Weak symplectic forms and differential calculus in Banach spaces

Jordan Bell
jordan.bell@gmail.com
Department of Mathematics, University of Toronto

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

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

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

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

[3] Andreas 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*.

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

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

so $B^\flat : E \to E^*$ is continuous.

We call a continuous bilinear form $B : E \times E \to F$ **weakly nondegenerate** if $B^\flat : E \to E^*$ is one-to-one. Since $B^\flat$ is linear, this is equivalent to the statement that $B^\flat(e) = 0$ 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 $T : E \to F$ that is continuous such that $T^{-1}F \to E$ is continuous. Equivalently, to say that $T : E \to F$ is an isomorphism of Banach spaces means that $T : E \to F$ is a bijective bounded linear map such that $T^{-1} : F \to E$ is a bounded linear map. It follows from the open mapping theorem that if $T : E \to F$ is an onto bounded linear isomorphism, hence is an isomorphism of Banach spaces.

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

For a real vector space V and a bilinear form $B : V \times V \to \mathbb{R}$, 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 $E_1, \ldots, E_p$ and F, let $\mathscr{L}(E_1, \ldots, E_p; F)$ denote the set of continuous multilinear maps $E_1 \times \cdots E_p \to F$. For a multilinear map $T : E_1 \times \cdots \times E_p \to F$ to be continuous it is equivalent that

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

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

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

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

For Banach spaces E, F, G, define

$$\phi : \mathscr{L}(E, F; G) \to \mathscr{L}(E; \mathscr{L}(F, G))$$

[4] Henri Cartan, *Differential Calculus*, p. 22, Theorem 1.8.1.
[5] Henri Cartan, *Differential Calculus*, p. 20, Theorem 1.7.3.

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

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 $f : U \to F$ is said to be **differentiable at** a if (i) f is continuous at a and (ii) there is a linear mapping $g : E \to F$ such that

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

as $x \to a$ 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 $f'(a) : E \to F$ is continuous at a and therefore, being linear, is continuous on E, namely $f'(a) \in \mathscr{L}(E;F)$.[7]

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

We say that $f : U \to F$ is C^1, also called **continuously differentiable**, if (i) f is differentiable on U and (ii) $f' : U \to \mathscr{L}(E;F)$ 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 $f : U \to F$ and $g : V \to G$ be continuous. Suppose that $a \in U$ and that $f(a) \in V$. We define $g \circ f : f^{-1}(V) \to G$ on $f^{-1}(V)$. One proves that if f is differentiable at a and g is differentiable at $f(a)$, then $h = g \circ f : f^{-1}(V) \to F$ is differentiable at a and satisfies[8]

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

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

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

satisfies[9]

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

4 Symplectic forms

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

[6]Henri Cartan, *Differential Calculus*, p. 23, §1.9.
[7]Henri Cartan, *Differential Calculus*, p. 25.
[8]Henri Cartan, *Differential Calculus*, p. 27, Theorem 2.2.1.
[9]Henri Cartan, *Differential Calculus*, p. 31, Theorem 2.4.4.

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

5 Hamiltonian functions

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

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

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

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

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

For a linear operator A in E, we define

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

$\mathscr{G}(A)$ is a linear subspace of $E \times E$. We say that A is **closed** if $\mathscr{G}(A)$ 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 $\mathscr{D}(A)$ with the norm

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

is a Banach space.

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

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

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

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

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

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

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

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

This shows that H is differentiable on the Banach space $\mathscr{D}(A)$, with derivative $H' : \mathscr{D}(A) \to \mathscr{D}(A)^*$ defined by[11]

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

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

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

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

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

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

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

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

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

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

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

What we established earlier tells us that

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

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 $x, y \in \mathscr{D}(A)$,

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

and thus $K(u) = K(0) = C$ for all $u \in \mathscr{D}(A)$. 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 $U_t \in \mathscr{L}(E)$ 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 $x \in \mathscr{D}(A)$,

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

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

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

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

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

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

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

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 $w \in \mathscr{D}(A)$. Using this and the fact that A is ω-skew,

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

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 $x \in \mathscr{D}(A)$, $AU_t x = U_t Ax$. (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,

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

□

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

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

7 Hilbert spaces

For a real vector space V, a **complex struture on** V is a linear map $J : V \to V$ 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 space W, we mean a function $h : W \times W \to \mathbb{C}$ 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 $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$ and let ω be a weak symplectic form on H. Then there is a complex structure $J : H \to H$ and a real inner product s on H such that*

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

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

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

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 $A \in \mathscr{L}(H)$ such that

$$\omega(x, y) = \langle Ax, y \rangle, \qquad x, y \in H. \tag{1}$$

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 $U \in \mathscr{L}(H)$ and some $P \in \mathscr{L}(H)$ with $P \geq 0$, such that

$$A = UP,$$

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

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

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

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

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

hence

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

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.

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

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

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 $S \in \mathscr{L}(H)$, $S \geq 0$, satisfying $S^2 = P$.[19] If $s(x, x) = 0$, we get

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

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.

$h : H \times H \to \mathbb{C}$ 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 $x_1, x_2, y \in H$,

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

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

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

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

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

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

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, $\omega^\flat : H \to H^*$ is an isomorphism of Banach spaces. We shall show that A, from (1), is onto. For $y \in H$, define $\lambda : H \to \mathbb{R}$ 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 $\mathscr{L}(H)$. Because $A = UP$ and A, U are invertible in $\mathscr{L}(H)$, P is invertible in $\mathscr{L}(H)$. Therefore S, $P = S^2$, $S \geq 0$, is invertible in $\mathscr{L}(H)$, whence

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

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

$$\|x\| \leq \left\|S^{-1}\right\| \|x\|_s, \qquad \|x\|_s \leq \|S\| \|x\|.$$

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

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 $v : U \to E$.

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 $\phi : J \to E$, where J is some open interval in $\mathbb{R}$ containing 0, that satisfies

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

If $\psi : I \to E$ and $\phi : J \to E$ 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, $\phi : J \to E$, is said to be **maximal** if there is no integral curve of v through x whose domain strictly includes J. If $X : E \to E$ 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 $\phi_x : J_x \to E$.[21] A vector field $v : E \to E$ is called **complete** when $J_x = \mathbb{R}$ for each $x \in E$. For a vector field $v : E \to E$, a C^1 function $f : E \to \mathbb{R}$ is called a **first integral of** v if for any integral curve $\phi : J \to E$ of v, $f \circ \phi : J \to E$ is constant. It is proved that if a vector field has a first integral $f : E \to \mathbb{R}$ 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 $\sigma : \Sigma_v \to E$, 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 $\sigma : \Sigma_v \to E$ is continuous.[23] It is also proved that for any $k \geq 1$, if v is C^k then $\sigma : \Sigma_v \to E$ is C^k.[24] If $(s,x), (t, \sigma(s,x)), (t+s, x) \in \Sigma_v$, then[25]

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

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

[20] Rodney Coleman, *Calculus on Normed Vector Spaces*, p. 194, Proposition 9.3.
[21] Rodney Coleman, *Calculus on Normed Vector Spaces*, p. 194, Theorem 9.2.
[22] Rodney Coleman, *Calculus on Normed Vector Spaces*, p. 207, Theorem 9.8.
[23] Rodney Coleman, *Calculus on Normed Vector Spaces*, p. 213, Theorem 10.1.
[24] Rodney Coleman, *Calculus on Normed Vector Spaces*, p. 222, Theorem 10.3.
[25] Yvonne Choquet-Bruhat and Cecile DeWitt-Morette, *Analysis, Manifolds and Physics, Part I*, p. 551.

9 Differential forms

For vector spaces V and W and for $p \geq 1$, a function $f : V^p \to W$ 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 $\mathscr{A}_p(E; F)$ the set of alternating elements of $\mathscr{L}_p(E; F)$. In particular, $\mathscr{A}_1(E; F) = \mathscr{L}_1(E; F) = \mathscr{L}(E; F)$. $\mathscr{A}_p(E; F)$ is a closed linear subspace of the Banach space $\mathscr{L}_p(E; F)$.[26] We define

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

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

$$\sigma(1) < \cdots < \sigma(p), \qquad \sigma(p+1) < \cdots < \sigma(p+q).$$

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

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

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

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

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

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

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

One proves that for $f \in \mathscr{A}_p(E; \mathbb{R})$ and $g \in \mathscr{A}_q(E; \mathbb{R})$, then[28]

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

It is also proved that for $f \in \mathscr{A}_p(E; \mathbb{R})$, $g \in \mathscr{A}_q(E; \mathbb{R})$, and $h \in \mathscr{A}_r(E; \mathbb{R})$, then[29]

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

[26] Henri Cartan, *Differential Forms*, p. 9.
[27] Henri Cartan, *Differential Forms*, pp. 12–14.
[28] Henri Cartan, *Differential Forms*, p. 14, Proposition 1.5.1.
[29] Henri Cartan, *Differential Forms*, p. 15, Proposition 1.5.2.

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

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

and that $f_1, \ldots, f_n \in E^*$ are linearly independent if and only if $f_1 \wedge \cdots \wedge f_n = 0$.[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

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

We abbreviate "differential form of degree p" as "differential p-form". In particular, a C^k differential 0-form is a C^k function $U \to \mathscr{A}_0(E;\mathbb{R}) = \mathbb{R}$. We denote 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 $f : U \to \mathbb{R}$, with $k \geq 1$, the derivative f' is C^{k-1} function $U \to \mathscr{L}(E;\mathbb{R}) = \mathscr{A}_1(E;\mathbb{R})$, 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 $\alpha \wedge : U \to \mathscr{A}_{p+q}(E;\mathbb{R})$ by

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

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. $\alpha : U \to \mathscr{A}_p(U;\mathbb{R})$ is a C^k function. Then the derivative is the C^{k-1} function

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

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

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

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

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

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

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 that[33]

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

[30] Henri Cartan, *Differential Forms*, p. 16, Proposition 1.6.1.
[31] Henri Cartan, *Differential Forms*, p. 19, §2.2.
[32] Henri Cartan, *Differential Forms*, pp. 20–21, §2.3.
[33] Henri Cartan, *Differential Forms*, p. 22, Theorem 2.4.2.

In particular, an element f of $\Omega_0^{(k)}(U,\mathbb{R})$ is a C^k function $U \to \mathbb{R}$, 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 $\phi : V \to U$ be a C^{k+1} function. The **the pullback of** α **by** f, denoted $\phi^*\alpha : V \to \mathscr{A}_p(F;\mathbb{R})$, is an element of $\Omega_p^{(k)}(V,\mathbb{R})$ satisfying[35]

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

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 $\phi : V \to U$ and $f : U \to \mathbb{R}$ 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 $\iota_v\alpha : U \to \mathscr{A}_{p-1}(E;\mathbb{R})$ by

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

(It is straightforward to check that indeed $(\iota_v\alpha)(x) \in \mathscr{A}_{p-1}(E;\mathbb{R})$.) It is proved that $\iota_v\alpha : U \to \mathscr{A}_{p-1}(E;\mathbb{R})$ 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 $U \to \mathbb{R}$, 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.

[34] Henri Cartan, *Differential Forms*, p. 23, Theorem 2.5.1.
[35] Henri Cartan, *Differential Forms*, p. 29, Proposition 2.8.1.
[36] cf. Serge Lang, *Differential and Riemannian Manifolds*, p. 137, V, §5.

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** is[37]

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

The Lie derivative satisfies

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

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

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

If there is some C^1 function $H : E \to E$ such that

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

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

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

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.)

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