \rightarrow E$, a C^1 function $f : E \rightarrow \mathbb{R}$ is called a **first integral of v** if for any integral curve $\phi : J \rightarrow E$ of v , $f \circ \phi : J \rightarrow \mathbb{R}$ is constant. It is proved that if a vector field has a first integral $f : E \rightarrow \mathbb{R}$ such that $f^{-1}(c)$ is a compact subset of E for each $c \in \mathbb{R}$, then v is a complete vector field.²²

The **flow of v** is the function $\sigma : \Sigma_v \rightarrow E$, where

$$\Sigma_v = \bigcup_{x \in E} J_x \times \{x\},$$

such that for each $x \in E$, $\sigma(t, x) = \phi_x(t)$, $t \in J_x$. It is proved that Σ_v is an open subset of $\mathbb{R} \times E$, and that $\sigma : \Sigma_v \rightarrow E$ is continuous.²³ It is also proved that for any $k \geq 1$, if v is C^k then $\sigma : \Sigma_v \rightarrow E$ is C^k .²⁴ If $(s, x), (t, \sigma(s, x)), (t+s, x) \in \Sigma_v$, then²⁵

$$\sigma(t+s, x) = \sigma(t, \sigma(s, x)).$$

When v is a complete vector field, its flow is called a **global flow**. In this case, for $t \in \mathbb{R}$ we define $\sigma_t : E \rightarrow E$ by $\sigma_t(x) = \sigma(t, x)$. Then $\sigma_t^{-1} = \sigma_{-t}$, and thus each σ_t is a C^k diffeomorphism $E \rightarrow E$.

9 Differential forms

For vector spaces V and W and for $p \geq 1$, a function $f : V^p \rightarrow W$ is called **alternating** if $(v_1, \dots, v_p) \in V^p$ and $v_i = v_{i+1}$ for some $1 \leq i \leq p-1$ imply that $f(v_1, \dots, v_p) = 0$.

For Banach spaces E and F and for $p \geq 1$, we denote by $\mathcal{A}_p(E; F)$ the set of alternating elements of $\mathcal{L}_p(E; F)$. In particular, $\mathcal{A}_1(E; F) = \mathcal{L}_1(E; F) = \mathcal{L}(E; F)$. $\mathcal{A}_p(E; F)$ is a closed linear subspace of the Banach space $\mathcal{L}_p(E; F)$.²⁶ We define

$$\mathcal{A}_0(E; F) = \mathcal{L}_0(E; F) = F.$$

Let Σ_n be the set of permutation $\{1, \dots, n\}$, which has $n!$ elements. Let $\text{Sh}_{p,q}$ be the set of permutations σ of $\{1, \dots, p, p+1, \dots, p+q\}$ for which

$$\sigma(1) < \dots < \sigma(p), \quad \sigma(p+1) < \dots < \sigma(p+q).$$

The set $\text{Sh}_{p,q}$ has $\binom{p+q}{p} = \binom{p+q}{q}$ elements.

²¹Rodney Coleman, *Calculus on Normed Vector Spaces*, p. 194, Theorem 9.2.

²²Rodney Coleman, *Calculus on Normed Vector Spaces*, p. 207, Theorem 9.8.

²³Rodney Coleman, *Calculus on Normed Vector Spaces*, p. 213, Theorem 10.1.

²⁴Rodney Coleman, *Calculus on Normed Vector Spaces*, p. 222, Theorem 10.3.

²⁵Yvonne Choquet-Bruhat and Cecile DeWitt-Morette, *Analysis, Manifolds and Physics, Part I*, p. 551.

²⁶Henri Cartan, *Differential Forms*, p. 9.

For $f \in \mathcal{A}_p(E; \mathbb{R})$ and $g \in \mathcal{A}_q(E; \mathbb{R})$, we define $f \wedge g : E^p \times E^q \rightarrow \mathbb{R}$ by

$$\begin{aligned} & (f \wedge g)(x_1, \dots, x_p, x_{p+1}, \dots, x_{p+q}) \\ &= \sum_{\sigma \in \text{Sh}_{p,q}} \text{sgn}(\sigma) f(x_{\sigma(1)}, \dots, x_{\sigma(p)}) g(x_{\sigma(p+1)}, \dots, x_{\sigma(p+q)}). \end{aligned}$$

It is proved that $f \wedge g \in \mathcal{A}_{p+q}(E; \mathbb{R})$.²⁷

For $f \in \mathcal{A}_p(E; \mathbb{R})$ and $g \in \mathcal{A}_q(E; \mathbb{R})$,

$$\begin{aligned} \|f \wedge g\| &= \sup_{\|x_1\| \leq 1, \dots, \|x_{p+q}\| \leq 1} |(f \wedge g)(x_1, \dots, x_p, x_{p+1}, \dots, x_{p+q})| \\ &\leq \sup_{\|x_1\| \leq 1, \dots, \|x_{p+q}\| \leq 1} \sum_{\sigma \in \text{Sh}_{p,q}} |f(x_{\sigma(1)}, \dots, x_{\sigma(p)}) g(x_{\sigma(p+1)}, \dots, x_{\sigma(p+q)})| \\ &\leq \sup_{\|x_1\| \leq 1, \dots, \|x_{p+q}\| \leq 1} \sum_{\sigma \in \text{Sh}_{p,q}} \|f\| \|g\| \\ &= \binom{p+q}{p} \|f\| \|g\|, \end{aligned}$$

showing that the operator norm of the bilinear map $(f, g) \mapsto f \wedge g$, $\mathcal{A}_p(E; \mathbb{R}) \times \mathcal{A}_q(E; \mathbb{R})$ is $\leq \binom{p+q}{p}$, and thus is continuous.

One proves that for $f \in \mathcal{A}_p(E; \mathbb{R})$ and $g \in \mathcal{A}_q(E; \mathbb{R})$, then²⁸

$$g \wedge f = (-1)^{pq} f \wedge g.$$

It is also proved that for $f \in \mathcal{A}_p(E; \mathbb{R})$, $g \in \mathcal{A}_q(E; \mathbb{R})$, and $h \in \mathcal{A}_r(E; \mathbb{R})$, then²⁹

$$(f \wedge g) \wedge h = f \wedge (g \wedge h).$$

It thus makes sense to speak about $f_1 \wedge \dots \wedge f_n$. We remind ourselves that $\mathcal{A}_1(E; \mathbb{R}) = \mathcal{L}(E; \mathbb{R}) = E^*$. It is proved that if $f_1, \dots, f_n \in E^*$, then $f_1 \wedge \dots \wedge f_n \in \mathcal{A}_n(E; \mathbb{R})$ satisfies

$$f_1 \wedge \dots \wedge f_n(x_1, \dots, x_n) = \sum_{\sigma \in \Sigma_n} \text{sgn}(\sigma) f_1(x_{\sigma(1)}) \dots f_n(x_{\sigma(n)}), \quad (x_1, \dots, x_n) \in E^n,$$

and that $f_1, \dots, f_n \in E^*$ are linearly independent if and only if $f_1 \wedge \dots \wedge f_n \neq 0$.³⁰

Let U be an open subset of the Banach space E . For $k \geq 0$ and $p \geq 0$, a C^k **differential form of degree p on U** is a C^k function

$$\alpha : U \rightarrow \mathcal{A}_p(E; \mathbb{R}).$$

We abbreviate “differential form of degree p ” as “differential p -form”. In particular, a C^k differential 0-form is a C^k function $U \rightarrow \mathcal{A}_0(E; \mathbb{R}) = \mathbb{R}$. We denote

²⁷Henri Cartan, *Differential Forms*, pp. 12–14.

²⁸Henri Cartan, *Differential Forms*, p. 14, Proposition 1.5.1.

²⁹Henri Cartan, *Differential Forms*, p. 15, Proposition 1.5.2.

³⁰Henri Cartan, *Differential Forms*, p. 16, Proposition 1.6.1.

by $\Omega_p^{(k)}(U, \mathbb{R})$ the set of C^k differential p -forms on U . It is apparent that this is a real vector space.

For a C^k function $f : U \rightarrow \mathbb{R}$, with $k \geq 1$, the derivative f' is C^{k-1} function $U \rightarrow \mathcal{L}(E; \mathbb{R}) = \mathcal{A}_1(E; \mathbb{R})$, hence $f' \in \Omega_p^{(k-1)}(U)$.

For $\alpha \in \Omega_p^{(k)}(U, \mathbb{R})$ and $\beta \in \Omega_q^{(k)}(U, \mathbb{R})$, we define $\alpha \wedge : U \rightarrow \mathcal{A}_{p+q}(E; \mathbb{R})$ by

$$(\alpha \wedge \beta)(x) = (\alpha(x)) \wedge (\beta(x)), \quad x \in U.$$

It is proved that $\alpha \wedge \beta \in \Omega_{p+q}^{(k)}(U, \mathbb{R})$.³¹

Suppose that $k \geq 1$ and $\alpha \in \Omega_p^{(k)}(U, \mathbb{R})$, i.e. $\alpha : U \rightarrow \mathcal{A}_p(U; \mathbb{R})$ is a C^k function. Then the derivative is the C^{k-1} function

$$\alpha' : U \rightarrow \mathcal{L}(E; \mathcal{A}_p(E; \mathbb{R})).$$

We define $d\alpha : U \rightarrow \mathcal{A}_{p+1}(E; \mathbb{R})$ by

$$(d\alpha)(x)(\xi_0, \xi_1, \dots, \xi_p) = \sum_{i=0}^p (-1)^i \alpha'(x)(\xi_i)(\xi_0, \dots, \hat{\xi}_i, \dots, \xi_p)$$

It is proved that $d\alpha \in \Omega_{p+1}^{(k-1)}(U, \mathbb{R})$.³²

In particular, if $f : U \rightarrow \mathbb{R}$ is a C^k function, then $df \in \Omega_1^{(k-1)}(U, \mathbb{R})$ is the function $df : U \rightarrow \mathcal{A}_1(E; \mathbb{R}) = \mathcal{L}(E; \mathbb{R})$ defined by

$$(df)(x)(\xi) = f'(x)(\xi), \quad x \in U, \quad \xi \in E.$$

Thus, $df = f'$.

For $\alpha \in \Omega_p^{(k)}(U, \mathbb{R})$ and $\beta \in \Omega_q^{(k)}(U, \mathbb{R})$ with $k \geq 1$, it is a fact that³³

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

In particular, an element f of $\Omega_0^{(k)}(U, \mathbb{R})$ is a C^k function $U \rightarrow \mathbb{R}$, for which, because $f \wedge \beta = f\beta$,

$$d(f\beta) = (df) \wedge \beta + f(d\beta).$$

For $\alpha \in \Omega_p^{(k)}(U, \mathbb{R})$, with $k \geq 2$,³⁴

$$d(d\alpha) = 0.$$

Let $\alpha \in \Omega_p^{(k)}(U, \mathbb{R})$, let V be an open subset of a Banach space F , and let $\phi : V \rightarrow U$ be a C^{k+1} function. The **pullback of α by f** , denoted $\phi^* \alpha : V \rightarrow \mathcal{A}_p(F; \mathbb{R})$, is an element of $\Omega_p^{(k)}(V, \mathbb{R})$ satisfying³⁵

$$(\phi^* \alpha)(y)(\eta_1, \dots, \eta_p) = \alpha(\phi(y))(\phi'(y)(\eta_1), \dots, \phi'(y)(\eta_p)), \quad (\eta_1, \dots, \eta_p) \in F^p.$$

³¹Henri Cartan, *Differential Forms*, p. 19, §2.2.

³²Henri Cartan, *Differential Forms*, pp. 20–21, §2.3.

³³Henri Cartan, *Differential Forms*, p. 22, Theorem 2.4.2.

³⁴Henri Cartan, *Differential Forms*, p. 23, Theorem 2.5.1.

³⁵Henri Cartan, *Differential Forms*, p. 29, Proposition 2.8.1.

The pullback satisfies, for $\alpha \in \Omega_p^{(k)}(U, \mathbb{R})$ and $\beta \in \Omega_q^{(k)}(U, \mathbb{R})$,

$$\phi^*(\alpha \wedge \beta) = (\phi^*\alpha) \wedge (\phi^*\beta),$$

which is an element of $\Omega_{p+q}^{(k)}(V, \mathbb{R})$. It also satisfies, if $\phi : V \rightarrow U$ and $f : U \rightarrow \mathbb{R}$ are C^1 ,

$$\phi^*(df) = d(\phi^*f),$$

where $(\phi^*f)(y) = f(\phi(y))$.

10 Contractions and Lie derivatives

Let U be an open subset of a Banach space E , let $k \geq 1$, $p \geq 1$, let v be a C^k vector field on U , and let $\alpha \in \Omega_p^{(k)}(U, \mathbb{R})$. We define $\iota_v \alpha : U \rightarrow \mathcal{A}_{p-1}(E; \mathbb{R})$ by

$$(\iota_v \alpha)(x)(v_1, \dots, v_{p-1}) = \alpha(v(x), v_1, \dots, v_{p-1}), \quad (v_1, \dots, v_{p-1}) \in E^{p-1}.$$

(It is straightforward to check that indeed $(\iota_v \alpha)(x) \in \mathcal{A}_{p-1}(E; \mathbb{R})$.) It is proved that $\iota_v \alpha : U \rightarrow \mathcal{A}_{p-1}(E; \mathbb{R})$ is C^k , and thus $\iota_v \alpha \in \Omega_{p-1}^{(k)}(U, \mathbb{R})$.³⁶ For $p = 0$, with $f \in \Omega_0^{(k)}(U, \mathbb{R})$, i.e. f is a C^k function $U \rightarrow \mathbb{R}$, we define $\iota_v f = 0$. We call $\iota_v \alpha$ the **contraction of α by v** .

It can be proved that if $\alpha \in \Omega_p^{(k)}(U, \mathbb{R})$ and $\beta \in \Omega_q^{(k)}(U, \mathbb{R})$,

$$\iota_v(\alpha \wedge \beta) = (\iota_v \alpha) \wedge \beta + (-1)^p \alpha \wedge \iota_v \beta.$$

Also, for a C^k vector field w on U ,

$$\iota_v(\iota_w \alpha) = -\iota_w(\iota_v \alpha),$$

and hence $\iota_v^2 \alpha = 0$. And $(v, \alpha) \mapsto \iota_v \alpha$ is bilinear.

For a C^k vector field v on U and $\alpha \in \Omega_p^{(k)}(U, \mathbb{R})$, the **Lie derivative of α with respect to v** is³⁷

$$\mathcal{L}_v \alpha = d(\iota_v \alpha) + \iota_v d\alpha \in \Omega_p^{(k)}(U, \mathbb{R}).$$

The Lie derivative satisfies

$$\mathcal{L}_v(\alpha \wedge \beta) = (\mathcal{L}_v \alpha) \wedge \beta + \alpha \wedge \mathcal{L}_v \beta.$$

If ω is a weak symplectic form on a Banach space E and v is a C^1 vector field on E , we say that v is a **symplectic vector field** if

$$\mathcal{L}_v \omega = 0.$$

If there is some C^1 function $H : E \rightarrow \mathbb{R}$ such that

$$\iota_v \omega = -dH,$$

³⁶cf. Serge Lang, *Differential and Riemannian Manifolds*, p. 137, V, §5.

³⁷cf. Serge Lang, *Differential and Riemannian Manifolds*, pp. 138–141, V, §5.

we say that v is a **Hamiltonian vector field with Hamiltonian function** H . If v is a Hamiltonian vector field with Hamiltonian function H , then

$$\mathcal{L}_v\omega = d(\iota_v\omega) + \iota_v d\omega = d(\iota_v\omega) = d(-dH) = -d^2H = 0,$$

showing that if a vector field is Hamiltonian then it is symplectic. (This is analogous to the statement that if a differential form is exact then it is closed.)