Mixba'al
Revista Metropolitana de Matemáticas
No. 1, Vol. II, junio 2011, 9–28

VARIEDADES LAGRANGIANAS Y APLICACIONES A LA FÍSICA

JOAQUÍN DELGADO

Resumen. Este artículo panorámico pretende ilustrar la importancia de las subvariedades Lagrangianas en la Física Matemática, particularmente en aquellas ramas de un alto contenido geométrico. En la primera sección revisamos los conceptos básicos de los espacios vectoriales simplécticos y los subespacios Lagrangianos, contrapartes de sus versiones no lineales, las variedad simpléctica y subvariedades Lagrangianas. Presentamos dos ejemplos que muestran su relevancia: la teoría de las transformaciones canónicas en términos de funciones generadoras y la formulación matemática de la Termodinámica.

Dedicado con aprecio a Ernesto A. Lacomba en su 65 aniversario

1. Formas diferenciales y derivada exterior

Una 1–forma diferencial en $\mathbb{R}^n$ con coordenadas $x = (x_1, \ldots, x_n)$ es una expresión formal

$$\alpha = \sum_{i=1}^{n} a_i(x) dx_i. \tag{1}$$

Una 2–forma diferencial en $\mathbb{R}^n$ es una expresión formal

$$\omega = \sum_{1 \leq i < j \leq n}^{n} a_{ij}(x) dx_i \wedge dx_j. \tag{2}$$

Las 1–formas diferenciales se integran sobre curvas: si $C\colon [0,1] \to \mathbb{R}^n$, $C(t) = (x_1(t), \ldots, x_n(t)) = x(t)$, es suave a trozos,

$$\int_C \alpha = \int_0^1 \sum_{i=1}^{n} a_i(x(t)) \frac{dx_i}{dt}\, dt,$$

y la integral no depende de la parametrización de la curva C.

Las 2–formas diferenciales se integran sobre superficies: si $S\colon [0,1] \times [0,1] \to \mathbb{R}^n$, $S(u,v) = (x_1(u,v), \ldots, x_k(u,v)) = x(u,v)$ es una superficie parametrizada suave excepto en un conjunto de medida cero, se define

$$\int_S \omega = \int_{[0,1]^2} \sum_{1 \leq i < j \leq n}^{n} a_{ij}(x(u,v)) \frac{\partial(x_i, x_j)}{\partial(u,v)}\, du\, dv$$

donde el determinante Jacobiano,

$$\frac{\partial(x_i, x_j)}{\partial(u,v)} = \begin{vmatrix} \dfrac{\partial x_i}{\partial u} & \dfrac{\partial x_i}{\partial v} \\ \dfrac{\partial x_j}{\partial u} & \dfrac{\partial x_j}{\partial v} \end{vmatrix}.$$

Estas integrales son invariantes bajo reparametrizaciones de la superficie que preservan la orientación.

2010 *Mathematics Subject Classification.* 37J05, 70H03, 70H05, 70H15, 53D05, 53D12.

Palabras clave. Variedades simplécticas, Transformación canónica, Mecánica, Termodinámica.

En general una k–forma diferencial es una expresión formal

$$\beta = \sum_{1\leq i_1<i_2<\cdots<i_k\leq n} a_{i_1 i_2\cdots i_k}(x)dx_{i_1}\wedge dx_{i_2}\wedge\cdots dx_{i_k} \tag{3}$$

que se integra sobre k-superficies parametrizadas,

$$\begin{aligned} S(u_1,u_2,\ldots,u_k) &= (x_1(u_1,u_2,\ldots,u_k),\ldots,x_n(u_1,u_2,\ldots,u_k)) \\ &\equiv x(u_1,u_2,\ldots,u_k), \quad (u_1,u_2,\ldots,u_k)\in[0,1]^k \end{aligned}$$

$$\begin{aligned} \int_S \beta &= \int_{[0,1]^k} \sum_{1\leq i<j\leq n}^{n} a_{i_1 i_2\cdots i_k}(x(u_1,u_2,\ldots,u_k))\times \\ &\qquad \frac{\partial(x_1,x_2,\ldots,x_k)}{\partial(u_1,u_2,\ldots,u_k)}\,du_1du_2\cdots du_k, \end{aligned}$$

donde

$$\frac{\partial(x_1,x_2,\ldots,x_k)}{\partial(u_1,u_2,\ldots,u_k)} = \begin{vmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_k}{\partial u_1} & \cdots & \frac{\partial x_k}{\partial u_k} \end{vmatrix}.$$

Definición 1. Sea β una k–forma diferencial (3), $f\colon\mathbb{R}^n\to\mathbb{R}^n$ un cambio de coordenadas $x=f(u)$; el *pullback* de β bajo el cambio de coordenadas es

$$\begin{aligned} f^*(\beta) &= \sum_{1\leq i<j\leq n}^{n} a_{i_1 i_2\cdots i_k}(x(u_1,u_2,\ldots,u_k))\times \\ &\qquad \frac{\partial(x_1,x_2,\ldots,x_k)}{\partial(u_1,u_2,\ldots,u_k)}\,du_1\wedge du_2\wedge\cdots\wedge du_k. \end{aligned} \tag{4}$$

La *diferencial exterior* de una 1–forma (1) se define como

$$d\alpha = \sum_{j=1}^{n} \frac{\partial a_i}{\partial x_j}(x)dx_j\wedge dx_i \tag{5}$$

y de la 2–forma (2) como

$$d\omega = \sum_k \sum_{1\leq i<j\leq n} \frac{\partial a_{ij}}{\partial x_k}(x)\,dx_k\wedge dx_i\wedge dx_j. \tag{6}$$

En general,

Definición 2. La *diferencial exterior de la k–forma* (3) es la $k+1$–forma

$$d\beta = \sum_l \sum_{1\leq i_1<i_2<\cdots<i_k\leq n} \frac{\partial a_{i_1 i_2\cdots i_k}}{\partial x_l}(x)\,dx_l\wedge dx_{i_1}\wedge dx_{i_2}\wedge\cdots dx_{i_k}. \tag{7}$$

Definición 3. Una k-forma β se dice *cerrada*, si $d\beta=0$ y *exacta*, si existe una $k-1$-forma γ tal que $\beta=d\gamma$.

El siguiente resultado se conoce como el *Lema de Poincaré*

Lema 1.1. *Si β es una k-forma cerrada entonces para cada punto $x\in\mathbb{R}^n$ existe una vecindad $U\subset\mathbb{R}^n$ de x donde β es exacta.*

Observación 1. El Lema de Poincaré depende fuertemente de la topología del dominio donde esté definida la forma cerrada. Por ejemplo si sólo está definida en una vecindad de x en $\mathbb{R}^n$, es suficiente que la vecindad sea estrellada respecto de x, i.e., si $y\in U$ entonces todo el segmento $\overline{xy}$ está contenido en U.

Los siguientes resultados se conocen como naturalidad del pullback.

Proposición 1.2. *Se satisfacen la siguientes propiedades*

1. $f^*(\alpha\wedge\beta)=f^*(\alpha)\wedge f^*(\beta)$.
2. $f^*(d\alpha)=d\,f^*(\alpha)$.

2. Formas simplécticas y campos Hamiltonianos

Una k–forma se puede contraer con un campo vectorial resultando en una $k-1$ forma.

Definición 4. La *contracción* de la k–forma β (3) y el campo vectorial

$$X = \sum_i X^i(x)\frac{\partial}{\partial x_i}$$

es la $k-1$–forma

$$i_X\beta \quad = \quad \sum_{i_1} \sum_{1\leq i_1<i_2<\cdots<i_k\leq n} a_{i_1 i_2\cdots i_k}(x)X^{i_1}(x) \wedge dx_{i_2} \wedge \cdots dx_{i_k}$$

Observación 2. En terminología tensorial se dice que $a_{i_1 i_2\cdots i_k}$ es un tensor covariante antisimétrico y que X^i es un tensor contravariante. Entonces $a_{i_1 i_2\cdots i_k}X^{i_1}$ es la contracción de los tensores respecto del índice i_1, donde el índice repetido i_1 indica suma.

Sea ω una 2–forma y X, Y dos campos vectoriales. Entonces podemos contraer repetidamente resultando en una función escalar de x

$$i_Y i_X\omega \equiv \omega(X,Y) \tag{8}$$

Definición 5. Una 2–forma se dice que es *no degenerada*, si $\omega(X,Y)=0$, para todo campo vectorial Y implica $X=0$. En otras palabras, la aplicación de campos vectoriales a 1–formas en $\mathbb{R}^n$, $\mathfrak{X}(\mathbb{R}^n) \to \Lambda^1(\mathbb{R}^n)$, $X \mapsto i_X\omega$ es inyectiva.

Definición 6. Una *forma simpléctica* ω en $\mathbb{R}^{2n}$ es una 2-forma cerrada no degenerada; si además $\omega = d\alpha$ la forma se dice *exacta simpléctica.* Se dice que $(\mathbb{R}^{2n},\omega)$ es un *espacio simpléctico.*

Ejemplo 1. La 2-forma en $\mathbb{R}^{2n}$ con coordenadas (q,p):

$$\omega = \sum_{i=1}^n dp_i \wedge dq_i \tag{9}$$

es una forma simpléctica exacta. Por ejemplo $\omega = d(\sum_i p_i dq_i)$, pero también $\omega = -d(\sum_i q_i dp_i)$. La 1–forma

$$\sum_{i=1}^n p_i\, dq_i. \tag{10}$$

La formas (10) y (9) se llaman la 1–forma y la 2–forma canónicas en $\mathbb{R}^{2n}$.

La forma (9) es no degenerada porque si

$$X = \sum_j \left(A^j \frac{\partial}{\partial q_j} + B^j \frac{\partial}{\partial p_j}\right)$$

entonces

$$i_X\omega = \sum_{i=1}^n B^i dq_i - A^i dp_i = 0 \tag{11}$$

lo cual implica $A^i = B^i = 0$, $i = 1,2,\ldots,n$.

Observación 3. La forma simpléctica (9) se llama la forma simpléctica canónica de $\mathbb{R}^{2n}$. Observe que la forma (9) tiene coeficientes constantes (no dependen de x). En este caso la 2-forma se puede identificar con una forma bilineal antisimétrica como sigue: Sean

$$X = \sum_i A^i \frac{\partial}{\partial q_i} + B^i \frac{\partial}{\partial p_i}, \quad Y = \sum_i L^i \frac{\partial}{\partial q_i} + M^i \frac{\partial}{\partial p_i}$$

campos vectoriales, entonces (8) se reduce a

$$\omega(X,Y) = i_Y(\sum_{i=1}^n B^i dq_i - A^i dp_i) = \sum_{i=1}^n B^i L^i - A^i M^i$$

que se puede escribir como

(12) $$\omega(X,Y) = (L,M)\begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix}.$$

La matriz

(13) $$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

se llama la matriz simpléctica canónica de $\mathbb{R}^{2n}$. Con esto $(\mathbb{R}^{2n}, J)$ es un *espacio vectorial simpléctico.*

Observación 4. El Teorema de Darboux afirma que localmente, toda forma simpléctica es equivalente a la forma canónica bajo un cambio de coordenadas simpléctico.

Definición 7. Un campo vectorial

$$X = \sum_{i=1}^{n} A_i(q,p)\frac{\partial}{\partial q_i} + B_i(q,p)\frac{\partial}{\partial p_i}$$

se dice *Hamiltoniano*, si existe una función escalar $H(q,p)$ diferenciable tal que

(14) $$i_X\Omega = -dH$$

Proposición 2.1. *Sea*

$$X = \sum_{i=1}^{n} A_i(q,p)\frac{\partial}{\partial q_i} + B_i(q,p)\frac{\partial}{\partial p_i}$$

un campo vectorial con Hamiltoniano $H(q,p)$. Entonces las componentes del campo vectorial son

$$X = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial}{\partial p_i}$$

y las curvas integrales satisfacen las ecuaciones de Hamilton

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}. \end{aligned}$$

Demostración. De la expresión (11) y por ser Hamiltoniano, se tiene

$$i_X\omega = \sum_{i=1}^{n} B^i dq_i - A^i dp_i = -\sum_{i=1}^{n} \frac{\partial H}{\partial q_i}dq_i + \frac{\partial H}{\partial p_i}dp_i$$

comparando coeficientes se sigue el resultado. □

3. Funciones generadoras de transformaciones canónicas

Las transformaciones simplécticas son aquellas que preservan la forma de las ecuaciones de Hamilton. Una clase importante se obtiene a partir de lo que se llama una *función generadora.* En esta sección recordamos los aspectos más elementales que son necesarios para desarrollar con soltura la ecuación de Hamilton–Jacobi.

Definición 8. *La forma simpléctica en $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ con coordenadas $z = (q,p)^T$* es la forma bilineal antisimétrica definida por la matriz

(15) $$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

donde I_n denota la matriz identidad $n \times n$, vgr.,

(16) $$\omega(z,w) = z^T J w, \quad z,w \in \mathbb{R}^{2n}.$$

Observación 5. Notese que la matriz J satisface las siguientes propiedades: $J^T = -J$, $J^2 = -I_{2n}$, $J^3 = -J$, $J^{-1} = -J$.

Definición 9. Una transformación lineal $M\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ se dice *simpléctica* si $M^TJM = J$, equivalentemente $\omega(Mz, Mw) = \omega(x, w)$. *Un difeomorfismo* $f\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ *(cambio de coordenadas) se dice simpléctico* (simpléctica) si $M = Df(x)\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ es simpléctica para toda $x \in \mathbb{R}^{2n}$.

Lema 3.1. *Si M es una matriz simpléctica, entonces $M^{-1} = -JM^TJ$.*

Demostración.

$$(-JM^TJ)M = -J(M^TJM) = -J^2 = I_{2n}$$

de donde se sigue el resultado. □

Proposición 3.2. *El conjunto de transformaciones lineales simplécticas forman un subgrupo del grupo general lineal $GL(n, \mathbb{R})$.*

Demostración. Es claro que la identidad es simpléctica. Si M y N son simplécticas entonces $M^TJM = J$ y $N^TJN = J$, luego

$$(MN)^TJ(MN) = N^T(M^TJ)MN = N^TJN = J$$

y por lo tanto MN es simpléctica. Para ver que M^{-1} es simpléctica, usaremos el lema anterior: $M^{-1} = -JM^TJ$.

$$\begin{aligned}(M^{-1})^TJM^{-1} &= (-JM^TJ)^TJM^{-1} = -(JMJ)JM^{-1}\\ &= JMM^{-1} = J.\end{aligned}$$

□

Lema 3.3. *Sea M una matriz simpléctica; entonces $|M| = \pm 1$.*

Demostración. Como $M^TJM = J$ se sigue que $|M^T||J||M| = |J|$, luego $|M|^2 = 1$, por lo tanto $|M| = \pm 1$. □

Observación 6. De hecho se puede probar que para una matriz simpléctica se satisface $|M| = 1$. La prueba es más elaborada y usa la idea de escribir el polinomio característico de una matriz antisimétrica como el cuadrado de un polinomio, llamado Pfafiano.

Teorema 3.4 (del valor propio simpléctico)**.** *Si $p(\lambda)$ es el polinomio característico de una matriz simpléctica M, entonces*

$$p(\lambda) = \lambda^{2n} p\left(\frac{1}{\lambda}\right).$$

Demostración. Como $M^{-1} = -JM^TJ$ se sigue que
$M = -J(M^T)^{-1}J$, luego

$$\begin{aligned}|M - \lambda I| &= |-J(M^T)^{-1}J - \lambda I| = |J|^2| - (M^T)^{-1} + \lambda I|\\ &= |-(M^T)^{-1} + \lambda M^T(M^T)^{-1}|\\ &= |-I + \lambda M^T|||(M^T)^{-1}| = |-I + \lambda M|\\ &= \lambda^{2n}|M - \frac{1}{\lambda}J| = \lambda^{2n} p\left(\frac{1}{\lambda}\right).\end{aligned}$$

□

Si $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ es una transformación de coordenadas y escribimos $f(Q, P) = (q(Q, P), p(Q, P))$, en términos de sus componentes, entonces la matriz Jacobiana puede particionarse como

$$M = \frac{\partial(q, p)}{\partial(Q, P)} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{17}$$

La siguiente proposición caracteriza entonces a una transformación simpléctica, en términos de la partición (17) de su matriz Jacobiana:

Proposición 3.5. *Un cambio de coordenadas (17) es simpléctico si y sólo si*
(a) *A^TC y B^TD son simétricas.*
(b) *$A^TD - C^TB = I_n$.*

En particular para toda j, l

$$\sum_i \left(\frac{\partial p_i}{\partial Q_j} \frac{\partial q_i}{\partial Q_l} - \frac{\partial p_i}{\partial Q_l} \frac{\partial q_i}{\partial Q_j} \right) = 0. \tag{18}$$

$$\sum_i \left(\frac{\partial p_i}{\partial P_j} \frac{\partial q_i}{\partial P_l} - \frac{\partial p_i}{\partial P_l} \frac{\partial q_i}{\partial P_j} \right) = 0. \tag{19}$$

$$\sum_i \left(\frac{\partial p_i}{\partial P_j} \frac{\partial q_i}{\partial Q_l} - \frac{\partial p_i}{\partial Q_l} \frac{\partial q_i}{\partial P_j} \right) = \delta_{jl}. \tag{20}$$

Demostración. La condición $M^T J M = J$ equivale a

$$\begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix} \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

Comparando bloques se tiene

$$\begin{aligned} A^T C - C^T A &= 0 \\ B^T D - D^T B &= 0 \\ A^T D - C^T B &= I_n \\ B^T C - D^T A &= -I_n. \end{aligned}$$

Las primeras dos igualdades dan la afirmación (a); las siguientes son equivalentes a la afirmación (b).

De (17) se tiene

$$\begin{aligned} (A_{ij}) &= \left(\frac{\partial q_i}{\partial Q_j} \right), \quad (C_{ij}) = \left(\frac{\partial p_i}{\partial Q_j} \right) \\ (B_{ij}) &= \left(\frac{\partial q_i}{\partial P_j} \right), \quad (D_{ij}) = \left(\frac{\partial p_i}{\partial P_j} \right), \end{aligned}$$

y de la condición $(A^T C - C^T A)_{lj} = 0$ se sigue

$$0 = \sum_i (A_{il} C_{ij} - C_{il} A_{ij}) = \sum_i \left(\frac{\partial q_i}{\partial Q_l} \frac{\partial p_i}{\partial Q_j} - \frac{\partial p_i}{\partial Q_l} \frac{\partial q_i}{\partial Q_j} \right).$$

Análogamente la condición $B^T D - D^T B = 0$ implica

$$\begin{aligned} 0 &= \sum_i (B_{il} D_{ij} - D_{il} B_{ij}) = \sum_i \left(\frac{\partial q_i}{\partial P_l} \frac{\partial p_i}{\partial P_j} - \frac{\partial p_i}{\partial P_l} \frac{\partial q_i}{\partial P_j} \right) \\ &= \sum_i \left(\frac{\partial q_i}{\partial P_l} \frac{\partial p_i}{\partial P_j} - \frac{\partial q_i}{\partial P_j} \frac{\partial p_i}{\partial P_l} \right) \end{aligned}$$

que son las identidades (18) y (19). Las relación (20) se obtiene de la identidad $(D^T A - B^T C)_{lj} = \delta_{lj}$. En efecto,

$$\begin{aligned} \delta_{jl} &= \sum_i (D_{ij} A_{il} - B_{ij} C_{il}) = \sum_i \left(\frac{\partial p_i}{\partial P_j} \frac{\partial q_i}{\partial Q_l} - \frac{\partial q_i}{\partial P_j} \frac{\partial p_i}{\partial Q_l} \right) \\ &= \sum_i \left(\frac{\partial p_i}{\partial P_j} \frac{\partial q_i}{\partial Q_l} - \frac{\partial p_i}{\partial Q_l} \frac{\partial q_i}{\partial P_j} \right). \end{aligned}$$

□

Corolario 3.6. *El conjunto de difeomorfismos simplécticos, Symp($\mathbb{R}^{2n}$) es un subgrupo del grupo de difeomorfismos de $\mathbb{R}^{2n}$, Diff($\mathbb{R}^{2n}$), ambos de dimensión infinita.*

Observación 7. Para simplificar la notación, en lo sucesivo índices repetidos indican suma.

Proposición 3.7. *Consideremos un cambio de coordenadas $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$, $(q, p) = f(Q, P)$ y las formas simplécticas canónicas $\omega = dp \wedge dq$, $\Omega = dP \wedge dQ$. La transformación de coordenadas es simpléctica si y sólo si $f^*(\omega) = \Omega$.*

Demostración.

$$
\begin{aligned}
f^*(dp_i\wedge dq_i) &= \left(\frac{\partial p_i}{\partial Q_j}dQ_j+\frac{\partial p_i}{\partial P_j}dP_j\right)\wedge\left(\frac{\partial q_i}{\partial Q_l}dQ_l+\frac{\partial q_i}{\partial P_l}dP_l\right)\\
&= \frac{\partial p_i}{\partial Q_j}\frac{\partial q_i}{\partial Q_l}dQ_j\wedge dQ_l+\frac{\partial p_i}{\partial P_j}\frac{\partial q_i}{\partial P_l}dP_j\wedge dP_l\\
&\quad+\frac{\partial p_i}{\partial Q_j}\frac{\partial q_i}{\partial P_l}dQ_j\wedge dP_l+\frac{\partial p_i}{\partial P_j}\frac{\partial q_i}{\partial Q_l}dP_j\wedge dQ_l\\
&= \left(\frac{\partial p_i}{\partial Q_j}\frac{\partial q_i}{\partial Q_l}-\frac{\partial p_i}{\partial Q_l}\frac{\partial q_i}{\partial Q_j}\right)dQ_j\wedge dQ_l\\
&+ \left(\frac{\partial p_i}{\partial P_j}\frac{\partial q_i}{\partial P_l}-\frac{\partial p_i}{\partial P_l}\frac{\partial q_i}{\partial P_j}\right)dP_j\wedge dP_l\\
&+ \left(\frac{\partial p_i}{\partial P_j}\frac{\partial q_i}{\partial Q_l}-\frac{\partial p_i}{\partial Q_l}\frac{\partial q_i}{\partial P_j}\right)dP_j\wedge dQ_l
\end{aligned}
$$

Los coeficientes de $dQ_j\wedge dQ_l$, $dP_j\wedge dP_l$ son cero debido a (18) y (19), por (20) el último término es

$$\delta_{jl}dP_j\wedge dQ_l = dP_j\wedge dQ_j=\Omega.$$

□

Consideremos ahora las 1–formas diferenciales en el contradominio $\mathbb{R}^{2n}$ con coordenadas (q,p) y en el dominio $\mathbb{R}^{2n}$ con coordenadas (Q,P)

(21) $$\alpha = pdq=\sum_{i=1}^{n}p_i dq_i,\qquad A=PdQ=\sum_{i=1}^{n}P_i dQ_i.$$

y observe que $\omega=d\alpha$, $\Omega=dA$, es decir las formas simplécticas son exactas.

Proposición 3.8. *Si el cambio de coordenadas $f\colon\mathbb{R}^{2n}\to\mathbb{R}^{2n}$ es simpléctico; α, A son las 1–formas (21), entonces existe una función $W(Q,P)$, posiblemente definida localmente, tal que*

(22) $$f^*(\alpha)-A=dW.$$

Demostración. Como

$$d\,(f^*(\alpha)-A)=d\,f^*(\alpha)-dA=f^*(d\alpha)-\Omega=\Omega-\Omega=0.$$

por Lema de Poincaré toda forma cerrada es localmente exacta, luego existe una función $W(Q,P)$ posiblemente localmente definida , satisfaciendo (22).

□

Observación 8. La relación (22) se acostumbra escribir como

(23) $$p_i dq_i - P_i dQ_i = dW.$$

"donde p_i y la diferencial dq_i se expresan como funciones de Q y P a través del cambio de coordenadas"(véase por ejemplo [7], p. 67) y W es función de Q. Alternativamente podemos considerar que el lado izquierdo de (23) es una 1–forma $\widetilde{\alpha}$ en el espacio producto $\mathbb{R}^{2n}\times\mathbb{R}^{2n}$ de coordenadas (Q,P,q,p). De hecho $\widetilde{\omega}=dp_i\wedge dq_i-dP_i\wedge dQ_i$ es una forma simpléctica en $\mathbb{R}^{2n}\times\mathbb{R}^{2n}$ exacta pues $\widetilde{\omega}=d\widetilde{\alpha}$. El espacio simpléctico $\left(\mathbb{R}^{2n}\times\mathbb{R}^{2n},\widetilde{\omega}\right)$ contiene a la gráfica de f:

(24) $$\Gamma_f=\{(Q,P,q,p)\mid (q,p)=f(Q,P)\}.$$

y si $i_{\Gamma_f}:\Gamma_f\to\mathbb{R}^{2n}\times\mathbb{R}^{2n}$ denota la inclusión, entonces

$$i^*_{\Gamma_f}(\widetilde{\omega})=i^*_{\Gamma_f}(d\widetilde{\alpha})=d\,(i^*_{\Gamma_f}(\widetilde{\alpha}))=0$$

Por el Lema de Poincaré

(25) $$i^*_{\Gamma_f}(\widetilde{\alpha})=dW$$

para alguna función local W definida sobre Γ_f con coordenadas (Q,P). La última expresión es precisamente (22).

Observación 9. La gráfica de la función f (24) es un ejemplo de *variedad Lagrangiana* en el espacio vectorial simpléctico producto $\mathbb{R}^{2n}\times\mathbb{R}^{2n}$ con la forma simpléctica $\widetilde{\omega}=dp_i\wedge dq_i-dP_i\wedge dQ_i$. Véase también la sección 5.

Lema 3.9. *Las siguientes 1–formas diferenciales en* $\mathbb{R}^{2n} \times \mathbb{R}^{2n}$

$$\begin{array}{rclrcl} F_1 & = & p_k dq_k - P_k dQ_k, & F_2 & = & q_k dp_k + P_k dQ_k, \\ F_3 & = & p_k dq_k + Q_k dP_k, & F_4 & = & q_k dp_k - Q_k dP_k, \end{array}$$

difieren por la diferencial de una función escalar.

Demostración. Se puede verificar que

$$F_1 = -F_2 + d(p_i q_i) = F_3 + d(Q_i P_i) = -F_4 + d(p_i q_i - P_i Q_i).$$

Verifiquemos la primera identidad para ilustrar el proceso conocido como *transformada de Legendre.* Primero notemos que

$$F_1 = p_k dq_k - P_k dQ_k = -(P_k dQ_k - p_k dq_k)$$

Las variables independientes son Q_k y q_k; debemos sustituir q_k por p_k como variables independientes así que sumamos y restamos

$$\begin{array}{rcl} F_1 & = & -(P_k dQ_k - p_k dq_k \pm q_k dp_k) \\ & = & -(P_k dQ_k + q_k dp_k - (p_k dq_k + q_k dp_k)) \\ & = & -(P_k dQ_k + q_k dp_k - d(p_k q_k)) \\ & = & -(P_k dQ_k + q_k dp_k) + d(p_k q_k) \\ & = & -F_2 + d(p_k q_k). \end{array}$$

□

Corolario 3.10. *Si el cambio de coordenadas* $f\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ *es simpléctico entonces existen funciones* $W_i(Q,P)$, $i = 1,2,3,4$ *tales que*

$$\begin{array}{rclrcl} f^*(p_k dq_k) - P_k dQ_k & = & dW_1, & f^*(q_k dp_k) + P_k dQ_k & = & dW_2, \\ f^*(p_k dq_k) + Q_k dP_k & = & dW_3, & f^*(q_k dp_k) - Q_k dP_k & = & dW_4 \end{array}$$

Teorema 3.11. *Si el cambio de coordenadas* $f\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ *es simpléctico entonces:*

1. Si $\left|\frac{\partial f_1}{\partial P}\right| \neq 0$, *entonces existe una función* $S(q,Q)$ *tal que*

$$\frac{\partial S}{\partial q_k} = p_k, \quad \frac{\partial S}{\partial Q_k} = -P_k. \tag{26}$$

2. Si $\left|\frac{\partial f_2}{\partial P}\right| \neq 0$, *entonces una función* $S(p,Q)$ *tal que*

$$\frac{\partial S}{\partial p_k} = q_k, \quad \frac{\partial S}{\partial Q_k} = P_k. \tag{27}$$

3. Si $\left|\frac{\partial f_1}{\partial Q}\right| \neq 0$, *entonces existe una función* $S(q,P)$ *tal que*

$$\frac{\partial S}{\partial q_k} = p_k, \quad \frac{\partial S}{\partial P_k} = Q_k. \tag{28}$$

4. Si $\left|\frac{\partial f_2}{\partial Q}\right| \neq 0$, *entonces existe una función* $S(p,P)$ *tal que*

$$\frac{\partial S}{\partial p_k} = q_k, \quad \frac{\partial S}{\partial P_k} = -Q_k. \tag{29}$$

Demostración. La relación

$$p_k dq_k - P_k dQ_k = dW_1$$

es una relación entre formas diferenciales sobre la variedad Γ_f de dimensión $2n$, la cual naturalmente está parametrizada por (Q,P), así que W_1 es función de (Q,P). Si el cambio de coordenadas es $q = f_1(Q,P)$, $p = f_2(Q,P)$ y se satisface la hipótesis del ítem 1, entonces es posible despejar $P = g_1(q,Q)$ y se puede parametrizar Γ_f con las variables (q,Q) como sigue

$$\Gamma_f = \{(Q,P,q,p) \mid P = g_1(q,Q), p = f_2(Q, g_1(q,Q))\}.$$

Si hacemos $S(q,Q) = W_1(Q, g_1(q,Q))$ entonces

$$p_k dq_k - P_k dQ_k = dS(q,Q),$$

con (q, Q) variables independientes, implica

$$p_k = \frac{\partial S}{\partial q_k} \tag{30}$$

$$-P_k = \frac{\partial S}{\partial Q_k}. \tag{31}$$

El resto de los casos se prueba de manera similar.

□

El teorema anterior tiene una versión recíproca: Supongamos que está dada una función S de $2n$ variables con las correspondientes ecuaciones de transformación (26), (27), (28) y (29), queremos dar condiciones suficientes para que dichas ecuaciones definan una transformación canónica; al menos localmente.

Teorema 3.12. *1. Sea $S(q, Q)$ una función escalar de clase C^2. Las ecuaciones implícitas*

$$p_k = \frac{\partial S}{\partial q_k}(q, Q) \tag{32}$$

$$-P_k = \frac{\partial S}{\partial Q_k}(q, Q), \tag{33}$$

bajo la condición de no degeneración

$$\left| \frac{\partial^2 S}{\partial Q_k \partial q_j} \right| \neq 0 \tag{34}$$

definen una transformación de coordenadas

$$q = f_2(Q, P), \quad p = \frac{\partial S}{\partial q}(f_2(Q, P), Q)$$

que es simpléctico.

2. Sea $S(p, Q)$ una función escalar de clase C^2. Las ecuaciones implícitas

$$q_k = \frac{\partial S}{\partial p_k}(p, Q) \tag{35}$$

$$P_k = \frac{\partial S}{\partial Q_k}(p, Q). \tag{36}$$

bajo la condición de no degeneración

$$\left| \frac{\partial^2 S}{\partial Q_k \partial p_j} \right| \neq 0 \tag{37}$$

definen una transformación de coordenadas

$$p = f_2(Q, P), \quad q = \frac{\partial S}{\partial p}(f_2(Q, P), Q)$$

que es simpléctica.

3. Sea $S(q, P)$ una función escalar de clase C^2. Las ecuaciones implícitas

$$p_k = \frac{\partial S}{\partial q_k}(q, P) \tag{38}$$

$$Q_k = \frac{\partial S}{\partial P_k}(q, P), \tag{39}$$

bajo la condición de no degeneración

$$\left| \frac{\partial^2 S}{\partial P_k \partial q_j} \right| \neq 0 \tag{40}$$

definen una transformación de coordenadas

$$q = f_2(Q, P), \quad p = \frac{\partial S}{\partial q}(f_2(Q, P), P)$$

que es simpléctico.

4. Sea $S(p,P)$ una función escalar de clase C^2. Las ecuaciones implícitas

$$q_k = \frac{\partial S}{\partial p_k}(p,P) \tag{41}$$

$$-Q_k = \frac{\partial S}{\partial P_k}(p,P), \tag{42}$$

bajo la condición de no degeneración

$$\left|\frac{\partial^2 S}{\partial P_k \partial p_j}\right| \neq 0 \tag{43}$$

definen una transformación de coordenadas

$$p = f_2(Q,P), \quad q = \frac{\partial S}{\partial p}(f_2(Q,P),P)$$

que es simpléctica.

4. Las ecuaciones de Hamilton

Si $H\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ es una función de clase C^2, las ecuaciones de Hamilton están definidas por el sistema dinámico

$$\dot{q} = \frac{\partial H}{\partial q} \tag{44}$$

$$\dot{p} = -\frac{\partial H}{\partial p}, \quad q,p \in \mathbb{R}^n \tag{45}$$

donde $\frac{\partial H}{\partial q} = (\frac{\partial H}{\partial q_1}, \ldots, \frac{\partial H}{\partial q_n})^T$, etc. que se pueden escribir como

$$\dot{z} = J\nabla H(z), \quad z = (q,p)^T. \tag{46}$$

Proposición 4.1. *Sea $f\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ un cambio simpléctico de coordenadas de clase C^1, $z = f(Z)$. Entonces las ecuaciones (46) se transforman en las ecuaciones de Hamilton*

$$\dot{Z} = J\nabla\widetilde{H}(Z), \quad Z = (Q,P)^T, \tag{47}$$

con Hamiltoniano $\widetilde{H}(Z) = H(f(Z))$.

Probaremos primero el siguiente lema

Lema 4.2. $\nabla\widetilde{H}(Z) = Df(Z)^T \nabla H(f(Z))$.

Demostración. Por definición de gradiente, para cualquier $W \in \mathbb{R}^{2n}$ se cumple,

$$\begin{aligned}\langle\nabla\widetilde{H}(Z), W\rangle &= D\widetilde{H}(Z)\cdot W = DH(f(Z))\cdot Df(Z)\cdot W \\ &= \langle\nabla H(f(Z)), Df(Z)\cdot W\rangle = \langle\nabla Df(Z)^T \cdot H(f(Z)), W\rangle,\end{aligned}$$

de donde se sigue la afirmación. □

Veamos ahora la demostración de la Proposición.

Demostración. Por la regla de la cadena y con $z = f(Z)$, de (46) se sigue

$$\begin{aligned} Df(Z)\cdot\dot{Z} &= J\nabla H(f(Z)) \\ \dot{Z} &= Df(Z)^{-1}J\nabla H(f(Z)) \\ &= -JDf(Z)^T J^2 \nabla H(f(Z)) \\ &= JDf(Z)^T \nabla H(f(Z)) \\ &= J\nabla\widetilde{H}(Z). \end{aligned}$$

□

5. Espacios vectoriales simplécticos

En esta sección veremos la geometría que subyace en la formulación de los sistemas Hamiltonianos, tal como se ha definido en la sección anterior. Posteriormente veremos que es posible globalizar estos conceptos al considerar una variedad simpléctica. Recuerde que una aplicación multilineal es una aplicación $\omega\colon \overbrace{V \times V \times \cdots \times V}^{k-veces} \to \mathbb{R}$ que es lineal en cada argumento por separado, manteniendo el resto constantes. Se reserva el término *forma* cuando la aplicación multilineal es simétrica o antisimétrica, vgr., cuando

$$\begin{aligned} \omega(v_{\sigma_1}, v_{\sigma_2}, \cdots, v_{\sigma_k}) &= \omega(v_1, v_2, \cdots, v_k) \ \text{(simétrica)} \\ \omega(v_{\sigma_1}, v_{\sigma_2}, \cdots, v_{\sigma_k}) &= \operatorname{sgn}(\sigma)\omega(v_1, v_2, \cdots, v_k) \ \text{(antisimétrica)} \end{aligned}$$

para cualquier permutación $(\sigma_1, \sigma_2, \ldots, \sigma_k)$ y $\operatorname{sgn}(\sigma)$ el signo de la permutación.

Definición 10. Un *espacio vectorial simpléctico* es un espacio vectorial sobre $\mathbb{R}$ equipado con una forma bilineal $\omega\colon V \times V \to \mathbb{R}$, antisimétrica y no degenerada. Esto último significa:

$$\omega(u, v) = 0, \quad \forall v \in V \Rightarrow u = 0.$$

Observación 10. La condición de no degeneración obliga a que el espacio vectorial, en caso de ser de dimensión finita, tenga dimensión par. Esto se sigue del hecho de que el determinante de una matriz antisimétrica de dimensión impar es cero. En lo sucesivo denotamos por $d = \dim(V)$, la dimensión del espacio vectorial simpléctico. La forma simpléctica define un isomorfismo de espacios vectoriales $V \to V^*$, mediante la fórmula $u \mapsto u^*$, $u^*(v) = \omega(u, v)$.

El Teorema de Darboux afirma que es posible escoger una base en la cual la matriz asociada a la forma simpléctica toma la forma canónica (13).

Teorema 5.1. *Si V es un espacio vectorial simpléctico y ω la forma simpléctica, entonces existe una base de V, e_i, f_i, $i = 1, 2, \ldots, d$ tal que*

$$\omega(e_i, e_j) = \omega(f_i, f_j) = 0, \quad \omega(e_i, f_j) = \delta_{i,j}.$$

es decir la matriz asociada a ω en esta base es precisamente la matriz simpléctica (13).

Definición 11. Si W es un subespacio del espacio vectorial simpléctico (V, ω) denotamos por $i_W^*\omega$ *la restricción de* ω *a* $W \times W$, es decir $i_W^*\omega = \omega|_{W\times W}$. Decimos que W es un subespacio:

a) *Simpléctico*, si $i_W^*\omega$ es no degenerada.
b) *Isotrópico*, si $i_W^*\omega = 0$.
c) *Lagrangiano*, si es isotrópico y $\dim(W) = \frac{1}{2}\dim(V)$.

5.1. Variedades simplécticas.

La versión global o “no lineal” de los espacios vectoriales simplécticos son las variedades simplécticas. Recuerde que una variedad diferenciable M de dimensión n es un espacio topológico Hausdorff dotado de un atlas de cartas coordenadas $(V_i, \phi_i)_{i\in I}$, donde V_i es un abierto en M, tales que $M = \cup_{i\in I} V_i$ y $\phi_i\colon V_i \to \mathbb{R}^n$ es una función continua con la propiedad de que si $V_i \cap V_j \neq \emptyset$ entonces el cambio de coordenadas

$$\phi_j \circ \phi_i^{-1}\colon \phi_i(V_i \cap V_j) \to \phi_j(V_i \cap V_j)$$

es una función diferenciable (C^∞) entre abiertos de $\mathbb{R}^n$. En lo sucesivo denotamos por $\phi(x) = (q_1(x), \ldots, q_n(x))$, las funciones coordenadas. En una variedad tiene sentido hablar de funciones (reales) diferenciables y curvas diferenciables. Decimos que una función $f\colon M \to \mathbb{R}^n$ es diferenciable, si para cualquier carta coordenada (U, ϕ), $f \circ \phi^{-1}\colon \phi(U) \to \mathbb{R}^n$ es diferenciable como función definida en el abierto $\phi(U) \subset \mathbb{R}^n$. Sean c_1, c_2 gérmenes de curvas alrededor del origen que pasan por el mismo punto $x \in M$. Decimos que son equivalentes, si para cualquier carta coordenada (U, ϕ) tal que $x \in U$ y $\phi(x) = p$, se satisface $\frac{d}{dt}(\phi \circ c_1)(0) = \frac{d}{dt}(\phi \circ c_2)(0)$. Una clase de equivalencia de curvas se llama un vector tangente en x y la colección de vectores tangentes forma el espacio tangente T_xM con la siguiente estructura natural de espacio vectorial: Si

$v_i = \frac{d}{dt}(\phi \circ c_i)(0)$, con c_i curvas como antes, entonces $c_3 \equiv \phi^{-1}(p + t(v_1 + v_2))$ es un germen de curva que define la suma de clases: $[c_3] = [c_1] + [c_2]$. Análogamente la diferencial de $f\colon M \to \mathbb{R}$ es una función lineal en cada fibra: $df_x\colon T_xM \to \mathbb{R}$, donde $df_x \cdot v = w$ se define como sigue: para una curva c representante del vector tangente $v \in T_xM$ entonces $df_x \cdot v = \frac{d}{dt}(f \circ c)(0)$. Sea (U, ϕ) una carta coordenada como antes; entonces $\phi^{-1}(p + te_i)$ es una curva cuya clase se denota por $\frac{\partial}{\partial q_i}(x)$.

Fibra a fibra, los espacios tangente y cotangente (el dual) pueden agruparse en el haz tangente y cotangente: $TM = \bigcup_{x \in M} T_xM$, $T^*M = \bigcup_{x \in M} T_xM^*$. Sean $\pi\colon TM \to M$, $\lambda\colon T^*M \to M$ las proyecciones. Tanto TM como T^*M son variedades diferenciables. Por ejemplo, dada una carta (U, ϕ) en M, $\pi^{-1}(U)$ es un abierto en TM y para $\xi \in \pi^{-1}(U)$ existen funciones $\dot{q}_i\colon \pi^{-1}(U) \to \mathbb{R}$ tales que $\xi = \sum_i \dot{q}_i(\xi) \frac{\partial}{\partial q_i}(x)$. Entonces $T\phi\colon \pi^{-1}(U) \to \mathbb{R}^n \times \mathbb{R}^n$, $T\phi = (q \circ \pi, \dot{q})$ es una carta coordenada. La base dual de $\frac{\partial}{\partial x_i}$ se denota por dq_i y permite definir una carta coordenada en $\lambda^{-1}(U) \subset T^*M$, asociando a $\alpha \in T_x^*M$ las coordenadas $\alpha = \sum p_i(\alpha) dq_i$, de modo que $T^*\phi = (q \circ \lambda, p)$ es una carta coordenada. Un campo vectorial es una sección de TM es decir $X\colon M \to TM$ y $\pi \circ X = id_M$. Una forma diferencial es una sección de T^*M.

Definición 12. Una *variedad simpléctica* es una variedad P dotada de una 2–forma $\omega \in \Lambda^2(P)$ es cerrada y no degenerada.

En otras palabras $d\omega = 0$ y para cada $x \in P$, $\omega_x(u, v) = 0$ para toda $v \in T_xP$ implica $u = 0 \in T_xP$.

El ejemplo canónico de variedad simpléctica es el haz cotangente a una variedad T^*M. La forma simpléctica canónica es la derivada exerior de la 1–forma canónica definida como sigue: La aplicación tangente de la proyección $\lambda\colon T^*M \to M$ es $T\lambda\colon T(T^*M) \to TM$, luego si $\xi \in T_\alpha(T^*M)$, con $\alpha \in T^*M$, sea entonces $x = \lambda(\alpha)$, luego $T\lambda \cdot \xi \in T_xM$ y podemos aplicar$\langle \alpha, T\lambda \cdot \xi \rangle$, donde $\langle , \rangle$ denota el apareamiento natural $T_x^*M \times T_xM \to \mathbb{R}$. Como la aplicación $\xi \mapsto \langle \alpha, T\lambda \cdot \xi \rangle$ depende linealmente de $\xi \in T_\alpha(T^*M)$ se ha definido una forma diferencial Γ en T^*M. La forma simpléctica es $\omega = d\Gamma$.

Observación 11. Tomando coordenadas $(q \circ \lambda, p)$ no es difícil ver que ω y Γ coinciden con la 2–forma y 1–forma canónicas en $\mathbb{R}^{2n}$ del Ejemplo 1, (10) y (9).

El Teorema de Darboux se extiende al caso de variedades

Teorema 5.2. *Si (P, ω) es una variedad simpléctica, para cada $x \in P$ existe una carta coordenada (U, ϕ) con coordenadas $\phi(x) = (q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n)$ tal que*

$$\omega|_U = \sum_{i=1}^{n} dp_i \wedge dq_i. \tag{48}$$

Definición 13. Una variedad L de una variedad simpléctica (P, ω) se dice *Lagrangiana*, si $\dim(L) = \dim(P)/2$ e $i_L^* \omega = 0$. En otras palabras $T_xL \subset T_xP$ es un subespacio Lagrangiano para cada x.

Como ω es cerrada entonces es exacta, al menos localmente digamos $\omega = d\alpha$. Si L es Lagrangiana, entonces $i_L^* \alpha$ es una forma cerrada, luego existe una función posiblemente definida sólo localmente, tal que $i_L^* \alpha = dF$. Llamamos a F una función generadora de la variedad Lagrangiana.

Usando el Teorema de Darboux, se puede obtener una forma canónica local de una variedad Lagrangiana, ya que $p_i \, dq_i = dF$, implica

$$L = \left\{ (q, p) \mid p_i = \frac{\partial F}{\partial q_i}(q, p) \right\}.$$

En la primera parte de este trabajo se han visto múltiples ejemplos de funciones generadoras de variedades Lagrangianas dadas por gráficas de transformaciones canónicas en la variedad simpléctica producto $\mathbb{R}^{2n} \times \mathbb{R}^{2n}$.

6. Sistemas Hamiltonianos

Definición 14. Sea (P, ω) una variedad simpléctica. Un campo vectorial $X \colon P \to TP$ se dice *Hamiltoniano*, si existe una función $H \in C^\infty(P)$ tal que

$$i_X \omega = -dH. \tag{49}$$

Recuerde que si X es un campo vectorial y α una k–forma entonces $i_X \alpha$ es una $k-1$ forma definida como

$$(i_X \alpha)_x(u_1, u_2, \ldots, u_k) = \alpha_x(X(x), u_1, u_2, \ldots, u_k).$$

Así, $i_X \omega$ es una 1-forma globalmente exacta, si el campo es Hamiltoniano.

Observación 12. En una carta de Darboux, si

$$X = A_i(q,p)\frac{\partial}{\partial q_i} + B_i(q,p)\frac{\partial}{\partial p_i}$$

entonces

$$i_X \omega = -A_i(q,p)dp_i + B_i(q,p)dq_i = -\frac{\partial H}{\partial q_i}dq_i - \frac{\partial H}{\partial p_i}dp_i$$

lo cual implica $A_i = \frac{\partial H}{\partial p_i}$, $B_i = -\frac{\partial H}{\partial q_i}$. Luego, a lo largo de una curva solución del campo se satisfacen las ecuaciones de Hamilton

$$\frac{dq_i}{dt} = A_i(q(t), p(t)) = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = A_i(q(t), p(t)) = -\frac{\partial H}{\partial q_i}.$$

7. Principios variacionales

Consideremos el siguiente problema variacional: sea $L(t, x, v)$ una función definida para $t \in [t_1, t_2]$, $(x, v) \in D \times \mathbb{R}^d$ y $D \subseteq \mathbb{R}^d$ un abierto, L de clase C^r, con $r \geq 2$. Definamos el funcional definido para una curva $x \colon [t_1, t_2] \to D$ por

$$I(x) = \int_{t_1}^{t_2} L(s, x(s), \dot{x}(s))\, ds, \quad \frac{dx}{dt} = \dot{x}(t)$$

Analicemos primero la existencia de minimales en el espacio de curvas de clase C^1 con extremos fijos

$$\Gamma = \{x \in C^1[t_1, t_2] \mid x(t_1) = x_1, \quad x(t_2) = x_2\}.$$

Definición 15. Decimos que $x^* \in \Gamma$ es *minimal en* Γ, si

$$I(x^*) \leq I(x), \quad \forall x \in \Gamma.$$

Observación 13. El ínfimo puede existir, sin que se alcance.

Considere por ejemplo $d = 1$, $L(t, x, v) = t^2 v^2$, $[t_1, t_2] = [0, 1]$ $x_1 = 0$, $x_2 = 1$. Para la familia de curvas $x_m(t) = t^m$ se tiene

$$I(x_m) = \frac{1}{m+3} \quad \text{por lo que} \quad \inf_{m \in \mathbb{N}} I(x_m) = 0,$$

pero $I(x) > 0$ para toda $x \in \Gamma$.

Sin embargo se tiene el siguiente resultado:

Proposición 7.1. *Si x es minimal en Γ entonces $x \in C^2[t_1, t_2]$ y*

$$\frac{\partial L}{\partial v_j}(t, x, \dot{x}) = \int_{t_1}^{t_2} \frac{\partial L}{\partial x_j}(s, x(s), \dot{x}(s))\, ds = const. \tag{50}$$

para toda $t \in [t_1, t_2]$.

Las ecuaciones (50) se llaman las *ecuaciones integrales de Euler–Lagrange.*

Definición 16. Una curva minimal $x^* \in \Gamma$ se dice *regular*, si $\det\left(L_{v_i, v_j}\right) \neq 0$ a lo largo de x^*, $v = \dot{x}^*$.

El siguiente resultado de regularidad aplica a curvas minimales.

Proposición 7.2. *Si $x^* \in \Gamma$ es una curva minimal regular, entonces $x^* \in C^2[t_1, t_2]$ y se satisface*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v_j}\right) = \frac{\partial L}{\partial x_j}, \quad v = \frac{dx^*}{dt}. \tag{51}$$

Definición 17. Una curva x^* que satisfaga las ecuaciones de Euler–Lagrange (51) se dice *extremal* en Γ.

Observación 14. La teoría detrás de los métodos variacionales es muy extensa y referimos al lector a la referencia [6] donde se caracterizan las propiedades topológicas de las curvas minimales y en particular los conjuntos de Mather.

8. La transformada de Legendre

En secciones anteriores hemos introducido la tansformada de Legendre en el contexto de pasar de una función generadora de un tipo a otro bajo hipótesis locales de convexidad. Aquí supondremos hipótesis suficientes para que la transformada de Legendre pueda definirse globalmente, lo cual nos permitirá relacionar la formulación Lagrangiana y Hamiltoniana de la Mecánica.

Definición 18. Un Lagrangiano $L(q, v)$ se dice de *Tonelli*, si la aplicación $L(x, \cdot)$ es de crecimiento superlineal y estrictamente convexa, es decir

$$\lim_{|v|\to\infty} \frac{L(x, v)}{|v|} = +\infty \quad \text{y} \quad \frac{\partial^2 L}{\partial v^2}(x, v)$$

es estrictamente definida positiva. Análogamente un Hamiltoniano $H(x, p)$ se dice de *Tonelli*, si la aplicación $H(x, \cdot)$ es de crecimiento superlineal y estrictamente convexa, es decir

$$\lim_{|p|\to\infty} \frac{H(x, p)}{|p|} = +\infty \quad \text{y} \quad \frac{\partial^2 H}{\partial p^2}(x, p)$$

es estrictamente definida positiva.

Observación 15. Si $L(x, v)$ es un Lagrangiano de Tonelli de clase C^2, entonces la transformación

$$p = \frac{\partial L}{\partial v}(x, v)$$

es invertible globalmente de manera continuamente diferenciable respecto de v y continuamente respecto de x. La demostración del siguiente resultado se deja al lector.

Proposición 8.1. *Sea $L(x, v)$ un Lagrangiano de Tonelli. Defina el Hamiltoniano asociado*

$$H(x, p) = v \cdot p - L(x, v) \tag{52}$$

donde, $v = v(x, p)$ de acuerdo a la última observación, entonces $H(x, p)$ es un Hamiltoniano de Tonelli. Viceversa, si $H(x, p)$ es un Hamiltoniano de Tonelli y $L(x, v) = v \cdot p - H(x, p)$ donde $p = p(x, v)$ de acuerdo a la útima observación, entonces $L(x, v)$ es de Tonelli.

Teorema 8.2. *Sea $L(x, v)$ un Lagrangiano de Tonelli de clase C^2 en $x \in D$ abierto de $\mathbb{R}^d$, $v \in \mathbb{R}^d$. Entonces la aplicación*
$\mathcal{L}\colon D \times \mathbb{R}^d \to D \times \mathbb{R}^d$,

$$(x, v) \mapsto \left(x, p = \frac{\partial L}{\partial v}(x, v)\right) \tag{53}$$

es un difeomorfismo global, con inversa

$$(x, p) \mapsto \left(x, v = \frac{\partial H}{\partial p}(x, p)\right). \tag{54}$$

Demostración. Sea $v(q, p)$ la función implícitamente definida por la primera relación en (53). Derivando H respecto de p,

$$\frac{\partial H}{\partial p} = v + p \cdot \frac{\partial v}{\partial p} - \frac{\partial L}{\partial v} \cdot \frac{\partial v}{\partial p} = v + \left(p - \frac{\partial L}{\partial v}\right) \cdot \frac{\partial v}{\partial p} = v$$

de la definición de p. ☐

Teorema 8.3. *Bajo las mismas hipótesis que el Teorema anterior. La transformada de Legendre (53) manda las ecuaciones de Euler-Lagrange en las ecuaciones de Hamilton y viceversa:*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial x} = 0 \Leftrightarrow \frac{dx}{dt} = \frac{\partial H}{\partial p}, \; \frac{dp}{dt} = -\frac{\partial H}{\partial x}.$$

Demostración. Nótese que de la definición (52) se satisface

$$\frac{\partial H}{\partial x} = -\frac{\partial L}{\partial x}.$$

Ahora, partiendo de las ecuaciones de Lagrange, de la definición de p (ver 53) se tiene

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial x} = 0 \Rightarrow \frac{dp}{dt} = -\frac{\partial H}{\partial x},$$

que es la segunda ecuación de Hamilton. Por otro lado,

$$\frac{dx}{dt} = v = \frac{\partial H}{\partial p} + p \cdot \frac{\partial v}{\partial p} - \frac{\partial L}{\partial v} \cdot \frac{\partial v}{\partial p} = \frac{\partial H}{\partial p},$$

donde se ha derivando (52) respecto de p, suponiendo $v = v(x, p)$. El recíproco se muestra de manera similar y se deja como ejercicio. □

9. Termodinámica de sistemas químicos

Las variables medibles de los sistemas macroscópicos cuando son perturbados, evolucionan de manera complicada y pueden variar de punto a punto (variables distribuidas). En muchos casos tienden a estabilizarse con el tiempo en valores determinados independientemente del punto donde son medidas (variables conglomeradas). Decimos que se ha alcanzado el equlibrio termodinámico. Los efectos disipativos pueden atenuarse perturbando poco a poco y ligeramente el sistema, esperando un tiempo suficiente a que los valores de las variables se hayan estabilizado. Un proceso de este tipo se llama quasiestático y se representa por una curva en el espacio de estas variables. El parámetro de esta curva no es el tiempo, sino que simplemente indica el orden en el que el proceso se lleva a cabo. De las variables conglomeradas se distinguen las de tipo extensivo que dependen proporcionalmente del volumen del sistema, vgr., si $r-1$ del total de r son aumentadas por un factor λ, entonces la variable restante varía en la misma proporción; las del tipo intensivo son independientes de esta transformación de escala[1].

Postulado I. *En los sistemas simples existen estados particulares, llamados estados de equilibrio termodinámico que están caracterizados macroscópicamente por la energía interna U, el volumen V y las cantidades molares $N_1, N_2, \ldots, N_r$ de los componentes químicos.*

Las variables a las que se hace referencia en el Postulado I son del tipo extensivo. Más adelante veremos que las variables intensivas están determinadas a partir de la relación funcional entre estas variables y otra más, la entropía S que será introducida más adelante, de modo que se puede suponer una relación funcional $U = U(S, V, N_1, \ldots, N_r)$; veremos que los valores de las variables intensivas están completamente determinados por las derivadas de primer orden de esta relación funcional. Equivalentemente se puede suponer una relación funcional $S = S(U, V, N_1, \ldots, N_r)$ en igual circunstancia. Más aún, veremos que más que la relación funcional de un tipo u otro, el contenido geométrico de la termodinámica está resumido en la suposición (postulado) de que los estados de equilibrio termodinámico están caracterizados por los valores de sus variables extensivas e intensivas, las cuales forman una *subvariedad Lagrangiana.*

La primera ley reconoce el calor como una forma de energía que aunque no es una variable de estado, el cambio en la energía interna se debe al calor suminstrado al sistema menos el trabajo realizado por éste.

(55) $$dU = \delta Q - \delta W,$$

[1]Caratheodory llama a las variables extensivas *variables de forma* y las variables intensivas, *variables de estado.*. Nosotros usaremos el término variables de estado para referirnos a ambas.

donde δQ representa el calor suministrado al sistema, una 1–forma diferencial y δW el trabajo producido por el sistema.

El trabajo $\delta W = -p\,dV + \sum_i \mu_i dN_i$. En general las 1–formas de calor y trabajo δQ, δW no son exactas por lo que el calor o el trabajo total dado por la integral sobre una trayectoria que especifica el proceso que siguen las variables extensivas $\gamma(t)$ en general depende de ésta. Un proceso se dice adiabático, si $\delta Q(\dot\gamma(t)) = 0$.

Segunda ley. *Alrededor de cualquier punto P y cualquier vecindad de éste, existen estados que no pueden ser alcanzados por un proceso adiabático. En particular, existen funciones T, S de las coordenadas extensivas $V, N_1, \ldots, N_r$ tales que $\delta Q = T\,dS$.*

Caratheodory muestra que la condición de la existencia de puntos inalcanzables por procesos adiabáticos es equivalente a pedir que la distribución definida por $\delta Q = 0$ (ecuación de Pfaff) sea integrable en el sentido de Frobenius, de lo cual se sigue la existencia de un factor integrante $1/T$ que hace que δQ sea exacta:

$$dS = \frac{dQ}{T}.$$

Para una sola sustancia química, de las tres variables V, U, N sólo dos son independientes, debido a la propiedad de homogeneidad, luego δQ es una 2–forma en un espacio de dimensión 2 y por lo tanto es integrable. Sin embargo, para más de una sustancia, la integrabilidad no es automática, por ello la segunda ley no se trata de un enunciado matemático, sino de una ley física,

La primera (55) y segunda leyes pueden enunciarse en el solo apartado

$$T\,dS = dU - p\,dV + \sum_i \mu_i\,dN_i. \tag{56}$$

La definición precisa de estados de equilibrio termodinámico es la siguiente: Consideremos el espacio vectorial simpléctico de pares de variables extensivas–intensivas y dentro de éste, el cono abierto

$$\mathcal{U}\colon (S, V, N_1, \ldots, N_c, T, -p, \mu_1, \ldots, \mu_c), \quad S, V, N_i > 0$$

con la forma simpléctica $\omega = d\alpha$, con

$$\alpha = TdS + p\,dV - \sum_i \mu_i\,dN_i \tag{57}$$

Los estados de equilibrio termodinámico forman una variedad Lagrangiana L: $0 = i_L^*\omega = d(i_L^*\alpha)$, lo cual implica que $i_L^*\alpha = dU$ para alguna función escalar U. Cuando L se puede parametrizar por $(S, V, N_1, \ldots, N_c)$ y entonces L es la gráfica de dU es decir $\alpha = dU$ con $U = U(S, V, N_1, \ldots, N_c)$; explícitamente

$$T = \frac{\partial U}{\partial S}, \quad p = -\frac{\partial U}{\partial V}, \quad \mu_i = \frac{\partial U}{\partial N_i}, \quad i = 1, 2, \ldots, c.$$

Así, la variedad Lagrangiana L se identifica con los estados de equilibrio termodinámico y sólo para ellos las variables extensivas toman valores bien definidos.

Postulado II. *Los estados de equilibrio de un sistema termodinámico forman una subvariedad Lagrangiana.*

Para una sóla especie se tiene que $U = U(S, V, N)$ y la propiedad extensiva de la energía interna implica que U es una función homogénea de grado 1. Comúnmente se fija la cantidad molar N o el volumen V y se habla de la entropía o energía interna molares $U = Nu$, $S = Ns$, $V = Nv$ o específicas $U = Vu$, $S = Vs$, $N = Vn$ En este caso el potencial químico es $\mu = \frac{\partial U}{\partial N} = 0$ y (57) se reduce a $\alpha = Tds + pdv$. En este caso la condición de integrabilidad para α es satisfecha automáticamente, ya que $d\alpha$ es una 2-forma en un abierto de $\mathbb{R}^2(s, v)$.

Para dos sustancias químicas *que no reaccionan* hay un potencial químico por cada fase de la sustancia $\mu_{1,2}^a$, $\mu_{1,2}^b$, ..., $\mu_{1,2}^\nu$. El equilibrio termodinámico de las ν fases se describe por las igualdades $\mu_1^j = \mu_2^j$ para $j = 1, 2, \ldots, \nu$; así las condiciones de equilibrio dan ν restricciones entre las $c+2$ variables, dando un total de $r = c + 2 - \nu$ grados de libertad.

9.1. Potenciales termodinámicos. La relación de estado $U = U(S, V, N_1, N_2, \ldots, N_c)$ no es sino una parametrización de la variedad Lagrangiana L, de ahí que en muchos casos sea conveniente usar otras coordenadas u otros potenciales, en vez de U, para describir mejor determinados procesos o propiedades.

Un cambio de coordenadas para el potencial U se lleva a cabo mediante un transformación de Legendre parcial. Por ejemplo, para una sustancia habiendo fijado el número de moles, $U = U(S, V)$. Si queremos expresar a U como función de T y V entonces hacemos $G = U - TS$, donde se supone que

$$T = \frac{\partial U}{\partial S}(S, V)$$

permite expresar a $S = S(T, V)$, luego $G = G(T, V)$, ahora

$$dG = dU - TdS - SdT = TdS - pdV - TdS - SdT = -pdV - SdT,$$

de donde se obtiene

$$-p = \frac{\partial G}{\partial V}, \quad -S = \frac{\partial G}{\partial T}.$$

El potencial termodinámico G se llama *el potencial de Helmholtz.* Vemos que el cambio de coordenadas se lleva a cabo agregando un término que es el producto de la variable extensiva que se quiere sustituir por su correspondiente variable extensiva (S–T en el caso de la función de Helmholtz). Para el estudio de reacciones químicas que generalmente se llevan a cabo a presión constante es más conveniente usar p y V como variables independientes; se define entonces la entalpía como $H = U + pV$, de donde

$$dH = dU + pdV + Vdp = TdS - pdV + pdV + Vdp = TdS + Vdp,$$

de donde la variedad Lagrangiana se describe por

$$T = \frac{\partial H}{\partial S}, \qquad V = \frac{\partial H}{\partial p}.$$

La energía libre de Gibbs es el potencial $G = U + PV - TS$ en el que se sustituyen (S, V) por variables intensivas (P, T). Se tiene $dG = -SdT + Vdp$.

Alternativamente, se puede expresar $U = U(T, V)$ al despejar $S = S(T, V)$ de la relación $T = \frac{\partial U}{\partial S}(S, V)$, luego $U(T, V) = U(S(T, V), V)$. Esta parametrización se usará explícitamente en el caso del gas ideal.

9.2. Relaciones de Maxwell. Son consecuencia de la exactitud de la forma $i_L^*\alpha$. Por ejemplo, tomando T, p como variables independientes para el potencial de Gibbs, $dG = -SdT + Vdp$ es una diferencial exacta, luego

$$-\frac{\partial S}{\partial p} = \frac{\partial V}{\partial T}. \tag{58}$$

Otras relaciones entre derivadas parciales pueden derivarse de manera parecida, por ejemplo para los potenciales químicos

$$\frac{\partial \mu_i}{\partial N_j} = \frac{\partial \mu_j}{\partial N_i}. \tag{59}$$

9.3. Ejemplo el gas ideal. Para un sistema PVT de una sola sustancia la capacidad calorífica se define como el número de calorías necesarias para elevar en un grado su temperatura. Es una cantidad extensiva que depende del proceso. La capacidad calorífica a volumen constante se define a partir de $U = U(T, V)$ como se mostró antes

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V \tag{60}$$

aunque hemos usado la notación clásica que enfatiza que en el proceso el volumen se mantiene constante, esto es innecesario.

La capacidad calorífica a presión constante se define a partir de $H = H(T, P)$ como se mencionó antes, luego

$$C_p = \frac{\partial H}{\partial T} \tag{61}$$

El uso de la energía interna en el primer caso y de la entalpía en el segundo es consistente con la definición física, ya que $\delta Q = dU$ si $dV = 0$ en el primer caso y $\delta Q = dH$, si $dp = 0$.

Para un gas ideal la ecuación de estado es

$$pV = NRT \tag{62}$$

donde R es la constante de los gases ideales.

Proposición 9.1. *Para un gas ideal la energía interna sólo depende de la energía.*

Demostración. Consideremos $U = U(T, V)$, luego

$$dU = \frac{\partial U}{\partial T} dT + \frac{\partial U}{\partial V} dV. \tag{63}$$

De la seguna ley y usando la ecuación de estado se tiene

$$dS = \frac{dU}{T} + \frac{NR}{V} dV \tag{64}$$

Sustituyendo (63) en (64) se tiene

$$dS = \frac{1}{T}\frac{\partial U}{\partial T} dT + \left[\frac{1}{T}\frac{\partial U}{\partial V} + \frac{NR}{V}\right] dV. \tag{65}$$

Sin embargo usando $S = S(T, V)$ y comparando con $dS = \frac{\partial S}{\partial T} dT + \frac{\partial S}{\partial V} dV$ se tiene de inmediato

$$\frac{\partial S}{\partial T} = \frac{1}{T}\frac{\partial U}{\partial T} \tag{66}$$

$$\frac{\partial S}{\partial V} = \frac{1}{T}\frac{\partial U}{\partial V} + \frac{NR}{V}. \tag{67}$$

Usando la igualdad de las parciales mixtas

$$\frac{1}{T}\left(\frac{\partial^2 U}{\partial V \partial T}\right) = -\frac{1}{T^2}\frac{U}{V} + \frac{1}{T}\frac{\partial^2 U}{\partial T \partial V} \tag{68}$$

cancelando las derivadas mixtas se obtiene finalmente

$$\frac{\partial U}{\partial V} = 0.$$

□

Agradecimientos

El autor agradece encarecidamente a los editores y a los árbitros su cuidadosa revisión de este trabajo.

Referencias

[1] Abbot, M.M. & Van Ness, H.C. *Termodinámica.* Serie de compendios Schaum. McGraw-Hill, 1975.
[2] Arnold V.I. *Mathematical Methods of Classical Mechanics.* Graduate Texts in Mathematics. Springer–Verlag. New York, Heidelberg Berlin, 1978.
[3] Guillemin, V. & Sternberg, Sh. *Symplectic techniques in physics.* Cambridge University Press, 1984.
[4] Kirkwood, John G. and Oppenheim, Irwin. *Chemical Thermodynamics* (pg. 36). McGraw-Hill, 1961.
[5] Maslov, V.P. & Nazaikinskii, V.E. Tunnel canonical operator in thermodynamics. Funct. Anal. and its Appl., 40, no.3, 173–187, 2006.
[6] Moser, J. Selected chapters in the calculus of variations. Notas no publicadas.
[7] Pollard, H. *Celestial Mechanics.* The Carus Mathematical Monographs No. 18, 1966.
[8] Schutz, B. *Geometric methods of mathematical physics.* Cambridge, 1980.

Dirección del autor:
Joaquín Delgado Fernández
Universidad Autónoma Metropolitana,
Unidad Iztapalapa,
División de Ciencias Básicas e Ingeniería,
Departamento de Matemáticas.
Av. San Rafael Atlixco 186, Col. Vicentina
Del. Iztapalapa, C.P. 09340 México, D.F.
e-mail: `jdf@xanum.uam.mx`