

25  
2ej.

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS



Geometría Simpléctica

T E S I S

QUE PARA OBTENER EL TITULO DE  
MATEMATICO

Eduardo Ricardo Uribe Vargas

MEXICO, D. F.

1992

FALLA DE ORIGEN



## **UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso**

### **DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# Índice.

	pag
Introducción	1
Capítulo I	
¿Formalmente, que es la geometría simpléctica?	3
Capítulo II	
Un poco de Óptica Geométrica.	
2.1 Óptica Gaussiana	
2.2 Óptica Lineal	
2.3 Principio de Fermat	
Capítulo III	
Calculo Variacional	44
Capítulo IV	
Descripción del flujo geodésico del toro plano y del toro de revolución	58
Capítulo V	
Ecuaciones de Hamilton	
5.1 Paso a las coordenadas $(q, p)$	74
5.2 Ecuaciones de Hamilton	78
Bibliografía	89

## Introducción

El material de esta tesis está basado en las notas de clase del curso de *Geometría Simpléctica*, impartido por el Dr. Santiago López de Medrano durante el primer semestre de 1990 en la facultad de Ciencias de la UNAM, y también en los siguientes textos:

1. Victor Guillemin; *Symplectic Technis in Physics*; Cambridge University.

2. Notas de Mecánica Celeste (versión preliminar) por Alain Chenciner; Université de paris VII.

3. V.I. Arnold, *Mathematical Methods of Classical Mechanics*; Springer Verlag, N.Y. 1978

En el capítulo I, se define formalmente lo que es la geometría simpléctica y se le compara con la geometría euclidiana. En el capítulo II, se ve algo de óptica gaussiana y óptica lineal, donde ya aparecen transformaciones lineales simplécticas y matrices simplécticas; luego se enuncia el Principio de Fermat y se ven algunos ejemplos de medios ópticos en los que la velocidad de la luz depende de la altura. En el capítulo III, se ve un poco de cálculo de variaciones, se introduce el concepto de "energía de una curva" y se define el concepto de geodésica en términos de la energía. En el capítulo IV, se hace una descripción de la geometría del flujo geodésico del toro plano y del toro de revolución. Finalmente, en el capítulo V se ven las ecuaciones de Hamilton, el flujo hamiltoniano determinado por éstas, y se demuestra que el flujo hamiltoniano preserva el volumen y la forma simpléctica. Dicho en otras palabras, el flujo hamiltoniano  $\varphi(x,t)$  consiste de transformaciones simplécticas  $\varphi(x)$ .

## Capítulo I

### ¿Formalmente, qué es la Geometría Simpléctica?

Formalmente, es la geometría definida en un espacio vectorial por medio de una forma bilineal antisimétrica y no-degenerada.

Para entender esto mejor, comparémosla con la euclidiana. En la geometría euclidiana tenemos un espacio vectorial  $V$  de dimensión finita sobre  $\mathbb{R}$  con producto interior.

Un producto interior es una forma bilineal

$$\alpha: V \times V \rightarrow \mathbb{R}$$

a) simétrica:  $\alpha(x,y) = \alpha(x,y)$

b) positiva definida:  $\alpha(x,x) > 0$  si  $x \neq 0$

$b) \Rightarrow b')$  no degenerada:  $\forall x \in V, x \neq 0 \exists y \in V$  tal que  $\alpha(x,y) \neq 0$

En  $(V, \alpha)$ , espacio euclidiano se tiene la notación usual  $\langle x, y \rangle = \alpha(x, y)$ . A partir de esto se define longitud de vectores:  $|x| = \alpha(x, x)^{1/2}$ ; luego con ayuda del cálculo obtenemos longitud de curvas.

Hay las isometrías:  $T: V \rightarrow V$  lineal es isometría si  $\alpha(T(x), T(y)) = \alpha(x, y)$ .

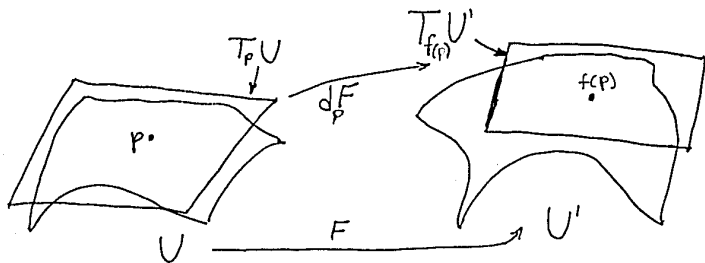
Se hace una extensión que es la geometría riemanniana; por ejemplo, considerando un abierto

$U \subset \mathbb{R}^n$  y dando una forma bilineal  $\alpha_p$  sobre todos los vectores a partir de un punto  $p \in U$ , para cada  $p \in U$  que varíe continua y diferenciablemente (suavemente) de punto a punto. En general, si  $e_1, \dots, e_n$  es base de  $V$  y escribimos  $\alpha(e_i, e_j) = (g_{ij})_p$ , entonces  $(g_{ij})_p$  es una matriz simétrica y es la matriz que representa a la forma bilineal en el punto  $p$ . Así, en geometría riemanniana tenemos que  $(g_{ij}) = (g_{ij})(p)$ , es decir la forma bilineal depende de  $p$ .

Con esto se extiende la noción de longitud.

Más en general, si  $U^n$  es una variedad, y si para todo punto  $p \in U^n$  se tiene  $\langle \cdot, \cdot \rangle_p: T_p U \times T_p U \rightarrow \mathbb{R}$ , donde  $\langle \cdot, \cdot \rangle_p$  varía suavemente de punto a punto, entonces tenemos una variedad riemanniana donde también podemos hablar de longitudes.

Decimos que  $F: U \rightarrow U'$  es una isometría si  $F \in C^\infty$  y  $dF_p$  es isometría  $\forall p \in U$   
 $dF_p: T_p U \rightarrow T_{f(p)} U'$ .



En la geometría simpléctica se tiene  $V$  espacio vectorial de dimensión finita sobre  $\mathbb{R}$  con una forma simpléctica. Una forma simpléctica es una forma bilineal  $\omega: V \times V \rightarrow \mathbb{R}$

a) antisimétrica:  $\omega(x,y) = -\omega(y,x)$

b) no degenerada:  $\forall x \in \mathfrak{D}, x \neq 0 \exists y \in \mathfrak{D}$  tal que  $\omega(y,x) \neq 0$

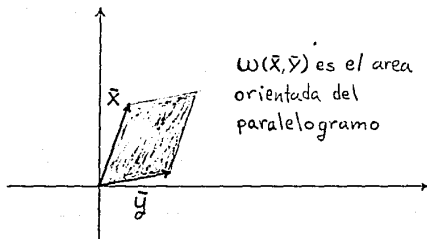
a) implica que  $\omega(x,x) = 0 \forall x \in V$ .

Un espacio vectorial con una forma simpléctica  $(V, \omega)$  es un espacio simpléctico o se dice que tiene estructura simpléctica. A cada pareja de vectores se le asocia un número real que es "algo" como un área orientada del paralelogramo formado por los vectores.

Ejemplo:

$$V = \mathbb{R}^2, \bar{x} = (x_1, x_2), \bar{y} = (y_1, y_2)$$

$$\omega(\bar{x}, \bar{y}) = x_1 y_2 - x_2 y_1 = \det$$



No hay ejemplos de dimensión 1, ya que si  $\dim(V) = 1$  y  $x, y \in \mathfrak{D}$ , entonces  $y = \lambda x$  para algún  $\lambda \in \mathbb{R}$ , lo cual implica que  $\omega(x, y) = \omega(x, \lambda x) = \lambda \omega(x, x) = 0$ , por lo que  $\omega$  es degenerada.

Otro ejemplo:

$$V = \mathbb{R}^4, x = (x_1, x_2, x_3, x_4), y = (y_1, y_2, y_3, y_4).$$

$$\omega(x, y) = x_1 y_2 - x_2 y_1 + x_3 y_4 - x_4 y_3 = \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} + \det \begin{pmatrix} x_3 & x_4 \\ y_3 & y_4 \end{pmatrix}$$

Por ejemplo, para  $x = (2, 2, 27, 2)$  tenemos que  $\exists y = (0, 0, 0, 1)$  tal que  $\omega(x, y) \neq 0$ .

Si  $(V, \omega)$  y  $(V', \omega')$  son espacios simplécticos y  $T: (V, \omega) \rightarrow (V', \omega')$  es lineal, se dice que  $T$  es una transformación simpléctica si  $\omega'(T(x), T(y)) = \omega(x, y)$ .

¿Cuándo  $T: (\mathbb{R}^2, \det) \rightarrow (\mathbb{R}^2, \det)$  es simpléctica?

$T$  es simpléctica  $\Leftrightarrow \det \begin{pmatrix} T(\bar{x}) \\ T(\bar{y}) \end{pmatrix} = \det \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \quad \forall \bar{x}, \bar{y} \in \mathbb{R}^2$

Si  $[T] = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , entonces

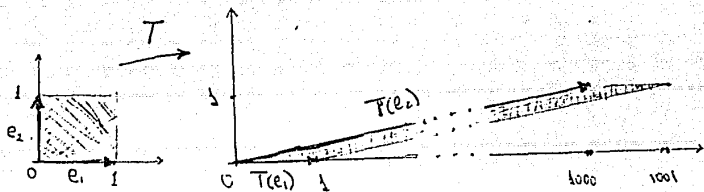
$$\det \begin{pmatrix} T(\bar{x}) \\ T(\bar{y}) \end{pmatrix} = (ad-bc)(x_1 y_2 - x_2 y_1) = (ad-bc) \det \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}$$

por lo tanto, en  $(\mathbb{R}^2, \det)$   $T$  es simpléctica  $\Leftrightarrow \det T = 1$ .

Ejemplo:

Veamos un caso sencillo en  $\mathbb{R}^2$

La transformación  $T$  cuya matriz es  $\begin{pmatrix} 1 & 1000 \\ 0 & 1 \end{pmatrix}$



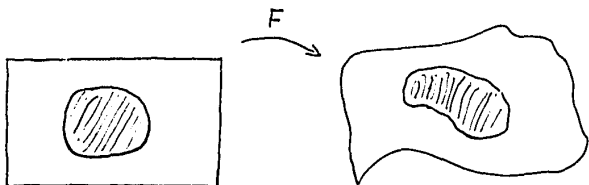
En  $(\mathbb{R}^2, \det)$  puede haber transformaciones no-lineales que sean simplécticas  $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ . Se considera  $dF_p: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $dF_p$  ya es lineal y entonces se pide que  $dF_p$  sea simpléctica para todo  $p \in \mathbb{R}^2$ :  $\det(dF_p) = 1 \forall p$ .

Ejemplo:

$$F(x,y) = (x, y + f(x)) \text{ con } f \in C^1, \text{ entonces}$$

$$dF(x,y) = \begin{pmatrix} 1 & 0 \\ f'(x) & 1 \end{pmatrix} \text{ y } \det(dF_p) = 1 \forall p \in \mathbb{R}^2$$

con  $\mathbb{R}^2$  en forma simpléctica podemos asignar un área orientada a cada superficie.



Se conservan las áreas orientadas y puede ser  $F$  no-lineal

También se hace geometría simpléctica en variedades simplécticas. Una variedad simpléctica es una variedad de dimensión par que, localmente, tiene la estructura de un espacio vectorial simpléctico.

Regresemos a la geometría euclidiana. Consideremos un espacio euclidiano  $(V, \alpha)$  con  $\alpha: V \times V \rightarrow \mathbb{R}$  la forma bilineal (producto interior) del espacio euclidiano con base  $e_1, \dots, e_n$ .

Si  $\alpha_{ij} = \alpha(e_i, e_j)$ , entonces  $M = (\alpha_{ij})$  es la matriz de  $\alpha$  respecto a la base  $e_1, e_2, \dots, e_n$ . Si  $e'_1, e'_2, \dots, e'_n$  es otra base para  $V$ , podemos escribir  $e'_i = \sum a_{ij} e_j$ , entonces  $\alpha'_{ij} = \alpha(e'_i, e'_j)$

$$= \alpha\left(\sum a_{ik} e_k, \sum a_{jl} e_l\right)$$

$$= \sum a_{ik} a_{jl} \alpha(e_k, e_l)$$

$$= \sum a_{ij} \alpha_i \alpha_j = (A^t M A)_{ij}$$

por lo que  $M^t = A^t M A$

Si  $(V, \alpha)$  es un espacio euclidiano  $(\alpha, V \times V \rightarrow \mathbb{R}$  producto interior), entonces se dice que

$e_1, e_2, \dots, e_n$  es una base ortonormal para  $V$  si

$$\alpha(e_i, e_i) = 1 \quad i = 1, \dots, n$$

$$\alpha(e_i, e_j) = 0 \quad \text{para } i \neq j$$

**Teorema:**

Todo espacio euclidiano  $(V, \alpha)$  tiene una base ortonormal.

**Demostración:** (Gramm-Schmidt)

$e_1, e_2, \dots, e_n$  es cualquier base

$$e'_1 = e_1 / |e_1|; \quad e'_2 = (e_2 - \alpha(e'_1, e_2)e'_1) / |(e_2 - \alpha(e'_1, e_2)e'_1)|; \text{ etc.}$$

**Corolario:**

Dos espacios euclidianos de la misma dimensión son isométricos.

**Demostración:**

Consideremos dos espacios euclidianos  $(V, \alpha)$  y  $(V', \alpha')$  con bases ortonormales  $e_1, \dots, e_n$

y  $e'_1, \dots, e'_n$  respectivamente, definimos  $T: V \rightarrow V'$  mediante  $T(e_i) = e'_i$ . Si  $x = x_1 e_1 + \dots + x_n e_n$ ,

$y = y_1 e_1 + \dots + y_n e_n$ ,  $x_i, y_i \in \mathbb{R}$ , entonces

$$\alpha'(T(x), T(y)) = \alpha'(T(\sum x_i e_i), T(\sum y_j e_j))$$

$$= \sum x_i y_j \alpha'(T(e_i), T(e_j))$$

$$= \sum x_i y_j \alpha'(e'_i, e'_j) = \sum x_i y_j \alpha(e_i, e_j)$$

$$= \alpha(x, y).$$

Por lo tanto,  $T$  es isometría.

Ahora consideremos un espacio vectorial simpléctico

$$(V, \omega), \omega: V \times V \rightarrow \mathbb{R} \text{ forma simpléctica.}$$

Una base simpléctica de  $(V, \omega)$  es una base de  $V$   $\{e_1, \dots, e_m, f_1, \dots, f_m\}$  con  $2m$  elementos que cumplen con esto:

$$\omega(e_i, f_i) = 1$$

$$\omega(e_i, e_j) = 0$$

$$\omega(f_i, f_j) = 0$$

$$\omega(e_i, f_j) = 0 \text{ si } i \neq j$$

**Ejemplo:**

$$(\mathbb{R}^2, \det), e_1 = (1, 0) \quad \omega(e_1, f_1) = 1 \quad f_1 = (0, 1)$$

$$M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{es la matriz de esta forma simpléctica.}$$

**Ejemplo:**

$$(\mathbb{R}^4, \omega) \text{ con } \omega(x, y) = x_1 y_2 - x_2 y_1 + x_3 y_4 - x_4 y_3$$

$$e_1 = (1, 0, 0, 0)$$

$$e_2 = (0, 1, 0, 0)$$

$$e_3 = (0, 0, 1, 0)$$

$$e_4 = (0, 0, 0, 1)$$

$$\omega(e_1, e_2) = 1$$



$\omega(e_3, e_4) = 1$ , los demás son 0.

$M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$  es la matriz de esta forma simpléctica

con la base simpléctica ordenada  $(e_1, e_3, e_2, e_4)$

Tenemos que  $M = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$

**Teorema:**

Todo espacio vectorial simpléctico tiene una base simpléctica.

**Demostración:**

Hagámosla por inducción sobre la dimensión del espacio. Para  $\dim=0, 1$ . No hay nada que demostrar.

Consideremos un vector  $e_1 \neq 0$ , como  $\omega$  es no degenerada tenemos que existe  $f_1 \neq 0$  tal que  $(e_1, f_1) = \lambda \neq 0$  entonces tomamos  $f_1 = f_1 / \lambda$ , así,  $\omega(e_1, f_1) = \omega(e_1, f_1 / \lambda) = 1 / \lambda \omega(e_1, f_1) = 1$

Ahora me gustaría otro vector linealmente independiente de  $e_1$  y  $f_1$ ; llamemos  $V_1$  al espacio generado por  $e_1$  y  $f_1$  (de dimension=2), y consideremos el ortogonal simpléctico

$V_1^\perp = \{x \in V : \omega(x, e_1) = \omega(x, f_1) = 0\}$

Afirmo que  $V = V_1 \oplus V_1^\perp$

Demostración de la afirmación

a)  $x \in V_1 \cap V_1^\perp \Rightarrow x \in V_1$  y  $x \in V_1^\perp$

$\therefore x = ae_1 + bf_1$

Como  $x \in V_1^\perp$ ,  $\omega(x, e_1) = 0$ , por lo que  $b = 0$

y  $\omega(x, f_1) = 0$ , por lo que  $a = 0$

así,  $x = 0$ , es decir,  $V_1 \cap V_1^\perp = \{0\}$

Ahora voy a demostrar que cualquier vector en  $V$  es la suma de uno en  $V_1$  y uno en  $V_1^\perp$ .

b) sea  $x \in V$

$x_1 = \omega(x, f_1)e_1 - \omega(x, e_1)f_1$  es la "proyección" de  $x$  en  $V_1$ ,  $x_1 \in V_1$ ; luego,  $x - x_1$  está en  $V_1^\perp$ , ya que:

i)  $\omega(x - x_1, e_1) = \omega(x, e_1) - \omega(x_1, e_1)$

$= \omega(x, e_1) - \omega(\omega(x, f_1)e_1 - \omega(x, e_1)f_1, e_1)$

$= \omega(x, e_1) - \omega(x, f_1)\omega(e_1, e_1) + \omega(x, e_1)\omega(f_1, e_1) = 0$

ii)  $\omega(x - x_1, f_1) = \omega(x, f_1) - \omega(x_1, f_1)$

$$= \omega(x_1, f_1) - \omega(x, f_1) e_1 - \omega(x, e_1) f_1, f_1$$

$$= \omega(x, f_1) - \omega(x, f_1) \omega(e_1, f_1) + \omega(x, e_1) \omega(f_1, f_1) = 0$$

entonces  $x = x_1 + (x - x_1)$  con  $x_1 \in V_1$  y  $(x - x_1) \in V_1^\perp$ .

Por hipótesis de inducción  $V_1^\perp$  tiene una base simpléctica  $\{e_2, \dots, e_m, f_2, \dots, f_m\}$ , entonces tenemos que  $\{e_1, e_2, \dots, e_m, f_1, f_2, \dots, f_m\}$  es una base simpléctica del espacio vectorial simpléctico  $(V, \omega)$ .

**Corolario 1:**

Todo espacio vectorial simpléctico tiene dimensión par.

**Corolario 2**

En un espacio de dimensión impar, cualquier forma bilineal antisimétrica es degenerada.

**Corolario 3**

Dos espacios vectoriales simplécticos de la misma dimensión son simplécticamente equivalentes.

**Demostración**

Consideremos dos espacios vectoriales simplécticos  $(V, \omega)$  y  $(V', \omega')$  con bases simplécticas  $g_1, \dots, g_{2n}$  y  $g'_1, \dots, g'_{2n}$  respectivamente, donde

$$g_i = e_i \text{ para } i \leq n$$

$$g_i = f_i \text{ para } n < i$$

$$\text{y } g'_i = e'_i \text{ para } i \leq n$$

$$g'_i = f'_i \text{ para } n < i$$

Definimos  $T: V \rightarrow V'$  mediante  $T(g_i) = g'_i$

si  $x = x_1 g_1 + \dots + x_{2n} g_{2n}$  y  $y = y_1 g'_1 + \dots + y_{2n} g'_{2n}$ ,  $x_i, g_i \in R$

Entonces  $\omega'(T(x), T(y)) = \omega'(\sum x_i T(g_i), \sum y_j T(g'_j))$

$$= \sum x_i y_j \omega'(T(g_i), T(g'_j))$$

$$= \sum x_i y_j \omega'(g'_i, g'_j)$$

$$= \sum x_i y_j \omega'(g_i, g_j) = \omega(x, y).$$

Por lo tanto,  $(V, \omega)$  y  $(V', \omega')$  son "isométricos" o simplécticamente equivalentes.

Sólo hay un espacio vectorial simpléctico de cada dimensión; el espacio vectorial simpléctico de dimensión  $2n$  es  $R^{2n}$ .

Consideremos la base usual de  $R^{2n}$

$$e_1 = (1, 0, \dots, 0)$$

$$\vdots$$

$$e_n = (0, \dots, 1, 0, \dots, 0)$$

$$f_1 = (0, \dots, 0, 1, \dots, 0)$$

$$\vdots$$

$$f_n = (0, \dots, 0, 0, \dots, 1)$$

Y definamos  $\omega$  de modo que  $\{e_1, \dots, f_n\}$  sea una base simpléctica:

$$\omega(x,y) = \sum_{i=1}^n \begin{vmatrix} X_i & X_{i+n} \\ Y_i & Y_{i+n} \end{vmatrix};$$

así,  $\{e_1, \dots, f_n\}$  es una base simpléctica y  $(\mathbb{R}^{2n}, \omega)$  un espacio vectorial simpléctico.

Una notación usual de la física es

$x = (q_1, \dots, q_n, p_1, \dots, p_n)$ , así podemos escribir  $\omega(x, x') = \sum (q_i p'_i - p_i q'_i)$

La matriz de esta forma respecto a la base estándar es

$\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = J$  donde  $I$  es la matriz identidad  $n \times n$ , por lo tanto podemos escribir

$$\omega(x,y) = x^t J y$$

Si  $T: \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$  es lineal, entonces como ya dijimos  $T$  es simpléctica  $\Leftrightarrow$

$$\omega(T(x), T(y)) = \omega(x, y) \quad \forall x, y \in \mathbb{R}^{2n}.$$

es decir, si y sólo si  $x^t A^t J A y = x^t J y \quad \forall x, y \in \mathbb{R}^{2n}$

Por lo tanto, si  $A$  es la matriz de  $T$  entonces  $T$  es simpléctica si y sólo si  $(Ax)^t J (Ay) = x^t J y$   
 $\forall x, y \in \mathbb{R}^{2n}$

Por lo que  $T$  es simpléctica  $\Leftrightarrow A^t J A = J$ .

Esto nos conduce a la siguiente definición:

Diremos que una matriz  $A$   $2n \times 2n$  es simpléctica si y sólo si  $A^t J A = J$  (esto implica que  $A$  debe de ser no singular).

Ejemplo:

$$\text{En } \mathbb{R}^2: \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A \quad A^t J A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$= \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} c & d \\ -a & -b \end{pmatrix}$$

$$= \begin{pmatrix} 0 & ad - bc \\ -(ad - bc) & 0 \end{pmatrix}.$$

Por lo tanto,  $A^t J A = J \Leftrightarrow \det A = 1$ , es decir,  $A$  es simpléctica  $\Leftrightarrow \det A = 1$ , como ya lo habíamos visto.

Si  $T: (V, \omega) \rightarrow (V', \omega')$  es lineal simpléctica

(con  $(V, \omega)$  con base simpléctica y  $(V', \omega')$  con base simpléctica)

Entonces la matriz  $A$  de  $T$  debe ser simpléctica.  $A^t J A = J$ .

Afirmación.

Si  $M$  es una matriz simpléctica, entonces  $M^{-1}$  también lo es.

Demostración.

$$M^t J M = J \Rightarrow (M^t J)(M J) = I \Rightarrow (M J^t)(M^t J) = I \Rightarrow M J M^t = J$$

Afirmación.

Si  $M$  es una matriz simpléctica, entonces  $M^{-1}$  también lo es.

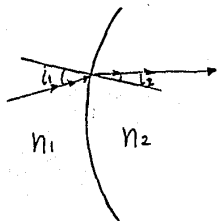
**Demostración. Ejercicio.**

Hemos definido formalmente lo que es la geometría simpléctica. Pero para entender más a fondo cual es su contenido, necesitamos remontarnos a sus orígenes y ver el contexto en que aparecen por primera vez las ideas geométricas que se irán formalizando hasta llegar a los conceptos anteriores. Este contexto es el de la óptica geométrica y la mecánica clásica. A continuación veremos un poco de óptica geométrica.

## Capítulo II

### Un poco de Óptica Geométrica

#### 2.1 Óptica Gaussiana

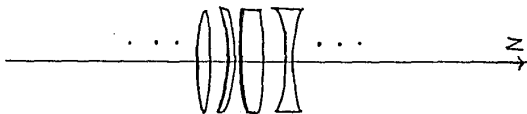


La óptica lineal es una aproximación lineal a la óptica geométrica, que es válida cuando los ángulos que entran en consideración son pequeños, y donde se asume que en el medio que hay entre dos superficies refractoras, el índice de refracción es constante (o equivalentemente: en el medio que hay entre dos superficies refractoras la velocidad de la luz es constante, ya que el índice de refracción  $n$  en un medio donde la luz viaja con velocidad constante  $v$ , se define como la razón entre las velocidades en el vacío y en el medio  $n=c/v$ ). En óptica lineal se hace la aproximación  $\sin\theta \approx \theta$ ,  $\cos\theta \approx 1$ ,  $\tan\theta \approx \theta$ ; así, por ejemplo, si en óptica geométrica la ley de Snell dice que si un rayo de luz

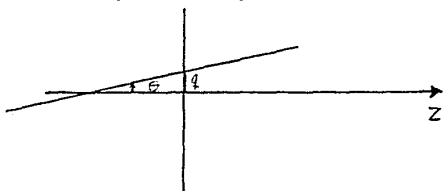
pasa de una región con índice de refracción  $n_1$  a una región con índice de refracción  $n_2$ , entonces  $n_1 \sin i_1 = n_2 \sin i_2$ , donde  $i_1$  e  $i_2$  son los ángulos que el rayo de luz forma con la normal a la superficie de separación entre las dos regiones. En óptica lineal la ley de Snell dice que  $n_1 i_1 = n_2 i_2$ , que es una buena aproximación si  $i_1$  e  $i_2$  son ángulos pequeños. Esta aproximación se conoce como ley de Ptolomeo.

La óptica gaussiana es una parte de la óptica lineal en la que sólo se consideran sistemas ópticos, en los que las superficies refractoras (o reflectoras) son simétricas por rotación respecto a un eje central que llamaremos eje óptico.

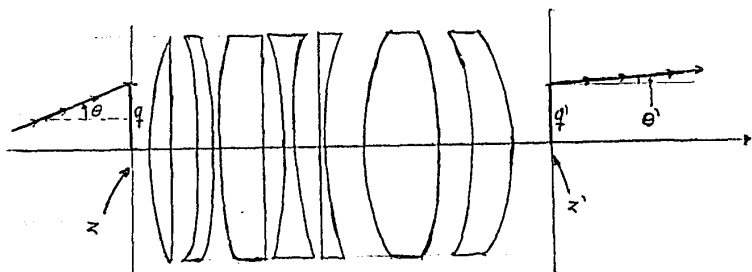
Introduzcamos un sistema coordenado de tal modo que el eje  $z$  coincida con el eje óptico. Por la simetría de rotación alrededor del eje óptico, podemos restringirnos al estudio de los rayos que se encuentran en un plano fijo que contenga al eje óptico, es decir al eje  $z$ .



Lo que queremos hacer ahora, es relacionar la línea recta de un rayo cuando entra a un sistema óptico, con la línea recta de éste cuando sale del sistema. Para esto, fijemos un plano perpendicular al eje óptico  $z$ . Para este plano ( $z=\text{cte.}$ ) especifiquemos cada recta mediante su altura  $q$  arriba del eje óptico en  $z$ , y con el ángulo  $\theta$  que esta recta forma con el eje óptico. El ángulo  $\theta$  lo mediremos en relación a la dirección positiva del eje  $z$  y en rotación contraria a las manecillas de un reloj. Esta especificación de las rectas puede hacerse para cada plano perpendicular al eje óptico.



Ahora, consideremos un plano  $z$  perpendicular al eje óptico. Entonces, lo que nos interesa ahora es poder relacionar la recta  $(q, \theta)$  de un rayo que incide (de izquierda a derecha) en el plano  $z$ , con la recta  $(q', \theta')$  cuando el rayo emerge del plano  $z'$ ; es decir, nuestro problema es tener  $(q', \theta')$  como función de  $(q, \theta)$



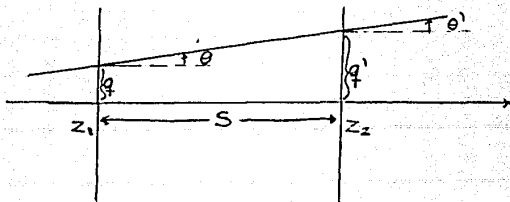
Vamos a asumir que nuestros sistemas ópticos no contienen espejos.

Consideremos el caso en el que entre los planos  $z_1$  y  $z_2$  no hay cambio de medio, es decir, el índice de refracción del medio que hay entre  $z_1$  y  $z_2$  es constante, y donde la distancia de  $z_1$  a  $z_2$  es  $s = z_2 - z_1$ . (ver la siguiente figura)

En este caso el rayo no sufre ninguna desviación y tenemos que  $\theta' = \theta$  y  $(q' - q)/s = \tan \theta$ , pero como convenimos que  $\tan \theta = \theta$  (ya que estamos considerando que  $\theta$  es pequeño) por lo que  $q' - q = s\theta$ ; así tenemos una traslación:

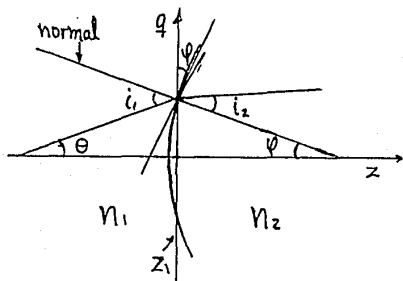
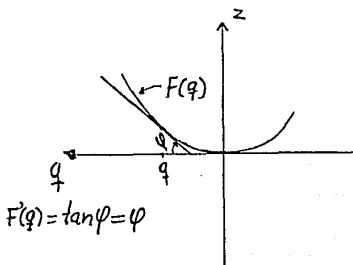
$$q' = q + s\theta$$

$$\theta' = \theta$$



por lo que podemos escribir que  $\begin{pmatrix} q' \\ \theta' \end{pmatrix} = \begin{pmatrix} 1 & S \\ 0 & 1 \end{pmatrix} \begin{pmatrix} q \\ \theta \end{pmatrix}$ , donde además  $\det \begin{pmatrix} 1 & S \\ 0 & 1 \end{pmatrix} = 1$ .

Ahora consideremos el caso donde sí hay cambio de medio en el trayecto de la luz, es decir, donde un rayo pasa de un medio con índice  $n_1$  a un medio con índice  $n_2$  con  $n_1 \neq n_2$ . En este caso el rayo  $(q, \theta)$  se transforma en el rayo  $(q', \theta')$  al atravesar la superficie de separación (por ejemplo la cara de un lente), por lo cual vamos a colocar a los dos planos  $z_1$  y  $z_2$  en el punto de contacto del rayo con la superficie refractora, o dicho de otro modo, el plano de referencia  $z_1$  lo tomamos "inmediatamente antes" de la superficie, y el plano  $z_2$  lo tomamos "inmediatamente después" de la superficie. En esta superficie de refracción el valor de  $q$  no cambia y el ángulo  $\theta$  cambia de acuerdo con la ley de Ptolomeo. Supongamos que la curva que da la intersección de la superficie con nuestro plano está dada por la función  $z=F(q)$ . Esta función es par, por la simetría respecto al eje  $z$ , y como suponemos que es suave, podemos expresarla como  $F(q)=z_1+1/2kq^2+\dots+\dots$ . Si ignoramos los términos de gradomayor que dos, tenemos que  $F'(q)=kq$ .



De acuerdo con los dibujos, tenemos que  $i_1 = \theta + \varphi = \theta + F'(q) = \theta + kq$  y similarmente obtenemos que  $i_2 = \theta' + kq' = \theta' + kq$ . Así que; usando la ley de Ptolomeo (ley de Snell linealizada) tenemos que  $n_1(\theta + kq) = n_2(\theta' + kq)$ , es decir,  $n_1\theta + n_2kq = n_2\theta' + n_2kq$  por lo que  $\theta' = n_1/n_2\theta + kq(n_1/n_2 - 1)$ . De esta manera tenemos la refracción:

$$\begin{cases} q' = q \\ \theta' = n_1/n_2\theta + k(n_1/n_2 - 1)q \end{cases} \text{ por lo que podemos}$$

escribir  $\begin{pmatrix} q' \\ \theta' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ k(n_1/n_2 - 1) & n_1/n_2 \end{pmatrix} \begin{pmatrix} q \\ \theta \end{pmatrix}$  pero sucede que  $\det \begin{pmatrix} 1 & 0 \\ k(n_1/n_2 - 1) & n_1/n_2 \end{pmatrix} = \frac{n_1}{n_2}$

y como  $n_1/n_2 \neq 1$ , la matriz no es simpléctica.\* Volvamos a empezar, pero buscando que la transformación tenga determinante 1. Como tenemos que  $n_1\theta + n_1kq = n_2\theta' + n_2kq$ , podemos hacer un cambio de coordenadas haciendo  $p = n\theta$  para así obtener  $p + n_1kq = p' + n_2kq$  por lo que  $p' = p + (n_1 - n_2)kq$  y  $q' = q$ ; así podemos escribir

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -P & 1 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}$$

donde  $P = (n_2 - n_1)k$ , que es lo que se conoce por el poder de la superficie refractora y ahora si

$$\det \begin{pmatrix} 1 & 0 \\ -P & 1 \end{pmatrix} = 1$$

Sólo falta ver cómo queda la traslación con el cambio de coordenadas.

Teníamos que  $q' = q + s\theta$  y  $\theta' = \theta$  que ahora pueden escribirse como  $q' = q + s/n p$  y  $p' = p$ ; ahora la traslación puede escribirse como:

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}$$

donde  $d = s/n$ , que es lo que se llama la distancia reducida, y se tiene que

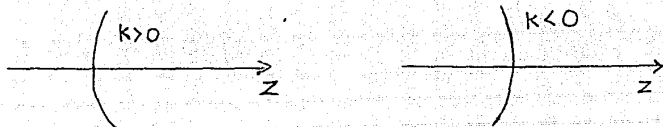
$$\det \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} = 1$$

Con esto podemos decir que a cada sistema óptico gaussiano, con un par de planos de referencia, podemos asignarle una matriz simpléctica. Recordemos que el signo de  $k$  depende de la forma de la función  $F(q)$ , ya que será útil en el siguiente ejemplo. (ver la figura)

---

\*Para el caso en el que se tiene un sistema con  $(k-1)$  superficies refractoras, donde el primer medio y el último medio tienen el mismo índice de refracción  $n_1 = n_k$  se tiene que  $\det M = \frac{n_1}{n_2} \frac{n_2}{n_3} \dots \frac{n_k}{n_1} = 1$  por lo que la matriz resulta ser simpléctica.

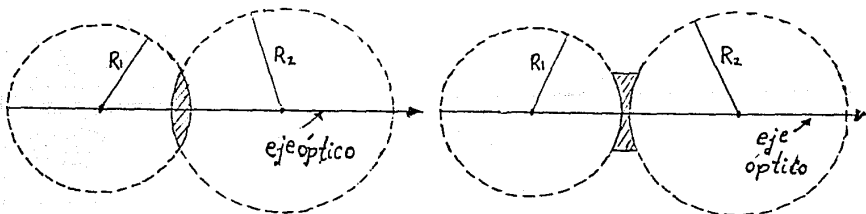




Ejemplo:

Fórmula del lente delgado.

Los lentes más comunes son objetos transparentes con dos caras, y por lo menos una de ellas es esférica. El estudio se simplifica si se restringe a los lentes delgados, que son los que tienen un ancho pequeño, en relación a los radios de las esferas que los limitan.



Vamos a suponer que nuestros lentes son tan delgados, que despreciaremos su espesor, es decir, vamos a pensar que una superficie se encuentra inmediatamente después de la otra. De esta manera obtenemos la matriz del lente delgado:

$$\begin{pmatrix} 1 & 0 \\ -P_2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -P_1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -P_2 - P_1 & 1 \end{pmatrix}$$

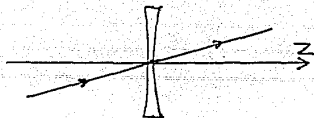
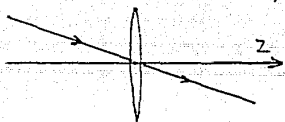
con  $P_1 = k(n_2 - n_1) = 1/R_1(n_2 - n_1)$  y  $P_2 = k_2(n_1 - n_2) = 1/R_2(n_1 - n_2)$ , por eso la matriz del lente delgado es:

$$\begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix}$$

donde  $1/f = (n_2 - n_1) (1/R_2 - 1/R_1)$  y a  $f$  se le llama distancia focal de la lente.

Una primera observación que podemos hacer es que un rayo de luz que atraviesa un lente delgado por el centro, no sufre cambio alguno:

$$\begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \rho \end{pmatrix} = \begin{pmatrix} 0 \\ \rho \end{pmatrix}$$



Consideraremos el caso más común en el que el medio que hay entre los lentes es aire ( $n_1 = 1$ ), y el medio del que están compuestos los lentes tiene índice de refracción  $n_2 > 1$ ; así tenemos que  $(n_2 - n_1) > 0$ , por lo que el signo de  $f$  es el mismo que el de  $(1/R_2 - 1/R_1)$ . En la figura siguiente se muestran los lentes mas comunes

a) biconvexo;



b) plano-convexo;

c) menisco convergente;

a) b) c)

d) bicóncavo;

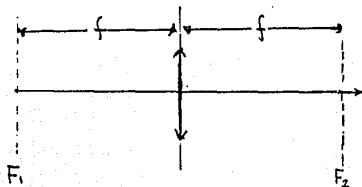


e) plano-cóncavo;

f) menisco divergente.

d) e) f)

Consideremos un lente delgado con distancia focal  $f > 0$  y representemoslo con el símbolo  $\uparrow$  y pensemos en el caso en el que el lente se encuentra entre un plano de referencia  $F_1$ , localizado a una distancia  $f$  (la distancia focal) a la izquierda del lente, y un plano de referencia  $F_2$  localizado a una distancia  $f$  a la derecha.

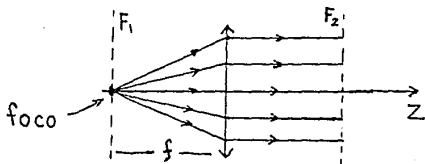


$$\begin{pmatrix} 1 & f \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix} \begin{pmatrix} 1 & f \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & f \\ -1/f & 0 \end{pmatrix}$$

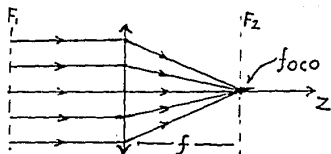
en este caso  $\begin{pmatrix} q' \\ p' \end{pmatrix} = \begin{pmatrix} 0 & f \\ -1/f & 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} fp \\ -q/f \end{pmatrix}$  por lo que  $q'$  no

depende de  $q$  y  $p'$  no depende de  $p$ .

Así se tiene que todo el haz de rayos que parten del foco del lado izquierdo (que es el punto que está a una distancia  $f$  a la izquierda del lente, sobre el eje óptico) atraviesan el lente y emergen paralelamente al eje óptico ya que  $q = 0$  implica  $p' = 0$ ;

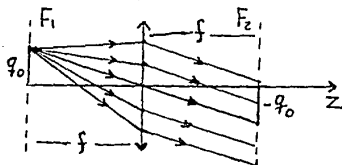


y también se tiene que los rayos de luz que inciden sobre el lente paralelamente al eje ( $p = 0$ ) convergen sobre el foco del lado derecho, ya que  $p = 0$  implica  $q' = 0$

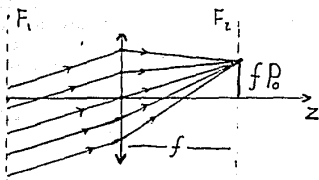


Se les llama planos focales a los planos de referencia (perpendiculares al eje óptico) que contienen a los focos.

Así tenemos que los rayos que parten de un punto de altura  $q_0$  en el plano focal izquierdo, atraviesan el lente y emergen paralelamente al rayo que atraviesa el centro de la lente, ya que  $p' = -q_0/f$ ;



y reciprocamente, un haz de rayos paralelos ( $q, p_0$ ), que atraviesa el lente, converge en un punto del plano focal de altura  $q' = fp_0$ .

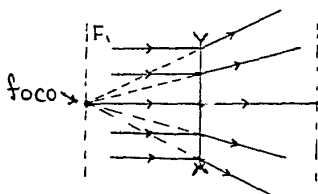


Ahora consideremos un lente delgado de distancia focal  $f > 0$  (es decir  $1/R_2 - 1/R_1 < 0$ ), representémoslo por el símbolo  $\chi$  y veamos que pasa si un rayo atraviesa el lente y luego hacemos una traslación hacia la izquierda:  $\begin{pmatrix} 1 & f \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ -1/f & 1 \end{pmatrix} = \begin{pmatrix} 0 & f \\ -1/f & 1 \end{pmatrix}$ ,

en este caso  $\begin{pmatrix} q' \\ p' \end{pmatrix} = \begin{pmatrix} 0 & f \\ -1/f & 1 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} fp \\ -q/f + p \end{pmatrix}$ ,

por lo que  $q'$  no depende de  $q$ .

Así, los rayos que inciden paralelamente al eje óptico, salen del lente como si procedieran del foco del lado izquierdo, ya que  $p = 0$  implica  $q' = 0$

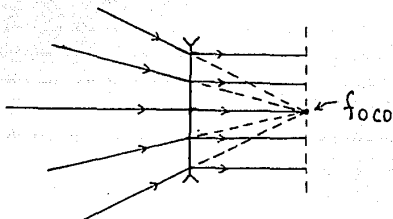


*Un observador del lado derecho del lente ve un punto luminoso en el foco.*

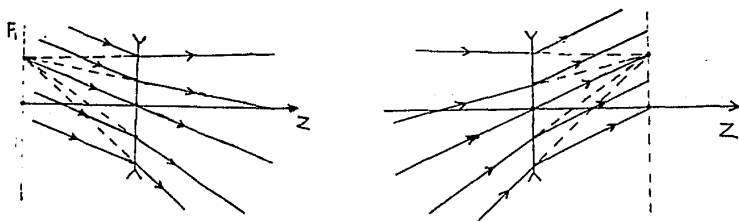
Ahora veamos que pasa con un haz de rayos que, antes de atravesar el lente, van dirigidos al foco del lado derecho; para esto, tomemos como primer plano de referencia el plano focal derecho y luego hagamos una traslación hacia la izquierda para tomar como segundo plano el mismo lente:  $\begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix} \begin{pmatrix} 1 & f \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & f \\ -1/f & 0 \end{pmatrix}$ ,

en este caso  $\begin{pmatrix} q' \\ p' \end{pmatrix} = \begin{pmatrix} 1 & f \\ -1/f & 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} q + fp \\ -q/f \end{pmatrix}$   
 por lo que  $p'$  no depende de  $p$ .

Así se tiene que los rayos que se dirigen hacia el foco del lado derecho, emergen paralelamente al eje óptico ya que  $q = 0$  implica  $p' = 0$



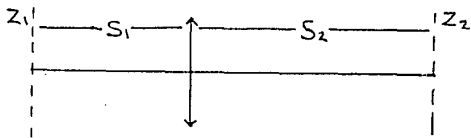
Para los planos focales, se hace un análisis similar al que hicimos para  $f > 0$ . Los resultados correspondientes se ven en las siguientes figuras



A los lentes con distancia focal  $f > 0$  se les llama convergentes, y a los lentes con distancia focal  $f < 0$  se les llama divergentes.

¿Qué pasa con un lente delgado si no estamos en los planos focales?

Tomemos un plano de referencia  $z_1$  a una distancia  $s_1$  a la izquierda del lente y un plano de referencia  $z_2$  a una distancia  $s_2$  a la derecha del lente.



La matriz correspondiente a estos dos planos es

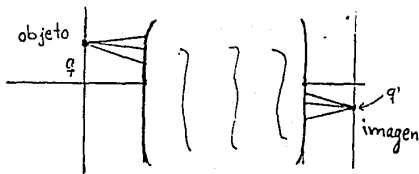
$$\begin{pmatrix} 1 & S_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix} \begin{pmatrix} 1 & S_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-S_2/f & S_2+S_1-S_1S_2/f \\ -1/f & 1-S_1/f \end{pmatrix}$$

### Definición:

Dos planos de un sistema óptico son conjugados, si la matriz tiene la forma

$$\begin{pmatrix} A & 0 \\ C & D \end{pmatrix} \quad \text{con } AD = 1.$$

Es decir,  $q' = Aq$  no depende de  $p$ , por lo que podemos decir, que dos planos  $z_1$  y  $z_2$  son conjugados si todos los rayos que salen de un punto de altura  $q$  en  $z_1$ , convergen en un punto de altura  $q'$  en  $z_2$ .



La magnificación de la imagen en respecto al objeto es  $q'/q = A$ .

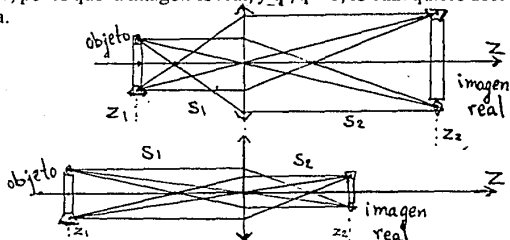
Para el caso del lente delgado los planos son conjugados si  $s_1 + s_2 - s_1s_2/f = 0$ , es decir, si  $1/s_1 + 1/s_2 = 1/f$ ; así la matriz toma la forma

$$\begin{pmatrix} 1-S_2/f & 0 \\ -1/f & 1-S_1/f \end{pmatrix}.$$

Consideremos un objeto luminoso delgado colocado en el plano  $z_1$ ; con lo visto hasta ahora es fácil dibujar su imagen y por tanto encontrar su plano conjugado  $z_2$ . Además la magnificación de la imagen respecto al objeto es  $q'/q = 1 - s_2/f = 1 - s_2(1/s_1 + 1/s_2) = -s_2/s_1$ .

Si  $f > 0$  y  $s_1 > f$ , entonces  $s_2 > 0$ , por lo que la imagen es real, y  $q'/q < 0$ , lo cual quiere decir que la imagen está invertida.

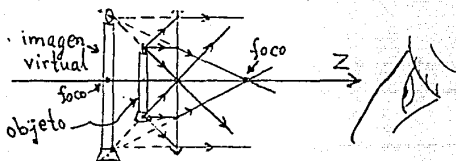
Si se pone una pantalla en el plano conjugado  $z_2$ , se verá una imagen real del objeto sobre la pantalla.



Si  $f > 0$  y  $f > s_1 > 0$ , entonces  $s_2 < 0$  por lo que la imagen es virtual, y  $q'/q > 0$  lo cual quiere decir que la imagen conserva la orientación. Además  $|s_2| > s_1$ , por lo que la imagen es más

grande que el objeto.

Através del lente no se ve el objeto sino su imagen virtual



Pregunta para el lector:

¿Cómo son las imágenes para el lente divergente?

(Hacer los esquemas de rayos)

Otro sistema óptico gaussiano importante es el telescopio astronómico que es un instrumento que se construye colocando un lente de distancia focal  $f_1$  (llamado objetivo) con un lente de distancia focal  $f_2$  (llamado ocular) separados por una distancia  $f_1 + f_2$  entre sí.

La matriz de este sistema es:

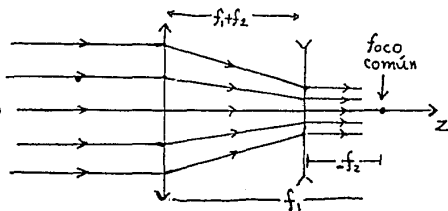
$$\begin{pmatrix} 1 & 0 \\ -1/f_2 & 1 \end{pmatrix} \begin{pmatrix} 1 & f_1 + f_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1/f_1 & 1 \end{pmatrix} = \begin{pmatrix} -f_2/f_1 & f_1 + f_2 \\ 0 & -f_1/f_2 \end{pmatrix}.$$

En este caso tenemos que  $p' = -f_1/f_2 p$ , por lo que los rayos paralelos (de una estrella distante, por ejemplo) que entran en este sistema óptico, son convertidos en otros rayos paralelos (que inciden en el ojo). La magnificación angular que se obtiene, es la razón de la pendiente de los rayos que salen con la pendiente de los rayos que entran:  $p'/p = -f_1/f_2$ . Como se ve, para obtener una buena magnificación angular es conveniente que la distancia focal del ocular sea pequeña comparada con la distancia focal del objetivo.

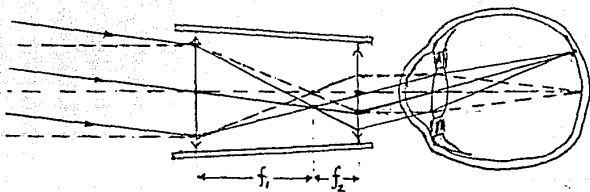
Telescopio Galileano

$f_1 > 0, f_2 < 0$ .

El aumento angular es positivo



$f_1 > 0, f_2 > 0$   
El aumento angular  
es negativo.



Diremos que una matriz es telescópica si tiene la forma  $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$   
con  $AD = 1$ ; la magnificación angular  $p'/p = D$

**Teorema:**

Toda matriz  $2 \times 2$  con  $\det = 1$  (en este caso es lo mismo que simpléctica) se puede realizar mediante un sistema óptico lineal (plano y simétrico respecto al eje z).

**Demstración:**

Sea  $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$  una matriz simpléctica:  $AD - BC = 1$ .

a) Si la matriz es telescópica ( $C=0$ ), quiero mostrar que si  $A \neq 1 \exists f_1, f_2 \in \mathbb{R}, f_1, f_2 \neq 0$  tales que

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1/f_2 & 1 \end{pmatrix} \begin{pmatrix} 1 & f_1+f_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1/f_1 & 1 \end{pmatrix};$$

es decir, busco  $f_1$  y  $f_2$  tales que  $A = -f_2/f_1, D = -f_1/f_2$  y  $B = f_1 + f_2$ . Si  $A \neq 1$ , tenemos que  $f_2 = -f_1 A$ , por lo que  $B = f_1 - f_1 A$ ; así,  $f_1 = B/(1 - A)$  y  $f_2 = AB/(A - 1)$ . Si  $A = 1$  entonces  $D = 1$ , y entonces tenemos una traslación  $\begin{pmatrix} 1 & B \\ 0 & 1 \end{pmatrix}$ .

**Segunda demostración del caso del telescopio:**

a\*) Si la matriz es telescópica, entonces tomemos  $p \neq 0$ . Así el producto  $\begin{pmatrix} 1 & 0 \\ -p & 1 \end{pmatrix} \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \begin{pmatrix} A & B \\ -pA & D-pB \end{pmatrix}$

es una matriz no telescópica, de modo que podemos escribir  $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ p & 1 \end{pmatrix} \begin{pmatrix} A & B \\ -pA & D-pB \end{pmatrix}$

y nos vamos al caso

b) (no-telescópico)  $c \neq 0$ . Consideremos el producto  $\begin{pmatrix} 1 & s_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 1 & s_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} A+s_1C & s_2(A+s_1C)+(B+s_1D) \\ C & s_2C+D \end{pmatrix} = \begin{pmatrix} A' & B' \\ C & D' \end{pmatrix}$

Ahora escojamos  $s_1$  y  $s_2$  de modo que  $A'=1$  y  $B'=0$ .  $A + s_1C = 1$  y  $s_2 + B + s_1D = 0$  implican que  $s_1 = (1 - A)/C$  y  $s_2 = -B - s_1D$ ; como las matrices son simplécticas su producto es simpléctico, y entonces  $D'=1$ , por lo que la matriz que nos queda es

$$\begin{pmatrix} 1 & 0 \\ C & 1 \end{pmatrix}; \text{ así, } \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & -s_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ C & 1 \end{pmatrix} \begin{pmatrix} 1 & -s_2 \\ 0 & 1 \end{pmatrix}$$

corresponde a dos traslaciones y una refracción. ♦

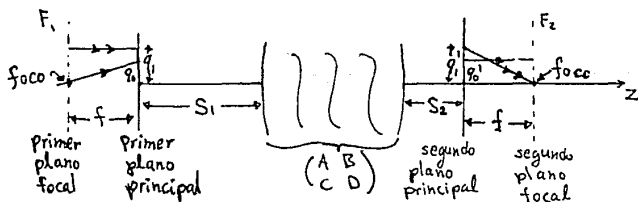


Nótese que  $s_1$  y  $s_2$  están determinados de forma única. Entonces para cualquier sistema óptico gaussiano no-telescópico,

$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$  con  $c \neq 0$ , hay dos planos únicos tales que la matriz entre ellos tiene la forma  $\begin{pmatrix} 1 & 0 \\ C & 1 \end{pmatrix}$ . Estos planos son conjugados y tienen magnificación 1; Gauss los llamó planos principales. Si estudiamos la matriz entre estos dos planos

$\begin{pmatrix} 1 & 0 \\ C & 1 \end{pmatrix}$ , podemos proceder exactamente como lo hicimos para el lente delgado, escribiendo  $C = -P = -1/f$ . Así los dos planos focales del sistema están colocados  $f$  unidades a la izquierda y derecha de los planos principales:

$$\begin{pmatrix} 1 & f \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix} \begin{pmatrix} 1 & f \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & f \\ -1/f & 0 \end{pmatrix}$$



Con lo anterior podemos afirmar que cualquier sistema óptico gaussiano no-telescópico queda determinado por tres parámetros: la localización de los planos principales y la distancia focal.

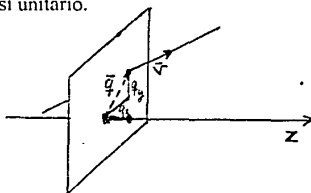
Podemos resumir los resultados de esta sección diciendo que hay un isomorfismo entre el grupo simpléctico  $SL(2, \mathbb{R})$  (de las matrices  $2 \times 2$  con  $\det = 1$ ) y la óptica gaussiana. Cada matriz corresponde a un sistema óptico; y la multiplicación de matrices corresponde a la composición de los correspondientes sistemas ópticos.

## 2.2 Óptica lineal

Veamos que pasa si no asumimos simetría rotacional pero seguimos con la aproximación lineal (los términos de orden mayor que uno en los ángulos y distancias pueden ser ignorados).

Para especificar un rayo necesitamos dar su intersección con el plano paralelo al plano  $x$ - $y$  que pasa por el punto  $z$  en el eje óptico, y su dirección, que puede especificarse por un vector de la forma  $(V_x, V_y, 1) = \vartheta$ . Es claro que  $\vartheta_x = \tan\theta_x = \theta_x$  y  $\vartheta_y = \tan\theta_y = \theta_y$ , como estamos asumiendo que los rayos en consideración son cercanos al eje óptico (los ángulos son pequeños) se tiene que  $\vartheta$  es casi unitario.

$(q_x, q_y)$  posición  
 $(\vartheta_x, \vartheta_y, 1)$  dirección

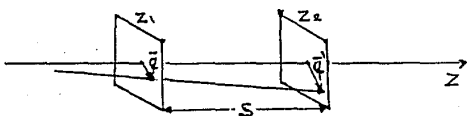


Si ahora ponemos  $p_x = nV_x$  y  $p_y = nV_y$  (donde  $n$  es el índice de refracción donde viaja la luz) entonces cualquier rayo queda determinado por las coordenadas  $(q_x, q_y, p_x, p_y)$ , aquí  $q_x = x$  y  $q_y = y$ .

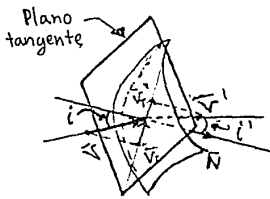
Veamos ahora como queda expresada la traslación, es decir, consideremos el caso en el que entre los planos  $z_1$  y  $z_2$  no hay cambio de medio. En este caso el rayo no sufre ninguna desviación y por tanto  $p'_x = p_x$  y  $p'_y = p_y$ ; si  $z_2 - z_1 = s$ , entonces  $(q'_x - q_x)/s = \tan\theta_x = V_x$  y  $(q'_y - q_y)/s = \tan\theta_y = V_y$ , por lo que

$$\begin{pmatrix} q'_x \\ q'_y \\ p'_x \\ p'_y \end{pmatrix} = \begin{pmatrix} q_x + d p_x \\ q_y + d p_y \\ p_x \\ p_y \end{pmatrix} \text{ es decir,}$$

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \begin{pmatrix} I & dI \\ 0 & I \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} \quad \text{donde } d = s/n.$$



Vamos a ver como es la refracción. (observese la siguiente figura)



Si la superficie refractante está dada por la función  $F(q_x, q_y) = z$ , entonces podemos pasarla como una superficie de nivel de la función  $G(q_x, q_y, z) = z - F(q_x, q_y)$  para  $G = 0$ ; por lo que la normal será  $N = \nabla G = (-\partial F/\partial q_x, -\partial F/\partial q_y, 1)$ .

Estamos pensando que el eje  $z$  es normal a la superficie, es decir  $\partial F(0,0)/\partial q_x = \partial F(0,0)/\partial q_y = 0$ , por lo tanto  $\partial F/\partial q_x$  y  $\partial F/\partial q_y$  son pequeñas cerca del  $(0,0)$  y  $N$  es casi unitario. En nuestra aproximación  $|\vartheta| = |\vartheta'| = 1$  por lo que  $\text{sen } i = |\vartheta \cdot N|$  y  $\text{sen } i' = |\vartheta' \cdot N|$ . Como los vectores  $\vartheta$ ,  $\vartheta'$  y  $N$  están en un mismo plano, entonces las proyecciones de  $\vartheta$  y  $\vartheta'$  sobre el plano tangente quedan en una sola línea y además tienen el mismo sentido, por lo que la ley de Snell puede escribirse así:  $n(\vartheta \cdot (\vartheta \cdot N)N) = n'(\vartheta' \cdot (\vartheta' \cdot N)N)$ , es decir,  $n\vartheta_\tau = n'\vartheta'_\tau$ . Esto nos da tres igualdades, pero antes de verlas observemos que  $n(\vartheta \cdot N) = n(-\vartheta_x \partial F/\partial q_x - \vartheta_y \partial F/\partial q_y + 1) = n - p_x \partial F/\partial q_x - p_y \partial F/\partial q_y = n - p \cdot \nabla F$  donde  $p = (p_x, p_y)$ .

Así,  $n\vartheta_\tau = n[(\vartheta_x, \vartheta_y, 1) \cdot n - p \cdot \nabla F] \cdot (-\partial F/\partial q_x, -\partial F/\partial q_y, 1)$   
 $= (p_x - n \partial F/\partial q_x - (p \cdot \nabla F) \partial F/\partial q_x, p_y + n \partial F/\partial q_y - (p \cdot \nabla F) \partial F/\partial q_y, p \cdot \nabla F)$ , análogamente  
 $n'\vartheta'_\tau = (p'_x - n' \partial F/\partial q_x - (p' \cdot \nabla F) \partial F/\partial q_x, p'_y + n' \partial F/\partial q_y - (p' \cdot \nabla F) \partial F/\partial q_y, p' \cdot \nabla F)$

Por lo tanto la ley de Snell nos da las siguientes tres igualdades:

$$p'_x - n' \partial F/\partial q_x - (p' \cdot \nabla F) \partial F/\partial q_x = p_x - n \partial F/\partial q_x - (p \cdot \nabla F) \partial F/\partial q_x$$

$$p'_y + n' \partial F/\partial q_y - (p' \cdot \nabla F) \partial F/\partial q_y = p_y + n \partial F/\partial q_y - (p \cdot \nabla F) \partial F/\partial q_y$$

$$p' \cdot \nabla F = p \cdot \nabla F$$

Sustituyendo la última igualdad en las dos primeras, obtenemos  $p'_x = p_x + (n - n') \partial F/\partial q_x$  y  $p'_y = p_y + (n - n') \partial F/\partial q_y$ .

Suponemos que  $F$  es suave y podemos expresarla como

$F(q_x, q_y) = z_1 + 1/2 a q_x^2 + b q_y^2 + 1/2 c q_x^2 + \dots$ , si ignoramos los términos de orden mayor que dos tenemos que  $\partial F/\partial q_x = a q_x + b q_y$  y  $\partial F/\partial q_y = b q_x + c q_y$ , por lo que

$$\begin{pmatrix} q'_x \\ q'_y \\ p'_x \\ p'_y \end{pmatrix} = \begin{pmatrix} q_x \\ q_y \\ p_x - (n - n')(a q_x + b q_y) \\ p_y - (n - n')(b q_x + c q_y) \end{pmatrix} \text{ es decir, } \begin{pmatrix} q' \\ p' \end{pmatrix} = \begin{pmatrix} I & 0 \\ -P & I \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}$$

$$\text{donde } P = (n - n')k \text{ y } k = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 F}{\partial q_x^2} & \frac{\partial^2 F}{\partial q_x \partial q_y} \\ \frac{\partial^2 F}{\partial q_y \partial q_x} & \frac{\partial^2 F}{\partial q_y^2} \end{pmatrix}$$

que es simpléctica; por lo tanto también  $p = p'$

Un rayo está caracterizado por un punto  $(q, p)$  en  $R^2$  cerca del cero; y la forma simpléctica está dada por

$$\omega((q_1, p_1), (q_2, p_2)) = q_1 \cdot p_2 - q_2 \cdot p_1.$$

Si tengo  $T: (q, p) \rightarrow (q', p')$  con  $T$  traslación o refracción, quiero ver si se preserva la forma simpléctica  $\omega$ .

Veamos si la traslación preserva la forma simpléctica.

$$T: q' = q + dp$$

$$p' = p$$

$$\omega((q'_1, p'_1), (q'_2, p'_2)) = (q_1 + dp_1) \cdot p_2 - (q_2 + dp_2) \cdot p_1 = q_1 \cdot p_2 - q_2 \cdot p_1$$

$$= \omega((q_1, p_1), (q_2, p_2)).$$
  $T$  es simpléctica

Para la refracción tenemos  $R: q' = q$

$$p' = p - Pq \text{ con } P = P'$$

$$\omega((q'_1, p'_1), (q'_2, p'_2)) = q_1 \cdot (p_2 - Pq_2) - q_2 \cdot (p_1 - Pq_1)$$

$$= q_1 \cdot p_2 - q_2 \cdot p_1 + q_2 \cdot Pq_1 - q_1 \cdot Pq_2$$

$$= q_1 \cdot p_2 - q_2 \cdot p_1, \text{ ya que } P = P' \text{ implica que } q_2 \cdot Pq_1 = q_1 \cdot Pq_2$$

$$= \omega((q_1, p_1), (q_2, p_2)).$$
  $R$  es simpléctica.

En ambos casos la matriz es simpléctica, por lo tanto todas las matrices ópticas son simplécticas.

Otra forma de verlo es como lo hicimos al principio: una transformación lineal preserva la forma simpléctica si y sólo si su matriz cumple con  $M^T J M = J$ , donde

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}. \text{ Es decir, para que la matriz}$$

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \text{ sea simpléctica deben cumplirse las condiciones:}$$

$$A^T C \text{ simétrica}$$

$$B^T D \text{ simétrica}$$

$$A^T D - C^T B = I.$$

Estas condiciones se cumplen para la traslación:  $\begin{pmatrix} I & dI \\ 0 & I \end{pmatrix}$ ;

$$\text{y también para la refracción: } \begin{pmatrix} I & 0 \\ -P & I \end{pmatrix}$$

$$\text{ya que } P = P'.$$

Como se ve, toda matriz óptica tiene  $\det = 1$ ; sin embargo,

$$N = \begin{pmatrix} N_1 & 0 \\ 0 & N_1 \end{pmatrix} \text{ con } N_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

cumple que  $\det(N) = 1$  pero no es óptica, pues

$$N^t J N = \begin{pmatrix} 0 & N_i^z \\ -N_i^z & 0 \end{pmatrix} \neq J.$$

Nota: El invariante que se preserva no es de parejas de coordenadas; es decir, si  $\omega_1((q_1, p_1), (q_2, p_2)) = q_{1x}p_{2x} - q_{2x}p_{1x}$  y  $\omega_2((q_1, p_1), (q_2, p_2)) = q_{1y}p_{2y} - q_{2y}p_{1y}$ , entonces  $\omega = \omega_1 + \omega_2$  y  $\begin{pmatrix} \mathbb{I} & 0 \\ -P & \mathbb{I} \end{pmatrix}$

preserva  $\omega$  pero no preserva  $\omega_1$  ni  $\omega_2$ :

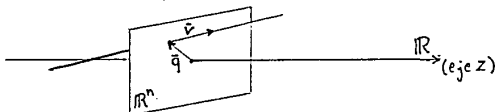
$$\begin{aligned} \omega_1((q'_1, p'_1), (q'_2, p'_2)) &= q_{1x}(p_{2x} + (n-n')(aq_{2x} + bq_{2y})) - q_{2x}(p_{1x} + (n-n')(aq_{1x} + bq_{1y})) \\ &= q_{1x}p_{2x} - q_{2x}p_{1x} + (n-n')b(q_{1x}q_{2y} - q_{2x}q_{1y}) \end{aligned}$$

$$= \omega_1((q_1, p_1), (q_2, p_2)) + (n-n')b(q_{1x}q_{2y} - q_{2x}q_{1y}); \text{ análogamente}$$

$$\omega_2((q'_1, p'_1), (q'_2, p'_2)) = \omega_2((q_1, p_1), (q_2, p_2)) + (n-n')b(q_{1x}q_{2x} - q_{2x}q_{1x}).$$

Generalizemos lo anterior pasando al caso en el que los rayos de luz viajan en  $\mathbb{R}^{n+1} = \{(q_1, \dots, q_n, z) : q_i, z \in \mathbb{R}\}$  y consideremos la aproximación lineal.

Para especificar un rayo necesitamos dar su intersección con el hiperplano perpendicular al eje óptico (eje  $z$ ) y su dirección, que puede especificarse por un vector de la forma  $(\vartheta_1, \vartheta_2, \dots, \vartheta_n, 1) = \vartheta$ . Es claro que  $\vartheta_i = \tan \theta_i = \theta_i$ . Como asumimos que los rayos en consideración son cercanos al eje óptico (los  $\vartheta_i$  son pequeños)  $\vartheta$  es casi unitario.



Si ponemos  $p_i = n\vartheta_i$ , tenemos que cualquier rayo queda determinado por las coordenadas  $(q_1, \dots, q_n, p_1, \dots, p_n)$ .

En el caso de la traslación el rayo no sufre ninguna desviación por lo que  $\bar{p}' = \bar{p}$ ; si  $s = z_2 - z_1$ , entonces  $(q'_i - q_i)/s = \tan \theta_i = V_i$  por lo que  $\bar{q}' = \bar{q} + d\bar{p}$ , es decir

$$\begin{pmatrix} \bar{q}' \\ \bar{p}' \end{pmatrix} = \begin{pmatrix} \mathbb{I}_n & d\mathbb{I}_n \\ 0 & \mathbb{I}_n \end{pmatrix} \begin{pmatrix} \bar{q} \\ \bar{p} \end{pmatrix} \quad \text{donde } d = s/n.$$

Para la refracción supondremos que la superficie refractora esta dada por  $z = F(q_1, \dots, q_n)$  por lo que la normal es  $N = -\partial F/\partial q_1, -\partial F/\partial q_2, \dots, -\partial F/\partial q_n, 1$ ; asumiremos que  $F$  es cuadrática  $F(q_1, \dots, q_n) = z_1 + \sum a_{ij}q_iq_j + \sum q_iq_j^2/2$  (el coeficiente de  $q_iq_j = q_jq_i$  denotamos de igual forma por  $a_{ij}$  que por  $a_{ji}$ )

Como  $\partial F(0,0)/\partial q_i = 0$  entonces  $\partial F/\partial q_i$  es pequeña cerca del cero para cada  $i$ , por lo que  $N$  es casi unitario. En nuestra aproximación  $|\vartheta| = |\vartheta'| = |N| = 1$  por lo que  $\text{seni} = |\vartheta - (\vartheta \cdot N)|$  y  $\text{seni}' = |\vartheta' - (\vartheta' \cdot N)|$ . Siguiendo el mismo camino que en  $\mathbb{R}^3$  se llega a que  $p'_i = p_i + (n-n')\partial F/\partial q_i$ , y como  $\partial F/\partial q_i = \sum \alpha_{ip}\theta_p$ , entonces  $\alpha_{ip} = \partial^2 F/\partial q_i \partial q_p = \partial^2 F/\partial q_i \partial q_p$ , por lo que  $\bar{p}' = \bar{p} - P\bar{q}$  donde  $P = (n'-n)k$  y

$$K = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 F}{\partial q_1^2} & \dots & \frac{\partial^2 F}{\partial q_1 \partial q_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial q_1 \partial q_n} & \dots & \frac{\partial^2 F}{\partial q_n^2} \end{pmatrix}$$

que es simétrica; así,  $(\bar{q}'^1) = \begin{pmatrix} I_n & 0 \\ -P & I_n \end{pmatrix} \begin{pmatrix} \bar{q} \\ \bar{p} \end{pmatrix}$

Es fácil verificar que  $\begin{pmatrix} I_n & 0 \\ -P & I_n \end{pmatrix}$  con  $P=P^t$ , y  $\begin{pmatrix} I_n & dI_n \\ 0 & I_n \end{pmatrix}$

son matrices simplécticas, por lo tanto toda matriz óptica es simpléctica.

El siguiente teorema es la generalización del teorema demostrado en óptica gaussiana. Para el caso de las matrices  $4 \times 4$  el teorema muestra que la óptica lineal es isomorfa al estudio del grupo simpléctico  $Sp(4, \mathbb{R})$ .

**Teorema:**

Toda matriz simpléctica  $2n \times 2n$  se puede realizar mediante un sistema óptico, es decir, se puede obtener multiplicando matrices de la forma

$\begin{pmatrix} I_n & dI_n \\ 0 & I_n \end{pmatrix}$  y de la forma  $\begin{pmatrix} I_n & 0 \\ -P & I_n \end{pmatrix}$

con  $P=P^t$ .

**Demostración.** (Gillemin & Sternberg, ver bibliografía)

**Corolario.**

Toda matriz simpléctica tiene determinante 1.

### 2.3 Principio de Fermat

La ley de Snell, con la cual ya hemos trabajado es consecuencia de un principio variacional, el principio de Fermat.

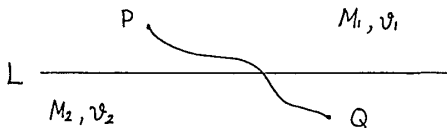
**Principio de Fermat:** La trayectoria que sigue la luz para llegar de un punto P a un punto Q es (entre todas las posibles trayectorias) la que hace que el tiempo de recorrido sea mínimo (o estacionario).

**Definición:**

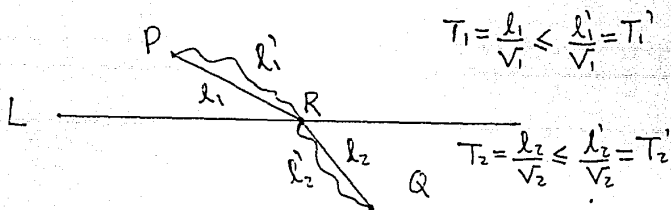
El tiempo  $T(\gamma)$  durante el cual se recorre la curva  $\gamma$  a la velocidad de la luz, se denomina (o se llama) la longitud óptica de la curva  $\gamma$ .

Con esta definición podemos enunciar el principio de Fermat como sigue: La trayectoria que sigue un rayo de luz para ir de un punto P a un punto Q es (entre todas las trayectorias posibles) la de longitud óptica mínima (o estacionaria).

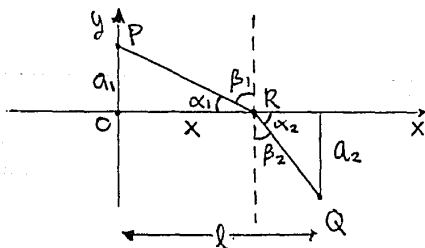
Veamos que el principio de Fermat implica la ley de Snell. Consideremos dos medios ópticos  $M_1$  y  $M_2$  en los que la luz viaja con rapidez  $v_1$  y  $v_2$  respectivamente, y cuya frontera es una recta L. Si P es un punto en  $M_1$  y Q es un punto en  $M_2$  ¿qué trayectoria sigue la luz para ir de P a Q?



Es claro que la trayectoria tiene que estar formada por dos segmentos de recta: uno que va de P a un punto R de L y otro de R a Q.



Tomemos a L como eje x, y como eje y a la recta perpendicular a L que pasa por P.



Podemos escribir el tiempo en recorrer la trayectoria en función de  $x=OR$ :

$T(x) = \sqrt{x^2 + a_{21}^2}/v_1 + \sqrt{a_{22}^2 + (l-x)^2}/v_2$ , entonces

$$T'(x) = \frac{x}{\sqrt{x^2 + a_{21}^2}} \cdot \frac{1}{v_1} - \frac{(l-x)}{\sqrt{(l-x)^2 + a_{22}^2}} \cdot \frac{1}{v_2}$$

de modo que  $T'(x)=0$  implica que  $\text{sen}\beta_1/v_1 = \text{sen}\beta_2/v_2$  ó  $n_1 \text{sen}\beta_1 = n_2 \text{sen}\beta_2$  (con  $n_i = c/v_i$ ,  $i=1,2$ ) donde  $\beta_1$  y  $\beta_2$  son los ángulos que forma el rayo con la normal a L. (Hay ocasiones en que conviene expresarlo como  $\text{cos}\alpha_1/v_1 = \text{cos}\alpha_2/v_2$  donde  $\alpha_1$  y  $\alpha_2$  son los ángulos que el rayo forma con L).

Supongamos que el plano está dividido por bandas horizontales --paralelas al eje x-- en cada una de las cuales la velocidad de la luz es constante





cumple la ecuación  $\text{sen}\beta \vartheta = K$  (También puede escribirse en términos del índice de refracción:  $n(y)\text{sen}\beta(y) = k$ ; o el ángulo  $\alpha$  que forma la tangente a la curva con el eje  $x$ :  $\text{cos}\alpha(y)\vartheta(y) = K$ ).

Ejemplo.

Consideremos las trayectorias de los rayos de luz en la atmósfera terrestre, despreciando la curvatura de la tierra. El índice de refracción es, en este caso, una función de la altura por lo que a lo largo del rayo se tendrá que

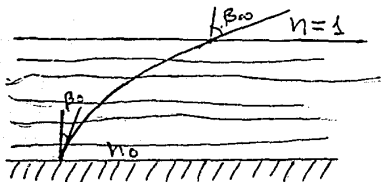
$$n\text{sen}\beta = \text{constante.}$$

De aquí se puede encontrar la desviación de los rayos de luz de una estrella debida a la refracción de la luz en la atmósfera. Si  $\beta_0$  es el ángulo que forma un rayo de luz con la vertical a gran altura, donde  $n=1$ ; entonces para el ángulo  $\beta_0$  correspondiente a la altura del observatorio, se tendrá

$$n_0\text{sen}\beta_0 = \text{sen}\beta_0, \quad (1)$$

donde  $n_0$  es el índice de refracción al nivel del observatorio. Si se miden  $n_0$  y  $\beta_0$  en el observatorio astronómico, la relación (1) nos da el ángulo  $\beta_0$  que es el que tiene interés en la astronomía.

Refracción de los rayos en la atmósfera plana.



Ejemplo.

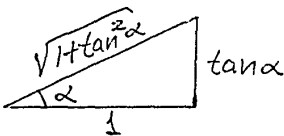
Modelo de Poincaré de la geometría de Lobachevski.

Consideremos al semiplano superior como un medio óptico en el que la velocidad de la luz en cualquier punto es proporcional a su ordenada:  $\vartheta(x, y) = cy$ .

¿Cuáles son las trayectorias que sigue la luz?

Resolvamos el problema usando el hecho de que a lo largo de cada trayectoria se cumple la relación  $\text{cos}\alpha(y)\vartheta(y) = \text{constante}$ .

Para eso primero calculemos  $\text{cos}\alpha(y) = y'$



$$\cos\alpha(y) = \frac{1}{\sqrt{1+\tan^2\alpha(y)}} = \frac{1}{\sqrt{1+y^2}}$$

Como  $\vartheta(y) = cy$ , entonces a lo largo de cada trayectoria se cumple que

$$\frac{1}{\sqrt{1+y^2}} \cdot \frac{1}{cy} = K_1 \quad \text{ó} \quad \frac{1}{ky} = \sqrt{1+y^2}$$

de modo que

$$\frac{dy}{dx} = \frac{\sqrt{1-K^2y^2}}{ky},$$

por lo que separando variables se tiene

$$\frac{ky dy}{\sqrt{1-k^2y^2}} = dx,$$

es decir,  $-\frac{1}{k} \sqrt{1-k^2y^2} = x-h$  ó  $(x-h)^2 + y^2 = \frac{1}{k^2}$

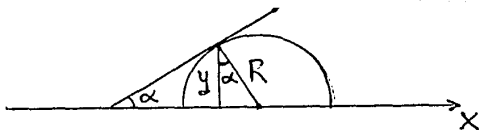
Por lo tanto las trayectorias son semicircunferencias con centro en el eje  $x$  y de radio  $1/K$  para  $K_1 \neq 0$

Si  $K_1 = 0$ , la relación  $\cos\alpha(y)\vartheta(y) = K_1$ , nos dice que las trayectorias son perpendiculares al eje  $x$ .

Otra manera de verlo es:

$\cos\alpha = y/R$ , es decir

$\cos\alpha/y = 1/R$ .



Si se toman por rectas las semicircunferencias con centro en el eje  $x$ ; por longitudes, las longitudes ópticas de los arcos de estas semicircunferencias; y los ángulos que forman estas rectas, son los ángulos que en el punto de su intersección, forman sus tangentes; se obtiene la geometría de Lobachevski, el llamado modelo de Poincaré.

Ejemplo.

Sean  $P$  y  $Q$  dos puntos. ¿Qué forma debe tener una curva  $\gamma$  que una a  $P$  y  $Q$  para que un peso puntual se deslice sobre ella sin fricción bajo la acción de la gravedad, y que partiendo de  $P$  con velocidad cero llegue a  $Q$  en el menor tiempo posible?

La curva que buscamos fue llamada por Johann Bernoulli, *braquistócrona*.



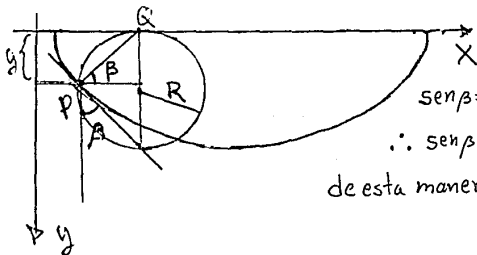
Tomemos por eje  $x$  a la recta horizontal que contiene a  $P$  y esta contenida en el plano vertical que contiene a  $P$  y  $Q$ ; al eje  $y$  lo orientamos verticalmente hacia abajo. Si tomamos (arbitrariamente)  $E=0$  como la energía total en  $P$ , entonces la energía total en cualquier otro punto de la trayectoria también será cero  $E = 1/2 m\dot{\vartheta}^2 - mgy = 0$ , es decir,

$$\dot{\vartheta} = \sqrt{2g\sqrt{y}} \quad (1)$$

Consideremos un medio óptico en el que la velocidad de la luz está dada por (1). La longitud óptica de la curva  $\gamma$ , que une a  $P$  con  $Q$ , coincide con el tiempo de caída a lo largo de esta curva. La trayectoria  $\gamma_0$  que en verdad sigue la luz para ir de  $P$  a  $Q$ , es la curva de longitud óptica mínima entre las curvas que unen a  $P$  con  $Q$ ; por lo tanto  $\gamma_0$  es la *braquistócrona*. Para  $\gamma_0$  se cumple la igualdad

$$\text{sen}\beta \dot{\vartheta} = \text{sen}\beta (\sqrt{2g\sqrt{y}}) = C_1, \text{ o bien } \text{sen}\beta / \sqrt{y} = K, K = C_1 / \sqrt{2g}$$

Pero esto es lo que caracteriza a la cicloide:

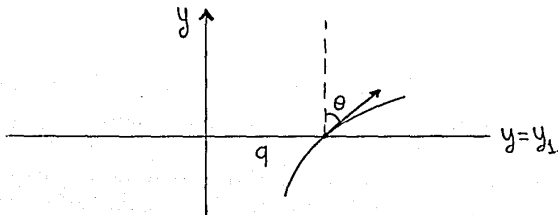


$$\text{sen}\beta = \frac{PQ}{2R} = \frac{y}{2R}$$

$$\therefore \text{sen}\beta = \frac{\sqrt{2R\sqrt{y}}}{2R} = \frac{\sqrt{y}}{\sqrt{2R}}$$

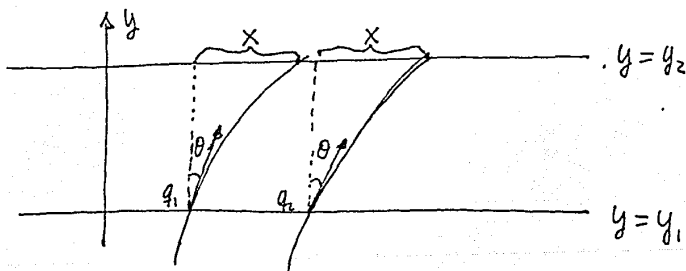
de esta manera  $\frac{\text{sen}\beta}{\sqrt{y}} = K$ , con  $K = \frac{1}{\sqrt{2R}}$

Supongamos que  $\dot{\vartheta} = \dot{\vartheta}(y)$  y consideremos dos rectas paralelas al eje  $x$ :  $y=y_1$  e  $y=y_2$ . En cada una de estas rectas podemos especificar cada rayo por el punto en que cruza a la recta, cuya distancia al eje  $y$  es  $q$ , y por el ángulo que forma el rayo con la vertical (eje  $y$ ) en el punto de intersección con la recta.



Como hicimos anteriormente, no vamos a tomar el ángulo como coordenada sino que ahora vamos a tomar  $p = n \sin \theta$ .

Lo que interesa ahora es conocer como es la transformación que nos dice cuales son las coordenadas  $(q', p')$ , en la recta  $y_2$ , de un rayo que incidió en la recta  $y_1$ , con coordenadas  $(q, p)$ . Para esto, observemos que, como la velocidad de la luz sólo depende de  $y$ , entonces la trayectoria de un rayo que cruza  $y_1$  con un ángulo  $\theta$  y altura  $q_1$  es idéntica a la trayectoria de un rayo que cruza  $y_2$  con el mismo ángulo  $\theta$  y altura  $q_2$ .



Es decir, la forma de las trayectorias no depende de  $q$ , sólo depende de  $p$ ; por lo tanto

$$q' = q + x$$

donde  $x$  es una función solamente de  $p$ , es decir,

$$q' = q + f(p)$$

Además, lo que caracteriza a las trayectorias de la luz en un medio óptico donde  $\vartheta = \vartheta(y)$ , es que a lo largo de cada trayectoria  $n \sin$  permanece constante, es decir,  $p' = p$ . Por lo tanto,

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = T \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} q + f(p) \\ p \end{pmatrix},$$

por lo cual

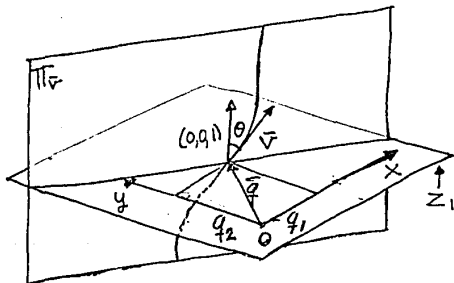
$$D T_{(q,p)} = \begin{pmatrix} 1 & \frac{\partial f}{\partial p} \\ 0 & 1 \end{pmatrix}$$

es decir, la transformación es simpléctica.

Ahora consideremos un medio óptico, en el espacio tridimensional, en el que la velocidad de la luz sólo depende de la altura  $\vartheta = \vartheta(z)$ , por ejemplo la atmósfera.

Dado un plano fijo  $z=z_1$ , podemos especificar cada rayo que lo atraviesa por sus coordenadas  $q_1=x$ ,  $q_2=y$ , y por un vector unitario  $\vec{v}=(v_x, v_y, v_z)$  que indica la dirección del rayo. Como  $\vec{v}$  es unitario, entonces  $v_z$  queda determinado por  $v_x$  y  $v_y$ :  $v_z=(1-v_x^2-v_y^2)^{1/2}$ . Por lo tanto la dirección del rayo se determina por el vector  $(v_x, v_y)$ . No vamos a tomar el vector  $(v_x, v_y)$  como coordenada, sino que vamos a tomar  $\vec{p}=(nv_x, nv_y)$ .

Dados dos planos fijos  $z_1, z_2$ , lo que interesa ahora es conocer como es la transformación que nos dice cuales son las coordenadas  $(\vec{q}', \vec{p}')$  en el plano  $z_2$  de un rayo que incidió en el plano  $z_1$ , con coordenadas  $(\vec{q}, \vec{p})$ .



Dado un rayo que incide en  $z_1$  con coordenadas  $(\vec{q}, \vec{p})$ , éste permanece en el plano  $\Pi\vec{v}$  que es paralelo al eje  $z$  y que contiene a  $\vec{v}$ . Este plano determina una recta en el plano  $z=z_1$ , cuya pendiente es  $m=v_y/v_x$  si  $v_x \neq 0$  (si  $v_x=0$  se toma  $m'=v_z/v_y$ ) y también determina una recta en el plano  $z=z_2$  con la misma pendiente  $m=v_y/v_x$ ; por lo tanto  $v_y=mv_x$  y  $v_y'=mv_x'$ . Si denotamos con  $\theta$  el ángulo formado por  $\vec{v}$  con el eje  $z$ , entonces

$$\text{sen } \theta = (1 - \cos^2 \theta)^{1/2} = (1 - (\vec{v} \cdot (0, 0, 1))^2)^{1/2} = (1 - v_z^2)^{1/2} = (v_x^2 + v_y^2)^{1/2}$$

por lo cual  $n \text{sen } \theta = ((nv_x)^2 + (nv_y)^2)^{1/2} = n \|\vec{p}\| = p$

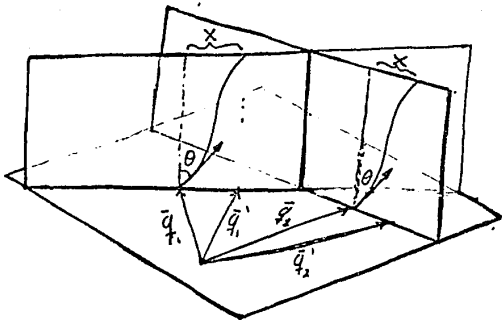
Como  $p = n \text{sen } \theta$  permanece constante a lo largo de cada trayectoria, entonces

$$n^2(v_x^2 + v_y^2) = n'^2(v_x'^2 + v_y'^2)$$

por lo que

$$n^2 v_z^2 (1 + m^2) = n'^2 v_z'^2 (1 + m^2)$$

Por lo tanto  $nv_x = n'v_x'$  y  $nv_y = n'v_y'$ , es decir,  $\vec{p}' = \vec{p}$ . Como la velocidad de la luz sólo depende de  $z$ , entonces la trayectoria de un rayo de luz que cruza  $z_1$  formando un ángulo  $\theta$  con la normal al plano en el punto  $q_1$ , es idéntica a la trayectoria que sigue un rayo que cruza  $z_2$ , formando el mismo ángulo  $\theta$  con la normal al plano en el punto  $q_2$ .



Es decir, la forma de las trayectorias, no depende de  $\bar{q}$ , sólo depende de  $\|\bar{p}\|$ , por lo tanto

$$x = |\bar{q}' - \bar{q}|$$

es una función solamente de  $p$ ; además el vector  $\bar{q}' - \bar{q}$  tiene la dirección de  $\bar{p}$ , por lo cual podemos escribir

$$\bar{q}' = \bar{q} + x\bar{p}/\|\bar{p}\|,$$

es decir,

$$\bar{q}' = \bar{q} + f(p)\bar{p}$$

o

$$\begin{pmatrix} q'_1 \\ q'_2 \end{pmatrix} = \begin{pmatrix} q_1 + f(p)P_1 \\ q_2 + f(p)P_2 \end{pmatrix}.$$

Por lo tanto

$$DT_{(\bar{q}, \bar{p})} = \begin{pmatrix} 1 & 0 & \frac{\partial}{\partial P_1}(f(p)P_1) & \frac{P_1 f'(p)P_2}{P} \\ 0 & 1 & \frac{P_2 f'(p)P_1}{P} & \frac{\partial(f(p)P_2)}{\partial P_2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

es decir,  $DT_{(\bar{q}, \bar{p})} = \begin{pmatrix} \mathbf{I} & \mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$

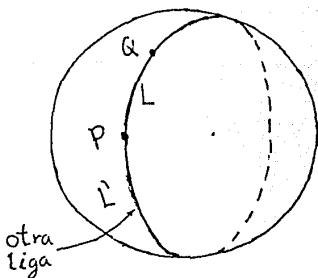
por lo tanto la transformación es simpléctica.

con  $\mathbf{Q}$  simétrica;

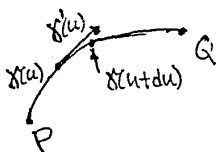
Los ejemplos anteriores son versiones diferentes del problema de encontrar geodésicas; otro ejemplo es cuando se tiene una esfera sólida y una liga, cuyos extremos están fijos sobre dos puntos P y Q de la superficie de la esfera.

El sistema es estacionario --la liga esta en reposo-- si la energía potencial es mínima (o estacionaria). (ver figura)

En ambos casos L y L', tenemos una geodésica. L permanece estable si la perturbamos en cualquier dirección, L' es estable sólo en la dirección de L'; L es una geodésica de longitud mínima y L' no. Sin embargo, si dividimos L' en pedacitos lo suficientemente pequeños, éstos sí son geodésicas de longitud mínima.



Regresando a la óptica supongamos que en  $\mathbb{R}^2$  la velocidad de la luz varía punto a punto, y consideremos una trayectoria  $\gamma: [a, b] \rightarrow \mathbb{R}^2$ ,  $\gamma(a) = P$  y  $\gamma(b) = Q$



$$r(u) = (x_1(u), x_2(u))$$

¿Cuál es el tiempo que tarda una partícula, que se mueve a la velocidad de la luz, para ir de P a Q por la trayectoria  $\gamma$ ?, es decir, ¿cuál es la longitud óptica de  $\gamma$ ?

Estamos suponiendo que el medio es isótropo, es decir, la velocidad de la luz varía de punto a punto, pero para cada punto, la velocidad de la luz es la misma en todas direcciones, y está dada por  $v(x, y)$ .

Si convenimos en tomar la velocidad de la luz en el vacío como  $c=1$ , entonces  $n(x, y) = 1/v(x, y)$ .

Para un  $du$  pequeño de tiene que  $|\gamma(u+du) - \gamma(u)|$  aproximadamente igual  $|\gamma'(u)|du$ , de modo que el tiempo para recorrer este pedacito es de  $dT = |\gamma'(u)|du/v(\gamma(u)) = n(\gamma(u))|\gamma'(u)|du$ , por lo que la longitud óptica de  $\gamma$  es

$$T(\gamma) = \int_a^b n(\gamma(u))|\gamma'(u)|du$$

Por el principio de Fermat, la trayectoria  $\gamma$  que recorre un rayo de luz debe ser tal que su longitud óptica  $T(\gamma)$  sea mínima (o estacionaria), entre todas las curvas que unen a  $P$  con  $Q$ ; es decir, buscamos la  $\gamma$  que minimice a  $T(\gamma)$ .

En la integral aparece  $|\gamma'(u)|$  que es la norma de un vector, y  $n(\gamma(u))$  que es un número. Otra forma de ver la integral anterior, es pensando que en cada punto  $(x,y)$  medimos vectores en forma distinta. Si  $|\xi|$  es la norma euclidiana del vector  $\xi$ , entonces cada punto  $(x,y)$  definimos una nueva norma dada por

$$|\xi|_{(x,y)} = n(x,y) |\xi| = \sqrt{n^2(x,y)\xi_1^2 + n^2(x,y)\xi_2^2},$$

de modo que podemos escribir

$$T(\gamma) = \int_a^b |\gamma'(u)|_{\gamma(u)} du,$$

donde  $|\gamma(u)$  es la norma que corresponde al punto  $\gamma(u)$ . Ahora podemos considerar el caso anisótropo, que es en el que la velocidad de la luz varía también para cada dirección; osea que  $\vartheta = \vartheta(P, \xi_p)$  donde  $\xi_p$  es un vector unitario en  $P$ , y asumimos que  $\vartheta(P, \xi_p) = \vartheta(P, -\xi_p)$  entonces  $n(P, \xi_p) = 1/\vartheta(P, \xi_p)$ , así el tiempo de recorrido para ir de  $P$  a  $Q$  (la longitud óptica de  $\gamma$ ) está dado por

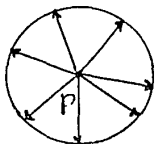
$$T(\gamma) = \int_a^b n(\gamma(u), \frac{\gamma'(u)}{|\gamma'(u)|}) |\gamma'(u)| du.$$

Si definimos  $n(\gamma, \gamma') = \frac{|\gamma'|}{\vartheta(\gamma, \gamma'/|\gamma'|)}$ ,

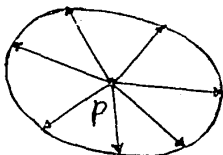
entonces  $T(\gamma) = \int_a^b n(\gamma(u), \gamma'(u)) du$ .

Otra manera de ver esto es pensar nuevamente que a cada vector  $\xi$  en  $P$  le asociamos una norma  $|\xi|_P$  que depende de  $P$  y de la dirección del vector  $\xi$ :

$$|\xi|_P = n(P, \xi)$$



vectores unitarios en la norma usual



vectores unitarios en la nueva norma



La hipótesis que hacemos es que  $|\xi|_p$  es una norma en  $\mathbb{R}^2$  dada por un producto interior, es decir, para cada punto P hay un producto interior

$$\xi, \eta \rightarrow \langle \xi, \eta \rangle_p$$

que varía suavemente de punto a punto. Así,  $|\xi|_p = \langle \xi, \xi \rangle_p^{1/2}$ , por lo cual la longitud óptica de una curva  $\gamma$  es

$$T(\gamma) = \int_a^b |\gamma'(u)|_{\gamma(u)} du = \int_a^b \langle \gamma'(u), \gamma'(u) \rangle_{\gamma(u)}^{1/2} du$$

Observemos que si  $\xi = (\xi_1, \xi_2)$  y  $\eta = (\eta_1, \eta_2)$  entonces

$$\langle \xi, \eta \rangle_p = a\xi_1\eta_1 + b\xi_1\eta_2 + b\xi_2\eta_1 + c\xi_2\eta_2 \text{ donde}$$

$$a = \langle e_1, e_1 \rangle_x, \quad b = \langle e_1, e_2 \rangle_x = \langle e_2, e_1 \rangle_x \quad \text{y} \quad c = \langle e_2, e_2 \rangle_x;$$

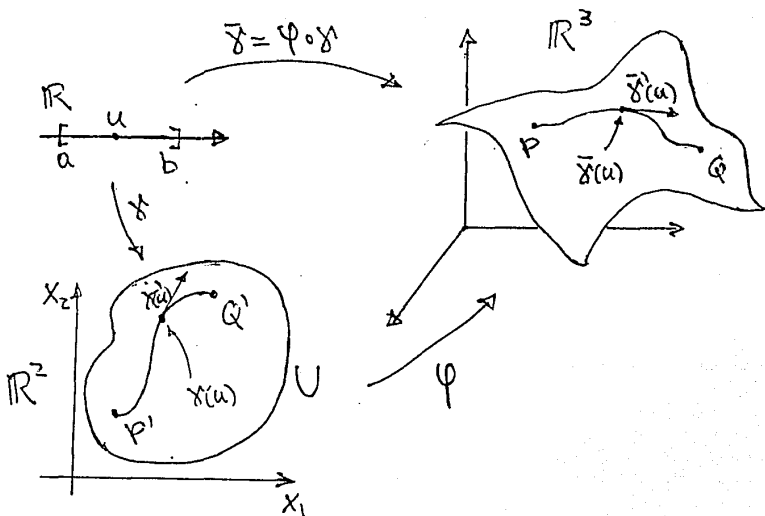
$$\langle \xi, \eta \rangle_{(x_1, x_2)} = a(x_1, x_2)\xi_1\eta_1 + b(x_1, x_2)\xi_1\eta_2 + b(x_1, x_2)\xi_2\eta_1 + c(x_1, x_2)\xi_2\eta_2 \quad \text{por tanto}$$

$$|\xi|_{(x_1, x_2)} = \sqrt{a(x_1, x_2)\xi_1^2 + 2b(x_1, x_2)\xi_1\xi_2 + c(x_1, x_2)\xi_2^2}$$

En los casos anteriores se dice que se tiene una métrica riemanniana.

Podemos plantear el problema de encontrar el camino mínimo que una dos puntos P y Q en una superficie S. Sea  $\varphi(P')=P$  y  $\varphi(Q')=Q$ .

Consideremos una curva parametrizada  $\bar{\gamma}: [a, b] \rightarrow S$ , con  $\bar{\gamma}(a)=P$  y  $\bar{\gamma}(b)=Q$ , es decir,  $\bar{\gamma}$  es la composición de  $\varphi$  con una curva parametrizada  $\gamma: [a, b] \rightarrow U$ , con  $\gamma(a)=P'$  y  $\gamma(b)=Q'$



La longitud de  $\bar{\gamma}$  está dada por  $\int_a^b |\bar{\gamma}'(u)| du = L(\bar{\gamma})$ . Ahora podemos medir las curvas en  $U$  de un modo distinto: a cada curva  $\gamma$  en  $U$  podemos asignarle la longitud de  $\bar{\gamma} = \varphi \circ \gamma$  en  $S$ ; para esto observemos que  $\bar{\gamma}(u) = \varphi \circ \gamma(u)$  y por tanto

$$\bar{\gamma}'(u) = \frac{\partial \varphi_i}{\partial x_1}(\gamma(u)) \gamma'_1(u) + \frac{\partial \varphi_i}{\partial x_2}(\gamma(u)) \gamma'_2(u) = \frac{\partial \varphi_i}{\partial \bar{x}} \cdot \gamma'(u),$$

donde  $\frac{\partial \varphi_i}{\partial \bar{x}} = \left( \frac{\partial \varphi_i}{\partial x_1}, \frac{\partial \varphi_i}{\partial x_2} \right)$ ; entonces

$$L(\gamma) = L(\bar{\gamma}) = \int_a^b |\gamma'(u)| du$$

$$= \int_a^b \left[ \left( \frac{\partial \varphi_1}{\partial \bar{x}} \cdot \gamma'_1(u) \right)^2 + \left( \frac{\partial \varphi_2}{\partial \bar{x}} \cdot \gamma'_1(u) \right)^2 + \left( \frac{\partial \varphi_3}{\partial \bar{x}} \cdot \gamma'_1(u) \right)^2 \right]^{1/2}$$

$$= \int_a^b \left[ \left( \frac{\partial \varphi}{\partial x_1} \cdot \frac{\partial \varphi}{\partial x_1} \right) (\gamma'_1(u))^2 + 2 \left( \frac{\partial \varphi}{\partial x_1} \cdot \frac{\partial \varphi}{\partial x_2} \right) \gamma'_1(u) \gamma'_2(u) + \left( \frac{\partial \varphi}{\partial x_2} \cdot \frac{\partial \varphi}{\partial x_2} \right) (\gamma'_2(u))^2 \right]^{1/2}$$

$$\left\{ \text{donde } \frac{\partial \varphi}{\partial x_i} = \left( \frac{\partial \varphi_1}{\partial x_i}, \frac{\partial \varphi_2}{\partial x_i}, \frac{\partial \varphi_3}{\partial x_i} \right) \right\}$$

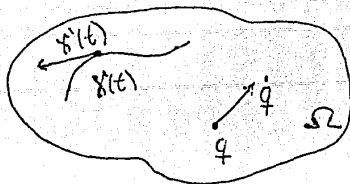
$$= \int_a^b \left( g_{11} [\gamma'_1(u)]^2 + 2g_{12} \gamma'_1(u) \gamma'_2(u) + g_{22} [\gamma'_2(u)]^2 \right)^{1/2}$$

$$= \int_a^b \sqrt{G(\gamma(u)) \gamma'(u) \cdot \gamma'(u)} du$$

$$\text{con } G(\gamma(u)) = \begin{pmatrix} g_{11}(\gamma(u)) & g_{12}(\gamma(u)) \\ g_{21}(\gamma(u)) & g_{22}(\gamma(u)) \end{pmatrix}, \text{ donde } g_{ij}(\gamma(u)) = \left( \frac{\partial \varphi}{\partial x_i} \cdot \frac{\partial \varphi}{\partial x_j} \right) \Big|_{\gamma(u)}$$

Podemos decir que  $\mathcal{L}(\gamma)$  en la métrica riemanniana  $G$  sobre  $U$ , y entonces la curva  $\tilde{\gamma}$  más corta en  $S$  me da una geodésica  $\gamma$  en  $U$  con la métrica riemanniana  $G$  en  $U$  inducida por  $S$ .

Por último sea  $\Omega$  un abierto en  $R^n$  (o una variedad), y supongamos que  $\Omega$  es un medio óptico no-homogeneo y anisótropo



Para especificar la velocidad de la luz en cada punto  $q \in \Omega$ , tomemos un vector unitario  $q$ ,  $|q|=1$ .

Decimos que la velocidad de la luz en el punto  $q$  vale  $\vartheta(q, q) = \vartheta(q, -q) \in R^+$ , de este modo si  $\gamma(t)$  es una trayectoria de la luz, entonces

$$|\gamma'(t)| = \vartheta(\gamma(t), \gamma'(t)/|\gamma'(t)|) \quad (1)$$

Las trayectorias de los fotones deben satisfacer la relación (1); aunque esto no es suficiente para determinarlas, ya que cualquier curva regular  $\gamma(t)$  se puede parametrizar de modo que cumpla con (1).

Si una partícula se mueve por una trayectoria a la velocidad de la luz, con  $\gamma: [a, b] \rightarrow \Omega$ ,  $\gamma(a)=P$  y  $\gamma(b)=Q$ , entonces el tiempo que tarda la partícula de ir de  $P$  a  $Q$  (la longitud óptica de  $\gamma$ ) es

$$T(\gamma) = \int_a^b \frac{|\gamma'(u)| \, du}{\vartheta(\gamma(u), \gamma'(u)/|\gamma'(u)|)}$$

Para encontrar la trayectoria de la luz hay que minimizar  $T(\gamma)$ .

Definamos  $n(q, q) = |q|/\vartheta(q, q/|q|)$  y observemos que

a)  $n(q, q) > 0$  si  $q \neq 0$ ;

b)  $n(q, \lambda q) = |\lambda|n(q, q)$ .

$n$  como función de  $q$  "parece" ser una "norma"; esta sería la norma con la que la luz mide distancias en cada punto  $q$  por donde pasa.

Escribimos  $|q|_q = n(q, q)$  y hacemos la hipótesis de que  $|q|_q$  es una norma definida por un producto interior  $\langle \cdot, \cdot \rangle_q$  que depende (continua y diferenciablemente) de  $q$ . (vamos a suponer  $|q|_q$  es una forma cuadrática).

Con todo esto se tiene que

$$T(\gamma) = \int_a^b |\dot{\gamma}'(u)|_{h(u)} du.$$

Con un producto interior así, se tiene una métrica riemanniana,

$$|q|_q = \langle q, q \rangle_q = (\sum g_{ij}(q) q^i q^j)^{1/2} = (G(q) q \cdot q)^{1/2}$$

donde  $G(q)$  es una matriz simétrica.

Si  $\gamma(t)$  es en verdad una trayectoria de la luz, entonces

$$|\dot{\gamma}'(t)| = \vartheta(\gamma(t), \dot{\gamma}'(t)/|\dot{\gamma}'(t)|),$$

es decir,  $|\dot{\gamma}'(t)|_{h(t)} = 1$ .

Sea  $\Omega$  un abierto en  $\mathbb{R}^n$  (o una variedad), si para cada  $q \in \Omega$  está definido un producto interior  $\langle \cdot, \cdot \rangle_q$  en  $\mathbb{R}^n$  que depende (continua y diferenciablemente) de  $q$  (en una variedad  $\Omega$  el producto interior  $\langle \cdot, \cdot \rangle_q$  está definido en el espacio tangente a  $\Omega$  en  $q$ ,

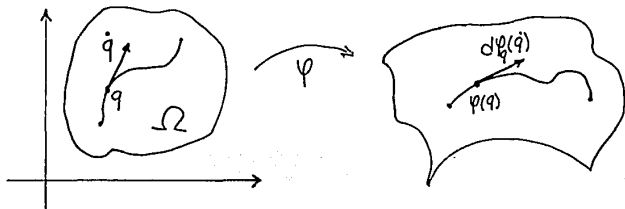
$$\langle \cdot, \cdot \rangle_q : T_q \Omega \times T_q \Omega \rightarrow \mathbb{R}.$$

El producto interior

$$\langle \xi, \eta \rangle_q = \sum g_{ij}(q) \xi_i \eta_j$$

puede provenir de varios casos:

- un medio óptico no-homogeneo e isótropo;
- un medio óptico homogéneo anisótropo;
- un medio óptico no-homogeneo anisótropo;
- de una parametrización de una variedad. (ver página 39)



e) puede ser la energía cinética de una partícula en un cierto sistema mecánico  $T = 1/2 m|\dot{x}|^2$ ;

f) etcétera;

En todos los casos buscamos una  $\gamma$  tal que

$$T(\gamma) = \int_a^b \langle \dot{\gamma}'(u), \dot{\gamma}'(u) \rangle_{h(u)} du$$

sea mínima. En cualquier caso, queremos sea mínimo (o estacionario) el tiempo, la longitud (en mecánica la acción). En todo caso tenemos un problema variacional; lo que sigue ahora es estudiar un poco de cálculo variacional.

### Capítulo III Cálculo Variacional

Consideremos un ejemplo simple.

Sea  $F=F(x, \dot{x}, t)$  una función real diferenciable con  $x \in \mathbb{R}$ ,  $\dot{x} \in \mathbb{R}$  y  $t \in [a, b]$ .

Dada  $f \in C^1_{[a,b]}$ , vamos a considerar la siguiente "funcional" que va a depender de  $f$ :

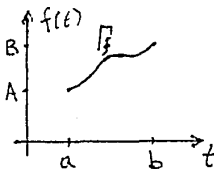
$$\Phi(f) = \int_a^b F(f(t), \dot{f}(t), t) dt$$

Hay que encontrar la  $f$  que haga que  $\Phi(f)$  sea mínima.

Ejemplo.

$$F(x, \dot{x}, t) = \sqrt{1 + \dot{x}^2}$$

Calculando la integral  $\Phi(f) = \int_a^b \sqrt{1 + \dot{f}^2(t)} dt$  se obtiene la longitud de la gráfica de  $f$  además se pueden dar las condiciones iniciales  $f(a) = A$  y  $f(b) = B$



Pero, ¿Qué es la funcional  $\Phi$ ?

El cálculo variacional está relacionado con el estudio de los extremales de funciones cuyo dominio es un espacio de dimensión infinita: el espacio de las curvas (o funciones). Tales funciones se llaman funcionales; de modo que por una funcional, se entiende una correspondencia que asigna un número real a cada función o curva perteneciente a cierta clase; es decir, una funcional es un tipo de función cuya variable independiente es una función o una curva.

Sea  $\Phi: C^1_{[a,b]} \rightarrow \mathbb{R}$  una funcional.

Busco la funcional lineal que se parezca más a

$$\Phi(f+h) - \Phi(f), \text{ con } h \in C^1_{[a,b]}$$

entonces escribo

$$\Phi(f+h) - \Phi(f) = T(h, f) + R(h, f),$$

donde  $T(h, f)$  es lineal en  $h$  y  $R(h, f) < c|h|^2$ , con  $c = \text{constante}$ ; es decir,  $R(h, f)$  tiende a cero más rápido que  $h$ .

Si existen  $T$  y  $R$  así, diremos que la funcional  $\Phi$  es diferenciable y que "T es la derivada de  $\Phi$  en el punto  $f$ ". También diremos que  $f$  es un "punto estacionario" para  $\Phi$  si la derivada de  $\Phi$  en  $f$  es cero, es decir, la funcional lineal cero:  $T(h, f) = 0 \forall h$ .

**Teorema:**

Sea  $F(x, \dot{x}, t)$  una función diferenciable de tres variables:  $x, \dot{x} \in \mathbb{R}, t \in [a, b]$ . Si definimos la funcional

$$\Phi(f) = \int_a^b F(f(t), \dot{f}(t), t) dt \text{ con } f \in C^1([a, b]),$$

entonces  $\Phi$  es diferenciable y

$$D\Phi_f(h) = \int_a^b \left( \frac{\partial F}{\partial x}(f(t), \dot{f}(t), t) h(t) + \frac{\partial F}{\partial \dot{x}}(f(t), \dot{f}(t), t) \dot{h}(t) \right) dt.$$

**Demostración.**

$$\Phi(f+h) - \Phi(f) = \int_a^b [F(f(t)+h(t), \dot{f}(t)+\dot{h}(t), t) - F(f(t), \dot{f}(t), t)] dt$$

$$\text{pero } F(x+h, \dot{x}+\dot{h}, t) - F(x, \dot{x}, t) = \frac{\partial F}{\partial x}(x, \dot{x}, t) h + \frac{\partial F}{\partial \dot{x}}(x, \dot{x}, t) \dot{h} + R(h, \dot{h})$$

donde  $R$  es tal que  $|R(h, \dot{h})| < K(|h|^2 + |\dot{h}|^2)$ , con  $k = \text{constante}$ .

De modo que la integral es

$$\int_a^b \left[ \frac{\partial F}{\partial x}(f(t), \dot{f}(t), t) h(t) + \frac{\partial F}{\partial \dot{x}}(f(t), \dot{f}(t), t) \dot{h}(t) \right] dt + \tilde{R}(h, \dot{h})$$

donde  $\tilde{R}$  es tal que  $|\tilde{R}(h, \dot{h})| < (b-a)K(|h|^2 + |\dot{h}|^2) \leq (b-a)K\|h\|_1^2$

La derivada de  $\Phi$  en  $f$  es la función lineal en  $h$

$$T(h, f) = \int_a^b \left[ \frac{\partial F}{\partial x}(f(t), \dot{f}(t), t) h(t) + \frac{\partial F}{\partial \dot{x}}(f(t), \dot{f}(t), t) \dot{h}(t) \right] dt = D\Phi_f(h)$$

Vamos a calcular esto de otra manera, haciendo la observación de que la integral

$$\int_a^b \frac{\partial F}{\partial \dot{x}}(f(t), \dot{f}(t), t) \dot{h}(t) dt$$

se puede calcular por partes:

$$\text{hagamos } u = \frac{\partial F}{\partial \dot{x}}(f(t), \dot{f}(t), t) \quad \text{y } dv = \dot{h}(t) dt$$

$$\text{así, } \int_a^b \frac{\partial F}{\partial \dot{x}}(f(t), \dot{f}(t), t) \dot{h}(t) dt = \frac{\partial F}{\partial \dot{x}}(f(t), \dot{f}(t), t) h(t) \Big|_a^b - \int_a^b h(t) \frac{d}{dt} \left( \frac{\partial F}{\partial \dot{x}}(f(t), \dot{f}(t), t) \right) dt,$$

por lo tanto

$$D\Phi_f(h) = \int_a^b \left[ \frac{\partial F}{\partial x}(f(t), \dot{f}(t), t) - \frac{d}{dt} \left( \frac{\partial F}{\partial \dot{x}}(f(t), \dot{f}(t), t) \right) \right] h(t) dt - \frac{\partial F}{\partial \dot{x}}(f, \dot{f}, t) h(t) \Big|_a^b$$

Vamos a considerar el caso en que  $f(a) = A$  y  $f(b) = B$ .

**Teorema:**

$f$  es un valor estacionario de la funcional

$$\Phi(f) = \int_a^b F(t, f(t), \dot{f}(t), t) dt \text{ (sujeta a que } f(a) = A \text{ y } f(b) = B)$$

si y solo si  $f$  satisface la ecuación de Euler-Lagrange:

$$\frac{\partial F}{\partial x}(t, f(t), \dot{f}(t), t) - \frac{d}{dt} \left( \frac{\partial F}{\partial \dot{x}}(t, f(t), \dot{f}(t), t) \right) = 0 \quad \forall t \in [a, b]$$

La demostración del teorema se basa en el siguiente lema.

**Lema:**

Si  $\alpha(t)$  es continua en  $[a, b]$  y  $\int_a^b \alpha(t)h(t)dt = 0$  para toda  $h \in C^1_{[a,b]}$  tal que  $h(a)=h(b)=0$ , entonces  $\alpha(t)=0$ .

**Demostración del lema:**

Supongamos que  $\alpha(t)$  es distinta de cero, digamos positiva, en algun punto  $t_0$ ,  $a < t_0 < b$ . Como  $\alpha$  es continua, entonces  $\alpha(t) > c > 0$  en una vecindad de  $t_0$ :  $(t_0-d, t_0+d)$ , contenida en  $[a, b]$ . Sea  $h(t) = 0$  fuera de  $(t_0-d, t_0+d)$ ,  $h(t) > 0$  en  $(t_0-d, t_0+d)$  y  $h(t) = 1$  en  $(t_0-d/2, t_0+d/2)$ . Entonces

$$\int_a^b \alpha(t)h(t)dt \geq dc > 0,$$

lo cual contradice la hipótesis del lema. Por tanto  $\alpha(t_0) = 0$  para todo  $t_0 \in [a, b]$   $\diamond$

Nota.

La construcción de  $h(t)$  se puede hacer tomando

$$g(x) = \begin{cases} e^{-x^2} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

luego  $g_1(x) = g(x-d/2)g(d-x)$ , despues

$$g_2(x) = \frac{\int_{-\infty}^x g_1(t)dt}{\int_{-\infty}^{\infty} g_1(t)dt}$$

finalmente  $h(t) = g_2(3/2d - |t - t_0|)$ .

**Demostración del teorema:**

Por el teorema anterior

$$D\Phi_f(h) = \int_a^b \left[ \frac{\partial F}{\partial x}(t, f, \dot{f}, t) - \frac{d}{dt} \left( \frac{\partial F}{\partial \dot{x}}(t, f, \dot{f}, t) \right) \right] h(t) dt - \frac{\partial F}{\partial x}(t, f, \dot{f}, t) h(t) \Big|_a^b$$

pero  $\frac{\partial F}{\partial x}(t, f, \dot{f}, t) h(t) \Big|_a^b = 0$  ya que  $h(a) = h(b) = 0$ .

Si  $f$  es punto estacionario de  $\Phi$ , entonces  $D\Phi_f(h) = 0$  para toda  $h \in C^1_{[a,b]}$  con  $h(a) = h(b) = 0$ ; es decir,

$$\int_a^b \alpha(t)h(t)dt = 0 \quad \forall h \in C^1_{[a,b]} \text{ con } h(a) = h(b) = 0, \text{ donde}$$



$$\alpha(t) = \frac{\partial F}{\partial x}(t, \dot{x}, t) - \frac{d}{dt} \left( \frac{\partial F}{\partial \dot{x}}(t, \dot{x}, t) \right),$$

Por el lema anterior se tiene que  $\alpha(t) \equiv 0$ .

Recíprocamente, si  $\alpha(t) \equiv 0$ , entonces

$$D\Phi_f(h) = \int_a^b \alpha(t) h_{\dot{x}}(t) dt = 0 \quad \forall h \in C^{\infty}_{[a,b]}$$

Nota.

Otra manera de escribir la ecuación de Euler-Lagrange (E-L) es:

$$\frac{\partial F}{\partial x}(t, \dot{x}, t) - \frac{\partial^2 F}{\partial x \partial \dot{x}}(t, \dot{x}, t) \dot{x}(t) - \frac{\partial^2 F}{\partial \dot{x}^2}(t, \dot{x}, t) \ddot{x}(t) - \frac{\partial F}{\partial t}(t, \dot{x}, t) = 0$$

que es una ecuación diferencial de segundo orden cuya solución depende de dos constantes arbitrarias que son determinadas por las condiciones de frontera  $f(a)=A$  y  $f(b)=B$ . El problema que se considera usualmente en la teoría de ecuaciones diferenciales es encontrar una solución definida en la vecindad de algún punto que satisfaga condiciones iniciales dadas. Sin embargo al resolver la ecuación de (E-L) buscamos una solución definida sobre una región fija y que satisfaga condiciones a la frontera dadas. Por lo que la cuestión de saber si un problema variacional tiene solución, no se reduce a los teoremas usuales de existencia para ecuaciones diferenciales. Sin embargo, la existencia de un extremo frecuentemente es clara del significado físico o geométrico del problema a resolver.

Ejemplo.

Sea  $\Phi: C^1_{[a,b]} \rightarrow \mathbb{R}$  definida por  $\Phi(f) = \text{longitud de } \Gamma_f =$

$$\int_a^b \sqrt{1 + \dot{f}^2(t)} dt.$$

$$F(x, \dot{x}, t) = \sqrt{1 + \dot{x}^2}; \quad \frac{\partial F}{\partial x} = 0; \quad \frac{\partial F}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}}$$

(E-L) es:

por lo tanto la ecuación de

$$-\frac{d}{dt} \left( \frac{\dot{f}(t)}{\sqrt{1 + \dot{f}^2(t)}} \right) \text{ si y sólo si } \frac{\dot{f}(t)}{\sqrt{1 + \dot{f}^2(t)}} = C_1$$

si y sólo si  $f(t) = C$

si y sólo si  $f(t) = Ct + D$

es decir, se trata de rectas, fijando  $f(a)=A$  y  $f(b)=B$  obtenemos  $C$  y  $D$ .

Ejemplo.

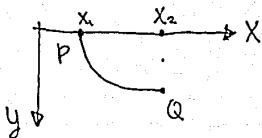
$$\Phi(f) = \int_a^b (1 + \dot{f}^2(t)) dt. \quad F(x, \dot{x}, t) = 1 + \dot{x}^2, \text{ por lo tanto } \frac{\partial F}{\partial \dot{x}} = 2\dot{x}$$

$$\text{y la ecuación de (E-L) queda } -\frac{d}{dt} (2\dot{f}(t)) = 0$$

es decir, si  $f(t)=C$ , por lo que  $f(t) = Ct + D$ . Nuevamente rectas.

Ejemplo.

La braquistócrona



$$1/2 m \dot{\theta}^2 = mgy, \text{ por lo que } \dot{\theta} = \sqrt{2g\sqrt{y}}.$$

$$dt = \frac{ds}{v} = \frac{\sqrt{1+y'^2} dx}{\sqrt{2g}\sqrt{y}}, \text{ por lo que el}$$

$$\text{tiempo de caída es } T = \frac{1}{\sqrt{2g}} \int_{x_1}^{x_2} \frac{\sqrt{1+y'^2}}{\sqrt{y}} dx$$

$$\text{Entonces } F(y, y', x) = \frac{\sqrt{1+y'^2}}{\sqrt{y}}; F_y = -\frac{\sqrt{1+y'^2}}{2y^{3/2}}; F_{y'} = \frac{y'}{\sqrt{y}\sqrt{1+y'^2}}$$

$$\text{por lo que } \frac{d}{dt} F_{y'} = \frac{2yy'' - (1+y'^2)y'^2}{2[y(1+y'^2)]^{3/2}} = 0,$$

$$\text{Así, la ecuación (E-L) queda: } F_y - \frac{d}{dt} F_{y'} = \frac{-(1+y'^2)^2 - 2yy'' + (1+y'^2)y'^2}{2y^{3/2}(1+y'^2)^{3/2}} = 0$$

por lo que  $-y'^2 - 1 - 2yy'' = 0$  ó  $2y''/(1+y'^2) = -1/y$ ; como  $y' = 0$  no es solución podemos escribir  $2y'y''/(1+y'^2) = -y'/y$

por lo cual  $\ln(1+y'^2) = \ln(c/y)$ , con  $c > 0$ ;

de este modo  $1+y'^2 = c/y$  por lo que  $\sqrt{1+y'^2} = \sqrt{c}/\sqrt{y}$ ,

es decir  $1/\sqrt{(1+y'^2)\sqrt{y}} = k$ . Pero  $1/\sqrt{(1+y'^2)} = \text{sen}\beta$ ,

donde  $\beta$  es el ángulo que forma la tangente con el eje  $y$ . Por lo tanto podemos escribir  $\text{sen}\beta/\sqrt{y} = K$ , que es lo que caracteriza la cicloide.

Ejemplo.

Sean  $(x_1, y_1), (x_2, y_2)$  fijos con  $y_1, y_2 > 0$

Sea  $E = \{f: [x_1, x_2] \rightarrow \mathbb{R} / f \text{ es diferenciable y } f(x_1) = y_1, f(x_2) = y_2\}$

Encontrar  $f \in E$  tal que el área de la superficie de revolución que se obtiene al girar la gráfica de  $f$  al rededor del eje  $x$  sea mínima. (A esta superficie se le llama catenoide, y a la curva que la genera se le llama catenaria)

$$A = 2\pi \int_{x_1}^{x_2} y\sqrt{1+y'^2} dx. \text{ Por lo tanto } F(y, y', x) = y\sqrt{1+y'^2};$$

$$F_y = \sqrt{1+y'^2}; F_{y'} = yy'/\sqrt{1+y'^2}; \frac{d}{dt} F_{y'} = \frac{(1+y'^2)(yy'' + y'^2) - yy'^2 y''}{(1+y'^2)^{3/2}}$$

Así, la ecuación de (E-L) queda:

$$F_y - \frac{d}{dt} F_{y'} = \frac{(1+y'^2) - (1+y'^2)(yy'' + y'^2) + yy'^2 y''}{(1+y'^2)^{3/2}}$$

de modo que  $1+y'^2 - yy'' = 0$  ó  $y''/(1+y'^2) = 1/y$ ; como  $y' = 0$  no da una solución, podemos escribir  $1/2 (2y'y''/(1+y'^2)) = y'/y$ , es decir  $\ln(1+y'^2)^{1/2} = \ln(cy)$ , con  $c > 0$ ;

es decir,  $\sqrt{(1+y'^2)} = cy$  (1)

o  $y' = \sqrt{(c^2 y^2 - 1)}$ ; separando variables obtenemos  $dx = dy/\sqrt{(c^2 y^2 - 1)}$ , es decir,  $x+k = (1/c)\ln|cy + \sqrt{(c^2 y^2 - 1)}|$ ;

por lo tanto,  $y = (1/c)\cosh(c(x+k))$ .

Antes de pasar al último ejemplo observaremos que la ecuación (1) se puede escribir como

$$1/(\sqrt{(1+y'^2)}(k/y)) = c_1, \text{ es decir, } \cos\alpha/(k/y) = c_1$$

Por lo tanto, una catenaria es la trayectoria que sigue un rayo de luz en un medio óptico, donde la velocidad de la luz es inversamente proporcional a la altura:  $v = K/y$ .

Ejemplo.

Sean  $(x_1, y_1)$ ,  $(x_2, y_2)$  y  $E$  como el ejemplo anterior.

Encontrar  $f \in E$  tal que

$$\int_{x_1}^{x_2} \frac{\sqrt{1+f'^2(x)}}{f(x)} dx \quad \text{sea mínima}$$

$$F(y, y', x) = \sqrt{(1+y'^2)}/y; \quad F_y = -\sqrt{(1+y'^2)}/y^2; \quad F_{y'} = y'/(y\sqrt{(1+y'^2)});$$

$$\frac{d}{dt} F_{y'} = \frac{y(1+y'^2)y'' - y y'^2 y'' - (1+y'^2)y'^2}{y^2(1+y'^2)^{3/2}}; \quad \text{de modo que la ecuación de (E-L) queda:}$$

por lo tanto  $1+y''+yy''=0$  ó  $y''/(1+y'^2) = -1/y$ ; como  $y'=0$  no da una solución, podemos escribir

$$(1/2)(2y'y''/(1+y'^2)) = -y'/y, \text{ es decir, } \ln(1+y'^2)^{1/2} = \ln(c/y), \text{ con } c>0, \text{ es decir, } (1+y'^2)^{1/2} = c/y.$$

Esta última ecuación ya la resolvimos en la página 32, y obtuvimos arcos de círculo con centro en el eje  $x$ , que corresponden a las trayectorias de la luz en el plano de Poincare.

Ahora generalizemos lo anterior, pero no con funciones sino con curvas. Sea  $F(x_1, \dots, x_n, \dot{x}_1, \dots, \dot{x}_n, t)$  una función de clase  $C^1$  de  $2n+1$  variables, con  $x_i, \dot{x}_i \in R$  y  $t \in [a, b]$ ; sea  $\gamma: [a, b] \rightarrow R^n$  una curva.

Consideremos la funcional

$$\Phi(\gamma) = \int_a^b F(\gamma_1(t), \dots, \gamma_n(t), \dot{\gamma}_1(t), \dots, \dot{\gamma}_n(t), t) dt.$$

Queremos calcular los puntos estacionarios  $\Phi$ ; consideremos las posibles variaciones de  $\gamma: \gamma(t)+h(t)$ .

Como  $D\Phi_\gamma(h)$  es lineal en  $h$  escribimos

$$D\Phi_\gamma(h) = D\Phi_\gamma(h_1, \dots, h_n) = D\Phi_\gamma(h_1, 0, \dots, 0) + \dots + D\Phi_\gamma(0, \dots, 0, h_n) \\ = [(F_{x_1} - \frac{d}{dt} F_{\dot{x}_1})h_1 + \dots + (F_{x_n} - \frac{d}{dt} F_{\dot{x}_n})h_n] dt$$

Esto nos conduce al siguiente teorema.

**Teorema:**

$\gamma$  es estacionaria para  $\Phi$  (sujeta a  $\gamma(a)=P$  y  $\gamma(b)=Q$ ) si y sólo si

$$\frac{\partial F}{\partial x_i}(\delta(t), \dot{\delta}(t), t) - \frac{d}{dt} \left( \frac{\partial F}{\partial \dot{x}_i}(\delta(t), \dot{\delta}(t), t) \right) = 0, \quad i=1, \dots, n.$$

Ejemplo.

(geogésica en  $\mathbb{R}$ ) Sean  $P$  y  $Q$  en  $\mathbb{R}$ , y  $x: [a, b] \rightarrow \mathbb{R}$ , con  $x(a)=P$  y  $x(b)=Q$ . Entonces  $F(x, \dot{x}) = |x|$ .

$$\partial F / \partial x = 0; \quad \partial F / \partial \dot{x} = \begin{cases} 1 & \text{si } \dot{x} > 0 \\ -1 & \text{si } \dot{x} < 0 \end{cases};$$

si pedimos que  $\dot{x} \neq 0$ , entonces  $\partial F / \partial \dot{x} = \begin{cases} 1 & \text{si } P < Q \\ -1 & \text{si } P > Q \end{cases},$

de cualquier forma se tiene que la ecuación de (E-L)

$$dF_2/dt = 0$$

es válida para cualquier curva  $x(t)$  con tal de que  $\dot{x}(t) \neq 0$

Ejemplo.

Sea  $F(x, \dot{x}, t) = x^2$  y sea  $x: [a, b] \rightarrow \mathbb{R}$  tal que  $x(a) = P$  y  $x(b) = Q$

$F_{\dot{x}} = 0$ ;  $F_t = 2x$  y  $dF_2/dt = d2x/dt = 0$ ; por lo que  $x=c$ , de modo que  $x=ct+d$ ;  $c$  y  $d$  quedan determinadas por  $P$  y  $Q$ , es decir,  $x(t)$  es única.

Ejemplo.

Consideremos dos puntos arbitrarios  $P$  y  $Q$  en  $\mathbb{R}^2$ .

¿Cuál es la curva más corta que une a  $P$  con  $Q$ ?



Sea  $\gamma: [a, b] \rightarrow \mathbb{R}^2$ ,  $\gamma(a) = P$ ,  $\gamma(b) = Q$ .

Entonces,  $L(\gamma) = \int_a^b |\dot{\gamma}(t)| dt = \int_a^b \sqrt{\dot{\gamma}_1^2(t) + \dot{\gamma}_2^2(t)} dt$ ,

por lo tanto,  $F(x_1, x_2, \dot{x}_1, \dot{x}_2, t) = \sqrt{\dot{x}_1^2 + \dot{x}_2^2}$ .

$\partial F / \partial x_i = 0$ ;  $\partial F / \partial \dot{x}_i = \dot{x}_i / \sqrt{\dot{x}_1^2 + \dot{x}_2^2}$ ;

$\partial F / \partial \dot{x}_2 = \dot{x}_2 / \sqrt{\dot{x}_1^2 + \dot{x}_2^2}$ . Entonces las ecuaciones de (E-L) son

$$\frac{d}{dt} \left( \frac{\dot{x}_i(t)}{\sqrt{\dot{x}_1^2(t) + \dot{x}_2^2(t)}} \right) = 0, \quad i = 1, \dots$$

Es decir,  $\dot{\gamma}_1(t) / \sqrt{\dot{\gamma}_1^2(t) + \dot{\gamma}_2^2(t)} = c_1$  y  $\dot{\gamma}_2(t) / \sqrt{\dot{\gamma}_1^2(t) + \dot{\gamma}_2^2(t)} = c_2$

Tomando el cociente de estas ecuaciones se tiene

$(\dot{\gamma}_1(t) / \dot{\gamma}_2(t) = C)$  si y sólo si  $(\gamma_1(t) = C\gamma_2(t))$  si y sólo si  $(\gamma_1(t) = C\gamma_2(t) + D)$ .

Entonces hay muchas soluciones, ya que se puede dar  $\gamma_2$  "arbitraria" y encontrar  $\gamma_1$ . Todas estas soluciones están sobre la recta  $x_1 = Cx_2 + D$ , donde  $C$  y  $D$  quedan determinadas por  $P$  y  $Q$ . Si pedimos que  $\gamma(t) \neq (0, 0)$  tenemos que la solución es el segmento  $PQ$  pero con distintas parametrizaciones (esto es natural porque la longitud de una curva, no depende de su parametrización). Si de la ecuación diferencial no se puede despejar explícitamente  $\ddot{q}$ , entonces no pueden aplicarse los teoremas de existencia y unicidad; y decimos que no

es una "verdadera ecuación diferencial de segundo orden" puede suceder que tengamos demasiadas soluciones.

Sea  $\gamma$  una curva,  $\gamma: [a, b] \rightarrow \mathbb{R}^n$ .

Definimos la energía de la curva  $\gamma$  como

$$E(\gamma) = \int_a^b |\dot{\gamma}(t)|^2 dt$$

Veamos otro problema: sean P y Q dos puntos arbitrarios en  $\mathbb{R}^2$ . Entre las curvas que unen a P con Q (¿cuál es la que minimiza la energía?)

$F(x_1, x_2, \dot{x}_1, \dot{x}_2, t) = \dot{x}_1^2 + \dot{x}_2^2$ ;  $\partial F / \partial x_i = 0$ . Por lo tanto, las ecuaciones de (E-L) son:

$$\frac{d}{dt} (2\dot{\gamma}_1(t)) = 0 \quad \text{y} \quad \frac{d}{dt} (2\dot{\gamma}_2(t)) = 0, \quad \text{es decir}$$

$$\begin{aligned} \dot{\gamma}_1(t) &= c_1 & \text{y} & & \dot{\gamma}_2(t) &= c_2, & \text{por lo que} \\ \gamma_1(t) &= c_1 t + d_1 & \text{y} & & \gamma_2(t) &= c_2 t + d_2, \end{aligned}$$

donde  $c_1, c_2, d_1, d_2$  quedan determinadas por P y Q.

De estos dos ejemplos se ve que no sólo es importante minimizar la longitud,

$$L(\gamma) = \int_a^b |\dot{\gamma}(t)| dt,$$

sino minimizar la energía

$$E(\gamma) = \int_a^b |\dot{\gamma}(t)|^2 dt;$$

la energía establece como debe ir parametrizada la curva. La ecuación (E,L) para las curvas estacionarias de la energía si es una "verdadera" ecuación diferencial de segundo orden (aquí si vamos a poder aplicar los teoremas de existencia y unicidad).

Veamos estos ejemplos en  $\mathbb{R}^2$ ; la ecuación (E-L) para  $F(q, \dot{q}, t)$  es

$$\begin{aligned} \frac{\partial F}{\partial q} - \frac{d}{dt} \left( \frac{\partial F}{\partial \dot{q}} \right) &= 0, \\ \frac{\partial F}{\partial q} &= \left( \frac{\partial F}{\partial q_1}, \dots, \frac{\partial F}{\partial q_n} \right), \\ \frac{\partial F}{\partial \dot{q}} &= \left( \frac{\partial F}{\partial \dot{q}_1}, \dots, \frac{\partial F}{\partial \dot{q}_n} \right). \end{aligned}$$

donde

$L(\gamma) = \int_a^b |\dot{\gamma}(u)| du$ . Entonces  $F(q, \dot{q}, t) = |\dot{q}| = (\dot{q} \cdot \dot{q})^{1/2}$  por lo que  $\frac{\partial F}{\partial q} = 0$  y  $\frac{\partial F}{\partial \dot{q}} = \frac{\dot{q}}{|\dot{q}|}$ ; de esta forma, la ecuación de (E-L) queda  $\frac{d}{dt} \left( \frac{\dot{q}}{|\dot{q}|} \right) = 0$ , es decir,  $\dot{q}/|\dot{q}| = \bar{C}$ .

Tenemos entonces rectas parametrizadas a cualquier velocidad; el vector unitario  $\bar{C}$  queda determinado por los puntos P y Q:  $\bar{C} = (Q-P)/|Q-P|$ .

$E(\gamma) = \int_a^b |\dot{\gamma}(u)|^2 du$ . Entonces,  $F(q, \dot{q}, t) = \dot{q} \cdot \dot{q}$ , por lo que  $\frac{\partial F}{\partial q} = 0$  y  $\frac{\partial F}{\partial \dot{q}} = 2\dot{q}$

de este modo  $\frac{d}{dt}(2\dot{q}) = 0$  si y solo si  $\dot{q} = \bar{C}_0$  si y solo si  $q = \bar{C}_0 t + \bar{C}_1$ . Obtenemos rectas parametrizadas a velocidad constante; hay unicidad.

Por lo que se ve, si reparametrizamos una curva, no cambia su longitud; sin embargo, si cambia su energía.

Consideremos un abierto  $\Omega \subset \mathbb{R}^n$  (o una variedad) donde esta definida una métrica riemanniana  $G$ :

$$G(q, \dot{q}) = |\dot{q}|_q^2 = \langle \dot{q}, \dot{q} \rangle_q = \sum (1/2) g_{ij}(q) \dot{q}_i \dot{q}_j = (1/2) g(q) \dot{q} \cdot \dot{q}.$$

Vamos a considerar dos problemas variacionales:

$$L(\gamma) = \int \gamma'(u) |_{\gamma(u)} du \quad y$$

$$E(\gamma) = \int \gamma'(u) |_{\gamma(u)}^2 du$$

según los ejemplos anteriores parece ser que las curvas extremales de  $L$  y de  $E$  son las mismas, pero las de  $E$  son parametrizadas a velocidad constante. Veamos cuales son las extremales de  $L$  y  $E$ , y las comparemos.

Vamos a considerar  $\int_{\gamma} |\dot{q}|_q^2 = \int_{\gamma} G$ ; entonces  $G$  debe satisfacer la (E-L) a lo largo de  $\gamma$ :

$$\frac{\partial G}{\partial q} - \frac{d}{dt} \left( \frac{\partial G}{\partial \dot{q}} \right) = 0;$$

como  $G(q, \dot{q}) = (1/2) g(q) \dot{q} \cdot \dot{q}$ , entonces

$$\frac{\partial G}{\partial q} = \frac{1}{2} \frac{\partial g(q)}{\partial q} \dot{q} \cdot \dot{q} \quad y \quad \frac{\partial G}{\partial \dot{q}} = g(q) \dot{q},$$

de modo que la (E-L) es

$$\frac{1}{2} \frac{\partial g(q)}{\partial q} \dot{q} \cdot \dot{q} - \frac{d}{dt} (g(q) \dot{q}) = 0$$

es decir,

$$\frac{1}{2} \frac{\partial g(q)}{\partial q} \dot{q} \cdot \dot{q} - \frac{\partial g(q)}{\partial q} \dot{q} \cdot \dot{q} - g(q) \ddot{q} = 0$$

de modo que podemos despejar  $\ddot{q}$  (esta si es una "verdadera" ecuación ...):

$$\ddot{q} = \frac{1}{2} \ddot{q}(q) \left( \frac{\partial g(q)}{\partial q} \dot{q} \cdot \dot{q} \right), \quad (0)$$

esta es la ecuación diferencial de las trayectorias extremales de  $E$  en esa métrica riemanniana.

Ahora consideremos los extremales de  $L(\gamma) = \int_{\gamma} G(q, \dot{q})^{1/2}$

$$F(q, \dot{q}) = G(q, \dot{q})^{1/2}, \text{ entonces } \frac{\partial F}{\partial q} = \frac{1}{2G^{1/2}} \frac{\partial G}{\partial q} \quad \text{y} \quad \frac{\partial F}{\partial \dot{q}} = \frac{1}{2G^{1/2}} \frac{\partial G}{\partial \dot{q}}$$

por lo que

$$\frac{\partial F}{\partial q} - \frac{d}{dt} \left( \frac{\partial F}{\partial \dot{q}} \right) = \frac{1}{2G^{1/2}} \frac{\partial G}{\partial q} - \frac{1}{2G^{1/2}} \frac{d}{dt} \left( \frac{\partial G}{\partial \dot{q}} \right) - \frac{d}{dt} \left( \frac{1}{2G^{1/2}} \right) \frac{\partial G}{\partial \dot{q}}$$

por lo tanto la ecuación de (E-L) para L es:

$$\frac{1}{2G^{1/2}} \left( \frac{\partial G}{\partial q} - \frac{d}{dt} \left( \frac{\partial G}{\partial \dot{q}} \right) \right) - \frac{d}{dt} \left( \frac{1}{2G^{1/2}} \right) \frac{\partial G}{\partial \dot{q}} = 0 \quad (1)$$

Si  $\gamma$  es extremal de L y  $\|\gamma'(t)\|_{\text{norm}} = G(\gamma(t), \dot{\gamma}(t)) = \text{constante}$ , entonces  $1/2 G^{1/2}$  es constante a lo largo de  $\gamma$ ; así, de (1) se deduce que  $\gamma$  es extremal de E. Es decir, las extremales de L parametrizadas a velocidad constante son extremales de E. Si demostramos que todas las extremales de E son parametrizadas a velocidad constante entonces de (1) se deduce la siguiente afirmación

**Afirmación.**

Las trayectorias estacionarias para E, son las trayectorias estacionarias de L, a velocidad constante.

(En particular la luz recorre trayectorias estacionarias para E con velocidad  $\|\gamma'(t)\|_{\text{norm}} = 1$ )  
Para demostrar la afirmación observemos que:

$$G(q, \dot{q}) = \|\dot{q}\|^2 = 1/2 g(q) \dot{q} \cdot \dot{q}$$

de modo que

$$\partial G / \partial \dot{q} = g(q) \dot{q}$$

por lo cual podemos escribir

$$G(q, \dot{q}) = (1/2) \dot{q} \partial G / \partial \dot{q} \quad (2)$$

Ahora vamos a ver que toda extremal de E, es parametrizada a velocidad constante. Si  $\gamma$  es estacionaria para E, entonces se satisface la ecuación de (E-L) a lo largo de  $\gamma$

$$\text{y como } \|\gamma'(t)\|_{\text{norm}}^2 = G(\gamma(t), \dot{\gamma}(t)), \text{ entonces } \frac{d}{dt} (\partial G / \partial \dot{q}) = \partial G / \partial q$$

$$\frac{d}{dt} (\|\gamma'(t)\|_{\text{norm}}^2) = \frac{d}{dt} G(\gamma(t), \dot{\gamma}(t))$$

que por (2) :

$$\begin{aligned} \frac{d}{dt} G(q, \dot{q}) &= (1/2) \frac{d}{dt} (\dot{q} \partial G / \partial \dot{q}) \\ &= (1/2) \frac{d}{dt} (\partial G / \partial \dot{q}) \dot{q} + (1/2) (\partial G / \partial \dot{q}) \dot{q} \end{aligned}$$

por (E-L): 
$$= 1/2 (\partial G/\partial q)\dot{q} + 1/2 (\partial G/\partial \dot{q})\dot{q}$$

$$= (1/2) \frac{d}{dt} G(q, \dot{q}) ;$$

o sea que a lo largo de  $\gamma$

$$\frac{d}{dt} G(\gamma(t), \gamma'(t)) = (1/2) \frac{d}{dt} G(\gamma(t), \gamma'(t))$$

Por lo tanto 
$$\frac{d}{dt} G(\gamma(t), \gamma'(t)) = 0,$$

de modo que a lo largo de  $\gamma$   $G(\gamma(t), \gamma'(t))$  es constante, es decir  $\|\gamma'(t)\|_{G_m}$  es constante.

Esto está relacionado con el problema de las ligas sobre una superficie: si se tiene una liga sobre una superficie, con los extremos de la liga fijos, la forma que toma la liga es la de una curva geodésica de la superficie; para esta curva la energía potencial de la liga es mínima (o estacionaria), y no sólo eso, sino que también la energía potencial de la liga es uniforme a lo largo de la curva, es decir, la liga no está más tensa en unos puntos que en otros.

#### Definición.

Decimos que  $\gamma$  es una geodésica para la métrica riemanniana  $G$ , se  $\gamma$  es estacionaria para  $E$ .

Veamos un ejemplo.

Considerese un rayo de luz, cuyo movimiento está restringido a una superficie de revolución en  $\mathbb{R}^3$ . En coordenadas cilíndricas la superficie está dada (localmente) en la forma  $r=r(z)$  o  $z=z(r)$ .

La energía de una curva en  $\mathbb{R}^3$  está dada por

$$E(\gamma) = \int_{\gamma} \frac{1}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

Como la luz se mueve por una superficie de revolución, entonces el lagrangiano es, en las coordenadas  $r$  y  $\varphi$

$$G(q, \dot{q}) = 1/2 (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = 1/2 (\dot{r}^2 + r^2 \dot{\varphi}^2 + z'^2(r) \dot{r}^2)$$

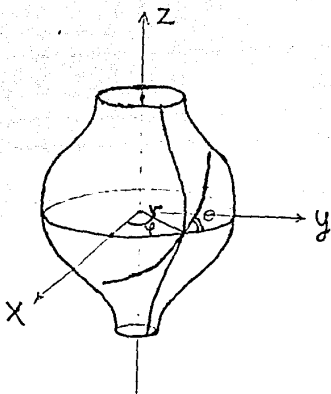
y en las coordenadas  $z$  y  $\varphi$ ,

$$G(q, \dot{q}) = 1/2 (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = 1/2 (\dot{r}^2 + r^2(z) \dot{\varphi}^2 + r'^2(z) \dot{z}^2)$$

En ambos casos el lagrangiano  $G(q, \dot{q})$  no depende de  $\varphi$ , por lo cual  $\frac{d}{dt} \left( \frac{\partial G}{\partial \dot{\varphi}} \right) = 0$ , es decir,  $r^2 \dot{\varphi}$  se conserva.

Denotemos por  $\theta$  el ángulo que forma la trayectoria con un paralelo





Entonces la componente horizontal de la velocidad es  $r\dot{\varphi} = |V\cos\theta|$ , pero  $|V|$  es constante, de modo que la conservación de  $\varphi r^2$  implica que

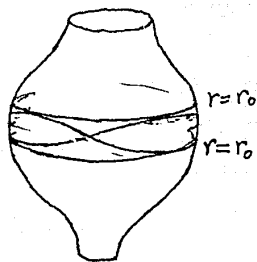
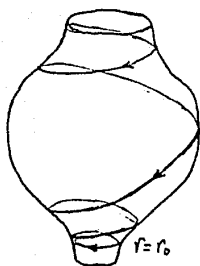
$$r\cos\theta = \text{constante (Teorema de Clairaut)}$$

Toda geodésica es una superficie de revolución, que satisface, la ecuación de Clairaut; sin embargo, no basta que una curva satisfaga dicha ecuación, para que la curva sea una geodésica en la superficie de revolución; por ejemplo, todos los paralelos de una esfera, (o de una superficie de revolución) satisfacen la ecuación de Clairaut, pero sólo el ecuador es geodésica.

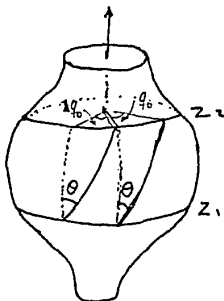
Consideremos ahora una geodésica, que cruza un ecuador de la superficie con un ángulo  $\theta$  pequeño; después de cruzar el ecuador,  $r$  disminuye, de modo que  $\cos\theta$  crece, es decir,  $\theta$  se hace más pequeño ya que la geodésica satisface la ecuación de Clairaut. Si  $r$  sigue disminuyendo, entonces, en algún  $r=r_0$ , se tiene que  $\theta=0$ , y a partir de ahí la geodésica se regresa,  $r$  vuelve a crecer y  $\theta$  también. Esto significa que hay geodésicas (o pueden haber), que están confinadas en una banda:  $r \geq r_0$ . (ver la siguiente figura)

Ahora fijemos dos paralelos:  $z=z_1$ , y  $z=z_2$ ; y fijemos un meridiano:  $\varphi=0$ . Se deja como ejercicio demostrar que los meridianos son geodésicas; entonces  $\varphi=0$  es una geodésica. Considerese las geodésicas cercanas a los meridianos ( $\theta$  cercano a  $\pi/2$ ). Cada rayo (geodésica) puede especificarse en el paralelo  $z=z_1$ , por su coordenada  $q=\varphi$  y por el ángulo  $\alpha$  que forma con la normal al paralelo  $z_1$  (la normal es un meridiano).

Las coordenadas que vamos a tomar para cada rayo son  $q=\varphi$  y  $P=n\sin\alpha$ , donde  $n=r$ ; es decir,  $P=r\cos\theta$ , por lo cual  $P'=P$ .



Es claro que si rotamos una geodésica en la dirección  $\varphi$ , la curva que se obtiene sigue siendo una geodésica, es decir,



la forma de las trayectorias no depende de  $q$ , sólo de  $P$ , por lo tanto  $q' = q + q_0$ , donde  $q_0$  sólo depende de  $P$ , es decir,  $q' = q + f(P)$  de esta forma si  $\begin{pmatrix} q' \\ p' \end{pmatrix} = T \begin{pmatrix} q \\ p \end{pmatrix}$

entonces 
$$DT_{(q,P)} = \begin{pmatrix} 1 & \frac{\partial f(P)}{\partial P} \\ 0 & 1 \end{pmatrix}$$

es decir, la transformación  $T \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} q' \\ p' \end{pmatrix}$  es simpléctica.

Antes de seguir adelante, obsérvese que la relación (0) de la página 52

$$\ddot{q} = \frac{1}{2} g^{-1}(q) \left( \frac{\partial g(q)}{\partial q} \dot{q} \cdot \dot{q} \right) \quad (0)$$

es un sistema de  $n$  ecuaciones diferenciales de segundo orden en un abierto  $\Omega \subset \mathbb{R}^n$ ; cualquier curva suave  $t \rightarrow q(t)$  en  $\Omega$  determina una curva  $t \rightarrow (q(t), \dot{q}(t))$  en  $\Omega \times \mathbb{R}^n$  (si se trata de una variedad en  $T\Omega$ ). Si  $q(t)$  es una geodésica en  $\Omega$ , entonces la curva  $(q(t), \dot{q}(t))$  en  $\Omega \times \mathbb{R}^n$  satisface el sistema de  $2n$  ecuaciones diferenciales de primer orden

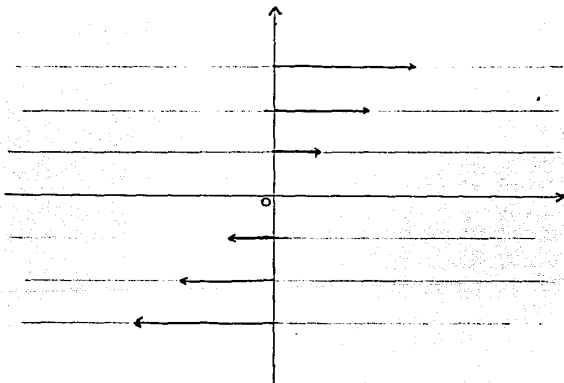
$$\begin{cases} \frac{d}{dt} q = \dot{q} \\ \frac{d}{dt} \dot{q} = \frac{1}{2} g^{-1}(q) \frac{\partial g(q)}{\partial q} \dot{q} \cdot \dot{q} \end{cases} \quad (3)$$

Dados  $q_0 \in \Omega$  y  $q_0' \in \mathbb{R}^n$  existe entonces una única geodésica  $\gamma(t)$  tal que  $\gamma(0) = q_0$  y  $\gamma'(0) = q_0'$ . Los sistemas (0) y (3) son equivalentes. Nosotros vamos a trabajar con el sistema (3), es decir, vamos a trabajar en  $\Omega \times \mathbb{R}^n$  (si se trata de una variedad en  $T\Omega$ ). Este sistema determina un flujo en  $\Omega \times \mathbb{R}^n$ ; a continuación estudiaremos el flujo del toro plano y del toro de revolución.

## Capítulo IV

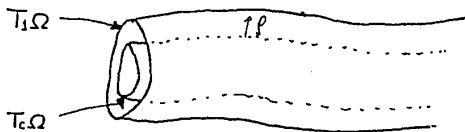
### Descripción del flujo geodésico del toro plano y del toro de revolución

Para comenzar veamos el flujo geodésico de  $\mathbb{R}$ . El flujo esta en  $\mathbb{R} \times \mathbb{R}$ ,



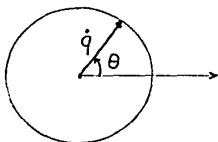
las geodésicas en  $\mathbb{R}$  son las parametrizaciones de las rectas a velocidad constante, la única diferencia entre ellas es la velocidad de recorrido, por lo cual basta con observar la geodésica recorrida a velocidad 1.

Los conjuntos  $T_c\Omega = \{(q, \dot{q}) \in \Omega \times \mathbb{R}^n \mid |\dot{q}|_g = G(q, \dot{q}) = C\}$  con  $C > 0$ , son invariantes por el flujo, ya que  $G$  es una constante a lo largo de cada geodésica. Además, hay una correspondencia natural entre las curvas integrales que están en  $T_cW$  y las que están en  $T_1\Omega$ . Definamos el transformación  $\rho: T_c\Omega \rightarrow T_1\Omega$  mediante  $\rho(q(t), \dot{q}(t)) = (q(t), \dot{q}(t)/c)$ ; entonces  $\varphi^t(q, \dot{q}) = \rho^{-1} \varphi^{t/c} \rho(q, \dot{q}) = \rho^{-1} \varphi^{ct} \rho(q, \dot{q})$ .



Dos curvas integrales que se corresponden —aunque son curvas distintas en  $\Omega \times \mathbb{R}^n$ — determinan la misma geodésica geométrica de  $(\Omega, G)$  pero recorrida a diferentes velocidades constantes. Para entender como es el flujo, nos basta con ver como es en  $T_1\Omega$ . Por esta razón, cuando se hable de flujo geodésico, se entenderá el flujo en  $T_1\Omega$ . Veamos como es el flujo geodésico en  $\mathbb{R}^2$  con la métrica usual.

Como ya sabemos, las geodésicas son las rectas recorridas a velocidad constante. Podemos identificar el conjunto de vectores unitarios en  $\mathbb{R}^2$  con  $S^1$



Así, las coordenadas  $(q_1, q_2, \theta)$  en  $T_1\mathbb{R}^2 = \mathbb{R}^2 \times S^1$  quedan definidas por las ecuaciones :

$$\begin{aligned} q_1 &= \cos\theta \\ q_2 &= \text{sen}\theta \end{aligned}$$

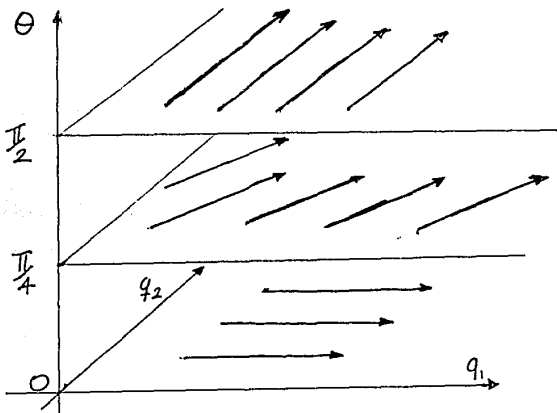
Entonces la ecuación de las rectas esta dada por

$$q_2 = \tan\theta q_1 + C$$

o su parametrización es

$$\begin{aligned} q_1 &= \cos\theta t + C_1 \\ q_2 &= \text{sen}\theta t + C_2; \end{aligned}$$

lo importante es que a lo largo de estas,  $\theta$  es constante (ver la figura siguiente)



Podemos pasar del plano euclidiano  $\mathbb{R}^2$ , al toro plano

$$T^2 = \mathbb{R}^2 / (2\pi\mathbb{Z})^2 \quad ((x,y) \rightarrow (x+2\pi k, y+2\pi l) \text{ con } k, l \in \mathbb{Z})$$

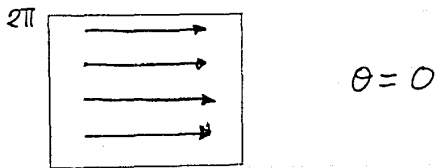
provisto de la métrica euclidiana en  $\mathbb{R}^2$ .

Como  $T^2$  tiene la métrica de  $\mathbb{R}^2$ , las geodésicas localmente son pedacitos de recta, por lo cual las geodésicas en  $T^2$  son las imágenes de las rectas en  $\mathbb{R}^2$  mediante la proyección de  $\Pi: \mathbb{R}^2 \rightarrow T^2$ .

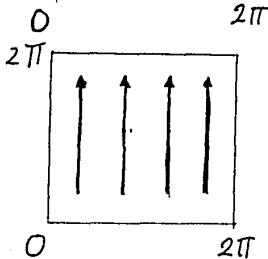
$$T_1 T^2 = T_1(S^1 \times S^1) \{((\varphi, \psi), v) \mid |v|=1\}$$

Pero como podemos identificar los vectores unitarios en  $\mathbb{R}^2$  con  $S^1$ , entonces vamos a considerar el conjunto

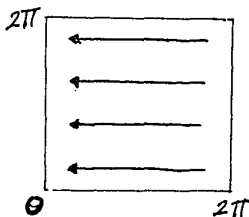
$$\tau_\theta = T_1 T^2_\theta = \{((\varphi, \psi), \theta)\} \text{ con } \theta \text{ fijo}$$



$$\theta = 0$$



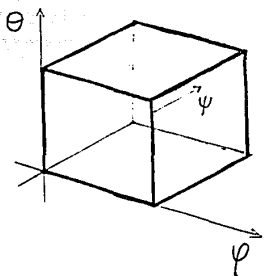
$$\theta = \frac{\pi}{2}$$



$$\theta = \pi$$

El flujo geodésico de  $T^2$  es el flujo sobre

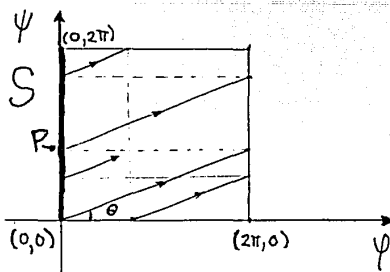
$$T_1(T^2) = T^3 = (R/(2\pi Z))^3 = \{(\varphi, \psi, \theta)\}$$



$T^3$

en este cubo  
identificamos  
las caras opuestas.

Veamos como son las trayectorias para un  $\theta$  fijo.



Partamos de  $(0,0)$  hasta llegar nuevamente al círculo  $\varphi=0$  en el punto  $P=(0,2\pi \tan \theta)$ , (llamemos  $S$  al círculo  $\varphi=0$ ). El conjunto de puntos de intersección de la trayectoria con  $S$ , es una órbita de la rotación del ángulo  $2\pi \tan \theta$ , es decir, el conjunto de puntos de la forma  $(0, \psi_l)$  donde  $\psi_l = k(2\pi \tan \theta) - l2\pi = 2\pi(k \tan \theta - l)$ , con  $k, l \in Z$  y  $(k \tan \theta - l) \in [0, 1)$ . La trayectoria se cierra si y sólo si existen  $k \neq 0, l \neq 0$  tales que  $k \tan \theta - l = 0$ , si y sólo si  $\tan \theta = l/k \in Q$ .

En este caso todas las trayectorias que corresponden a  $\theta$  (aunque no partan de  $(0,0)$ ) son cerradas.

Para  $\theta$  tal que  $\tan \theta \notin Q$ , ninguna trayectoria se cierra.

**Lema.**

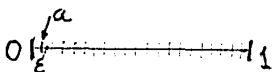
Si  $\lambda \in \Theta$ , el conjunto de puntos

$$D = \{(k\lambda/l) \in [0,1) / k, l \in \mathbb{Z}\}$$

es denso.

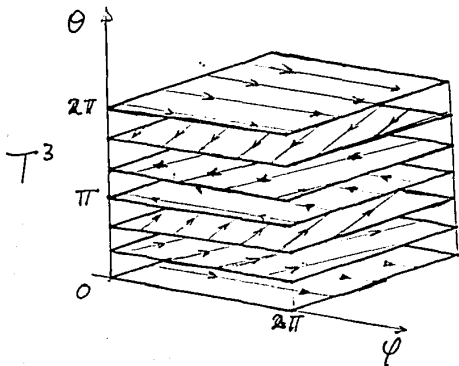
**Demostración.**

Sea  $\varepsilon > 0$ , digo que existen  $a_1, a_2 \in D$ , tales que  $a_2 > a_1$  y  $a_2 - a_1 < \varepsilon$ , ya que, como  $D$  es un conjunto infinito, entonces por Bolzano-Weierstrass, hay un punto de acumulación. Como  $a_1 = k_1\lambda - l_1$  y  $a_2 = k_2\lambda - l_2$  entonces  $a = (a_2 - a_1) \in D$  y por lo tanto  $ma \in D$  para  $m \in \mathbb{N}$  y  $ma < 1$ . De esta forma, para cualquier punto  $x \in [0,1)$ , existe un punto  $m_a \in D$  tal que  $|x - m_a| < \varepsilon$ .



De aquí se desprende que cuando  $\theta$  es tal que  $\tan\theta \notin \Psi \cup \mathbb{Q}$ , las trayectorias no se cierran y cada trayectoria es densa en el cuadrado.

El comportamiento del flujo geodésico de  $T^2$  es similar al flujo geodésico de  $\mathbb{R}^2$ . Los planos de ecuación  $\theta = \theta_0$ , definen toros  $T_{\theta_0} \subset T^3$  invariantes bajo el flujo, sobre los cuales el flujo geodésico  $g^t$  es lineal:  $g^t(\varphi, \psi, \theta_0) = (\varphi + t\cos\theta_0, \psi + t\sin\theta_0)$



La restricción a cada toro  $T_{\theta_0}$  de la proyección,  $(\varphi, \psi, \theta) \rightarrow (\varphi, \psi)$  de  $T^3$  sobre  $T^2$  es un difeomorfismo

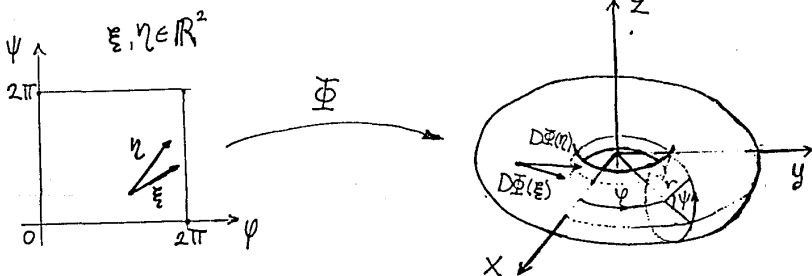
$$(\varphi, \psi, \theta_0) \rightarrow (\varphi, \psi);$$



cada curva integral en  $\mathcal{T}_\theta$  es periódica o densa en  $\mathcal{T}_\theta$  según  $\tan\theta$ , sea racional o irracional; además se puede demostrar que cuando  $\tan\theta$  es irracional cada curva correspondiente a este valor de  $\theta$ , está equidistribuida, es decir, dado un intervalo pequeño de longitud  $\varepsilon$  en el círculo  $\psi = 0$  (o en el círculo  $\varphi=0$ ) la frecuencia con la que pasa la curva por ahí, sólo depende de su longitud  $\varepsilon$ .

Para describir el flujo geodésico del toro de revolución, consideremos la parametrización  $\Phi: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  definida por

$$\Phi(\varphi, \psi) = ((1+r\cos\psi)\cos\varphi, (1+r\cos\psi)\sin\varphi, r\sin\psi) \quad , \quad \text{con } r < 1$$



Se puede obtener una métrica riemanniana en el cuadrado, mediante el producto interior ordinario en  $\mathbb{R}^3$  vía  $\Phi$ . Dados dos vectores  $\xi, \eta$  en  $\mathbb{R}^2$  —basados en  $p$ — mediante la derivada  $D\Phi_p$ , obtenemos los vectores correspondientes en  $\mathbb{R}^3$ :  $D\Phi_p(\xi), D\Phi_p(\eta)$ . Así, se puede definir un producto interior para  $\xi, \eta$  mediante

$$\langle \xi, \eta \rangle_{((p,\psi))} = D\Phi_{(p,\psi)}(\xi) \cdot D\Phi_{(p,\psi)}(\eta)$$

$$D\Phi_{(\varphi,\psi)} = \begin{pmatrix} -(1+r\cos\psi)\sin\varphi & -r\sin\psi\cos\varphi \\ (1+r\cos\psi)\cos\varphi & -r\sin\psi\sin\varphi \\ 0 & r\cos\psi \end{pmatrix}$$

por lo tanto  $\langle e_1, e_1 \rangle_{(\varphi,\psi)} = (1+r\cos\psi)^2$ ;  $\langle e_1, e_2 \rangle_{(\varphi,\psi)} = \langle e_2, e_1 \rangle_{(\varphi,\psi)} = 0$  y  $\langle e_2, e_2 \rangle_{(\varphi,\psi)} = r^2$   
Es así que se obtiene

$$g(\varphi, \psi) = \begin{pmatrix} (1+r\cos\psi)^2 & 0 \\ 0 & r^2 \end{pmatrix},$$

es decir, el lagrangiano esta dado por

$$G(\varphi, \psi, \dot{\varphi}, \dot{\psi}) = \|\dot{\varphi}, \dot{\psi}\|_{(q,v)} = 1/2(1+r\cos\psi)^2\dot{\varphi}^2 + r^2\dot{\psi}^2$$

y las ecuaciones de (E-L) estan dadas por

$$\frac{d}{dt} (1+r\cos\psi)^2\dot{\varphi} = 0$$

$$\frac{d}{dt} (r^2\dot{\psi}) = r\sin\psi(1+r\cos\psi)\dot{\varphi}^2$$

La primera de estas ecuaciones nos dice que hay una constante de movimiento

$$(1+r\cos\psi)^2\dot{\varphi} = \text{constante}$$

ésta es resultado de la simetría de rotación alrededor del eje z. En el toro plano, también hay una constante de movimiento ( $\theta = \text{constante}$ ), por lo cual el problema se redujo a dos dimensiones. es decir, cada curva integral se mueve en algo de dos dimensiones (los toros invariantes  $\mathcal{T}_{\theta_0}$ ).

En el caso del toro de revolución, la constante de movimiento ( $(1+r\cos\psi)^2\dot{\varphi} = \text{constante}$ ) también nos reduce el problema a dos dimensiones.

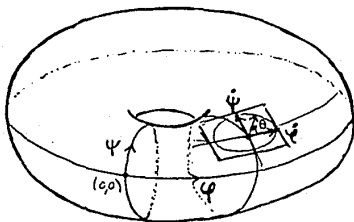
Vamos a considerar el conjunto

$$S_1 = \{(\varphi, \psi, \dot{\varphi}, \dot{\psi}) / \|\dot{\varphi}, \dot{\psi}\|_{(q,v)} = 1\}$$

Sea  $\theta = \angle \{(\dot{\varphi}, \dot{\psi}), (1, 0)\} = \angle \{\dot{q}, e_1\}$  (ver la siguiente figura)

$$\cos\theta = \frac{\langle (\dot{\varphi}, \dot{\psi}), (1, 0) \rangle_{(q,v)}}{\|(\dot{\varphi}, \dot{\psi})\|_{(q,v)} \|(1, 0)\|} = \frac{(1+r\cos\psi)^2\dot{\varphi}}{1 \cdot (1+r\cos\psi)} = (1+r\cos\psi)\dot{\varphi}$$

$$y \quad \sin\theta = \frac{\langle (\dot{\varphi}, \dot{\psi}), (0, 1) \rangle_{(q,v)}}{\|(\dot{\varphi}, \dot{\psi})\|_{(q,v)} \|(0, 1)\|} = \frac{r^2\dot{\psi}}{1 \cdot r} = r\dot{\psi}$$



Cada punto de  $S_1$ , está descrito por las coordenadas  $(\varphi, \psi, \theta)$ , de modo que podemos describir el flujo geodésico del toro de revolución en  $T^3 = \{(\varphi, \psi, \theta)\}$ ; por lo tanto podemos escribir, nuestra constante de movimiento de otra manera

$$(1+r\cos\psi)\cos\theta = \text{constante};$$

esta es la integral de Clairaut (teorema de Clairaut), para las geodésicas de una superficie de revolución. (Esta ecuación podíamos haberla obtenido directamente, como se vió al final del capítulo III, ya que se trata de una superficie de revolución).

Si se deriva la ecuación de Clairaut respecto a  $t$  y se despeja  $\dot{\theta}$ , entonces ya se tienen las ecuaciones que definen el flujo geodésico:

$$d\varphi/dt = \cos\theta/(1+r\cos\psi)$$

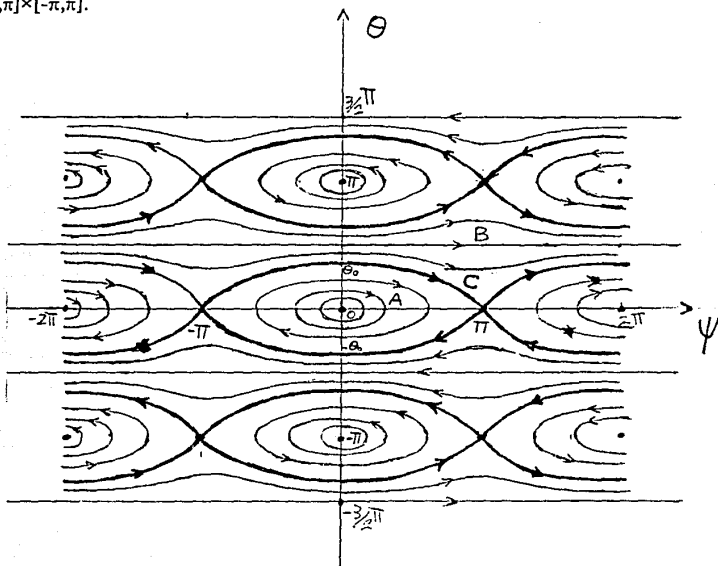
$$d\psi/dt = \text{sen}\theta/r$$

$$d\theta/dt = \cos\theta\text{sen}\psi/(1+r\cos\psi)$$

Estas ecuaciones no dependen de  $\varphi$ , y la integral de Clairaut tampoco; entonces las superficies de nivel de la función  $C(\psi, \theta) = 1+r\cos\psi\cos\theta$

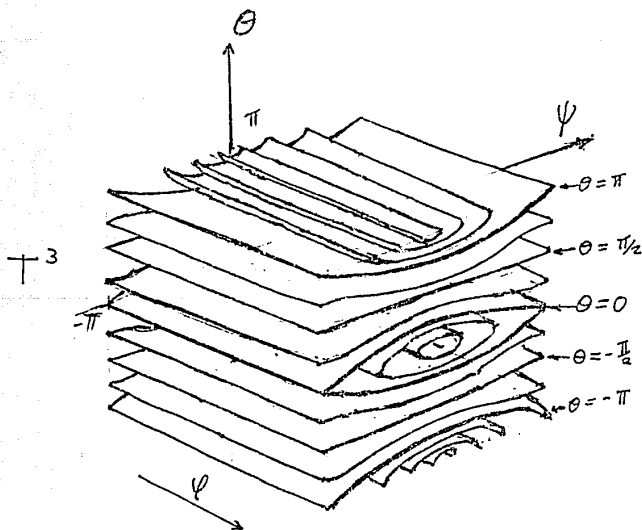
no dependen de  $\varphi$  lo cual quiere decir que las curvas de nivel de  $C(\psi, \theta)$  en cada plano (toro plano)  $\varphi = \varphi_0$ , son las mismas para toda  $\varphi_0$ .

Veamos como son las curvas de nivel de la integral de Clairaut en el plano  $(\psi, \theta)$ , en  $[-\pi, \pi] \times [-\pi, \pi]$ .



De las ecuaciones de flujo geodésico, se puede ver que el paralelo máximo ( $\psi=0$ ;  $\theta=0$  ó  $\theta=\pi$ ) y el paralelo mínimo ( $\psi=\pi$ ;  $\theta=0$  ó  $\theta=\pi$ ) son geodésicas.

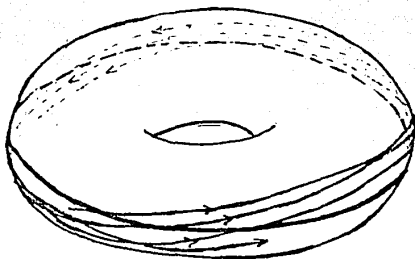
Si para cada curva de nivel de la integral de Clairaut, se considera su producto cruz con el círculo de las  $\varphi$ , se obtienen toros invariantes bajo el flujo geodésico, estos toros son uniones de curvas integrales y son superficies de nivel de la integral de Clairaut en  $T^3$ .



*Toros invariantes*

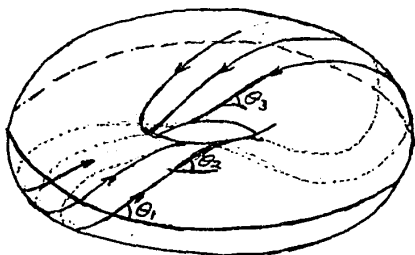
Consideremos las curvas de nivel que encierran al  $(0,0)$  —en la figura están marcadas con la letra A— y sus toros invariantes correspondientes; Las geodésicas del toro de revolución correspondientes a estos toros invariantes —a las que llamaremos geodésicas de tipo A— nunca tocan un paralelo mínimo y además, una geodésica con condiciones iniciales cercanas al paralelo máximo permanece a él todo el tiempo. (ver figura siguiente)

Se puede decir lo mismo de las geodésicas que corresponden a las curvas de nivel, que encierran al  $(0, \pi)$ ; de hecho, son las mismas geodésicas solo que recorridas en sentido contrario.



geodésica  
de tipo A.

Consideremos ahora los toros invariantes correspondientes a las curvas de nivel que no encierran al  $(0,0)$  ni al  $(0, \pi)$  y que en la figura están marcadas con la letra B. Todas las geodésicas del toro de revolución correspondientes a estos toros invariantes —a las que llamaremos geodésicas de tipo B— se enrollan alrededor del toro de revolución.



geodésica  
de tipo B

$$\theta_3 < \theta_2 < \theta_1$$

Observemos que  $\dot{\psi}$  y  $\dot{\theta}$  se anulan simultáneamente cuando  $\psi = K\pi$  y  $\theta = l\pi$ , con  $K, l \in \mathbb{Z}$ ; pero nos basta con considerar  $\psi = 0, \pi$  y  $\theta = 0, \pi$ . Por ninguno de los cuatro puntos  $(0,0)$ ,  $(0,\pi)$ ,  $(\pi,0)$ ,  $(\pi,\pi)$ , pasa ninguna curva de nivel A, B o C; de hecho, cada uno de estos puntos es un conjunto de nivel de la integral de Clairaut; y se dice que estos puntos son soluciones estacionarias o puntos de equilibrio del sistema

$$\begin{aligned} \dot{\psi} &= \text{sen}\theta/r \\ \dot{\theta} &= \cos\theta \text{sen}\psi / (1 + r\cos\psi) \end{aligned}$$

en el plano  $(\psi, \theta)$  —o en el toro plano  $(\psi, \theta)$ —. Al hacer el producto cruz de estos puntos con el círculo de las  $\varphi$ , se obtienen cuatro curvas integrales, de las cuales dos corresponden

al paralelo máximo en los dos sentidos de recorrido respectivamente; y las otras dos corresponden al paralelo mínimo en los dos sentidos de recorrido respectivamente.

En los puntos de equilibrio, que corresponden al paralelo máximo, hay estabilidad en el sentido de que las geodésicas con condiciones cercanas al paralelo máximo —o las curvas de nivel condiciones iniciales cercanas a  $(0,0)$  o al  $(0, \pi)$ — permanecen cercanas a él, —respectivamente permanecen cercanas al  $(0,0)$  o al  $(0, \pi)$ —. Sin embargo en los puntos de equilibrio que corresponden al paralelo mínimo, no hay la estabilidad de la que hablamos, ya que cualquier geodésica con condiciones iniciales cercanas al paralelo mínimo se aleja mucho de él; de hecho, todas las geodésicas, excepto el paralelo mínimo, tocan al paralelo máximo.

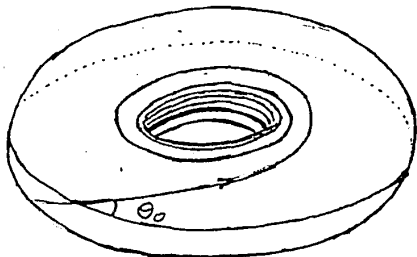
Ahora consideremos las curvas de nivel marcadas en la figura con la letra C, estas curvas son las que pasan por los puntos  $(0, \pm\theta_0)$  y  $(0, \pi \pm \theta_0)$ . Estas curvas no contienen a sus extremos, que son los puntos  $(\pi, 0)$  y  $(\pi, -\pi)$  correspondientes al paralelo mínimo; por lo tanto, al hacer el producto cruz de estas curvas con el círculo de las  $\varphi$ , no se obtiene toros invariantes sino cilindros invariantes. Las geodésicas correspondientes a estos cilindros invariantes son las que hacen la transición entre las geodésicas de tipo A y las de tipo B —a estas las llamaremos geodésicas de tipo C—. Cada geodésica de tipo C se enrolla por encima del paralelo mínimo, acercándose indefinidamente a él, pero sin tocarlo nunca cuando se recorre en sentido contrario tiene el mismo comportamiento pero por abajo del paralelo mínimo. Este comportamiento se debe a que  $\theta_0$  es el ángulo para el cual la ecuación de Clairaut se escribe como:

$$(1 + r \cos \psi) \cos \theta = 1 - r;$$

eso significa que cuando  $\theta$  tiende a cero,  $\psi$  tiende a  $\pi$ ; y si  $\theta$  llegara a ser cero, entonces  $\psi$  sería  $\pi$ , pero eso no puede ser, porque entonces, para cada  $\varphi$  habrían dos geodésicas que pasan por el punto  $(\varphi, \pi, 0)$  lo cual contradice el teorema de unicidad de ecuaciones diferenciales ordinarias. Por otra parte, cuando  $(\psi, \theta)$  tiende a  $(\pi, 0)$  o a  $(\pi, \pi)$ ; se tiene que

$$\dot{\psi} = \cos \theta / (1 + r \cos \psi)$$

tiende a  $\dot{\psi}_{00} = 1/(1-r) > 0$  ó  $\dot{\psi}_{00} = 1/(1-r) < 0$ ; es decir,  $\varphi$  crece o decrece monotonamente, es decir, la geodésica permanece dando vueltas.



geodésica  
de tipo C

### Lema.

Sobre cada uno de sus toros invariantes, el flujo geodésico del toro de revolución, tiene el mismo comportamiento que el flujo geodésico del toro plano: Las curvas integrales son todas periódicas o todas densas.

### Demostración.

Consideremos uno de los toros invariantes bajo el flujo geodésico, y en éste, escojamos un círculo  $S$  de la forma  $(\psi, \theta) = (\psi_0, \theta_0)$ .  $S$  está parametrizado por el ángulo  $\varphi$  y corta transversalmente a todas las curvas integrales contenidas en el toro invariante.

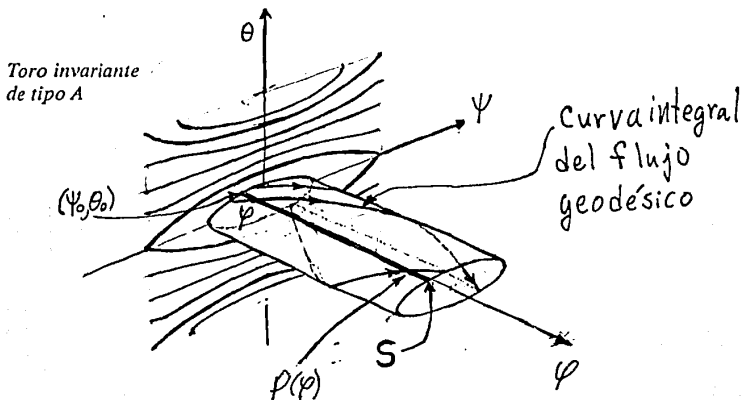
Consideremos la transformación  $P: S \rightarrow S$  que asigna a cada punto  $\varphi$  en  $S$  el primer punto  $P(\varphi)$  donde la curva integral que pasa por  $(\varphi, \psi_0, \theta_0)$  vuelve a cortar a  $S$ , (ver las figuras siguientes).  $P(\varphi)$  es una transformación que conmuta con las rotaciones en  $\varphi$ , por lo tanto, debe ser una rotación de un ángulo  $\Delta\varphi = 2\pi\lambda$ .

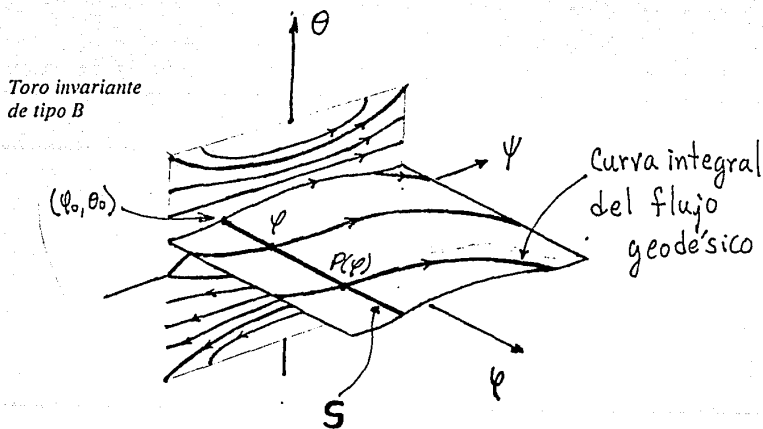
Del lema demostrado para el toro plano, se deduce que la órbita de un punto  $\varphi \in S$  (aplicando sucesivamente  $P$ ) es :

periódica si y solo si  $\lambda \in \mathbb{Q}$

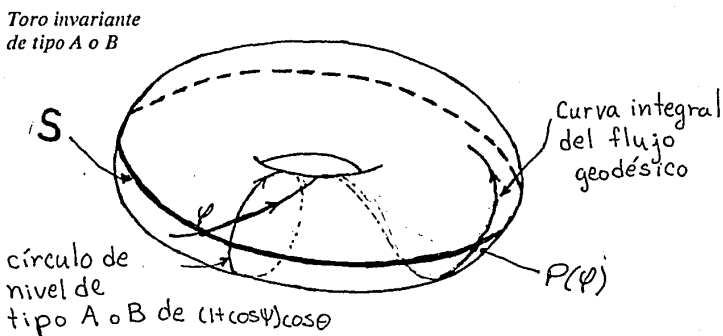
densa en  $S$  si y solo si  $\lambda \notin \mathbb{Q}$

Cabe señalar que una vez que fijamos el toro invariante, (es decir, una vez que fijamos la constante de la integral de Clairaut) el ángulo  $\Delta\varphi = 2\pi\lambda$  no depende del punto  $(\psi_0, \theta_0)$  que hayamos escogido sobre la curva de nivel asociada al toro invariante. Para ver que esto es cierto, supongamos que el ángulo  $\Delta\varphi$  varía al modificar el punto  $(\psi_0, \theta_0)$ ; en tal caso  $\Delta\varphi$  debe depender en forma continua del círculo  $S_{(\varphi_0, \theta_0)}$ , por lo cual  $\lambda = \Delta\varphi/2\pi$  toma valores tanto racionales como irracionales y, por tanto, las curvas integrales del flujo geodésico en este toro invariante son densas y periódicas a la vez, lo cual no puede suceder.





Las dos figuras anteriores pueden resumirse en la figura siguiente, ya que ambas representan toros invariantes.

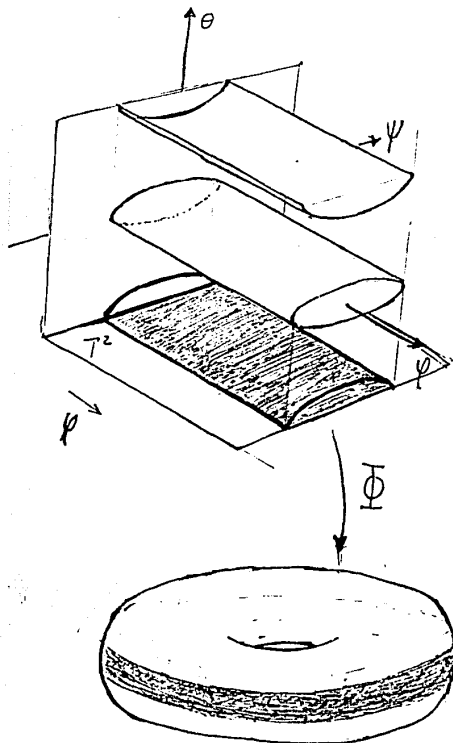




En el toro plano se tiene que la restricción a cada toro invariante de la proyección de  $T^3$  sobre  $T^2$

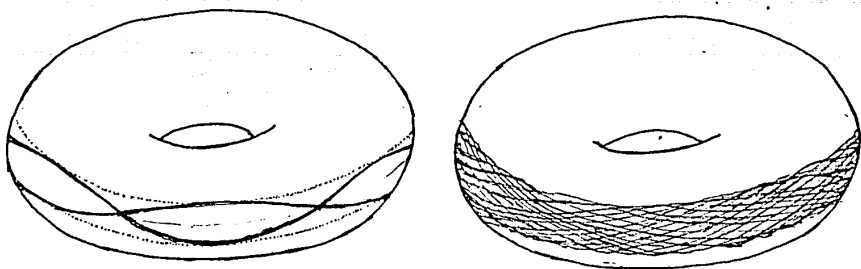
$$(\varphi, \psi, \theta) \rightarrow (\varphi, \psi),$$

es un difeomorfismo. Sin embargo, en el toro de revolución se tiene que los toros invariantes que corresponden a las geodésicas de tipo A no se proyectan biunívocamente sobre  $T^2$  y la integral de Clairaut determina un valor máximo para  $\psi$ :  $|\psi| \leq \psi_0 < \pi$ . La imagen del toro invariante bajo la proyección, es una banda que a su vez es la unión de geodésicas, obtenidas todas a partir de una de ellas por rotación en la dirección  $\varphi$ .

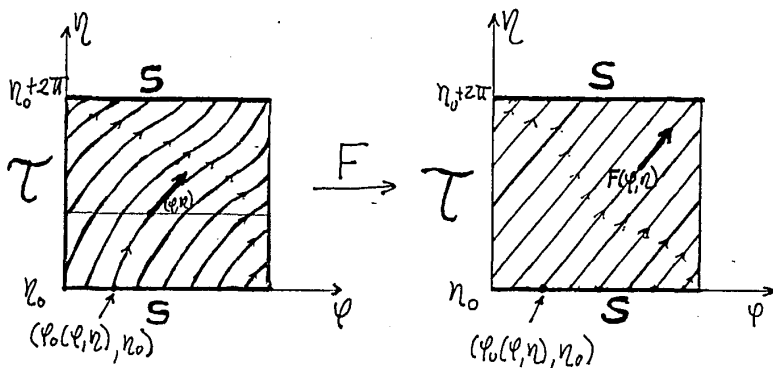


Proyección invariante de dos toros de tipo A, que corresponden a las mismas geodésicas, pero recorridas en sentidos opuestos

Nótese que una misma banda, es la proyección de dos toros invariantes distintos, que corresponden a las mismas geodésicas en el toro geométrico, solo que recorridas en sentidos opuestos. Estas geodésicas son o bien periódicas, o bien densas en la banda.



Se pueden dar otras coordenadas, para las cuales el flujo geodésico en los toros invariantes sea lineal. Para esto, fijemos un toro invariante  $\mathcal{T}$ , y designemos con  $\eta$  una coordenada angular sobre el "circulo de nivel" de la integral de Clairaut correspondiente al toro invariante  $\mathcal{T}$ .



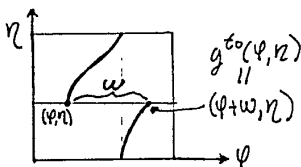
Consideremos un círculo  $S$ , de ecuación  $\eta = \eta_0$  sobre  $\mathcal{T}$ . Vamos a construir un difeomorfismo  $F: \mathcal{T} \rightarrow \mathcal{T}$  de la siguiente manera.

Dado un punto  $(\varphi, \eta)$  en el toro  $\mathcal{T}$ , lo llevamos hacia atrás (en el tiempo) hasta tocar por primera vez  $S$ , es decir, Al punto  $(\varphi, \eta)$  se le asocia —por un lado— el primer punto de intersección  $(\varphi_0(\eta, \varphi), \eta)$  de la curva integral que pasa por  $(\varphi, \eta)$  con  $S$ , recorrida en sentido negativo; y por otra parte, también se le asocia el tiempo  $t(\varphi, \eta) = t(\eta)$  que tarda el flujo en ir de  $(\varphi_0(\eta, \varphi), \eta)$  a  $(\varphi, \eta)$  es decir  $g^{t(\eta)}(\varphi_0(\eta, \varphi), \eta) = (\varphi, \eta)$ .

Entonces se da un difeomorfismo del toro  $\mathcal{T}$  en sí mismo mediante

$$F(\varphi, \eta) = (\varphi_0(\varphi, \eta) + \omega t(\eta)/t_\omega, \eta_0 + 2\pi t(\eta)/t_\omega) = (\tilde{\varphi}, \tilde{\eta})$$

donde  $\omega$  es el ángulo de la rotación  $P$ , y  $t_\omega$  es el tiempo que tarda en dar vuelta.



## Capítulo V

### Ecuaciones de Halmiton

#### 5.1 Paso a las coordenadas (q,p)

Observese que en el toro de revolución, se tiene que el lagrangiano  $G(q, \dot{q}) = 1/2 g(q) \dot{q}^2$  no depende de  $\varphi$ ; entonces si en el toro de revolución se considera un campo vectorial  $X$  paralelo al ecuador, se tiene que  $G(q, \dot{q})$  no cambia a lo largo de las trayectorias del campo  $X$ , ya que  $G(q, \dot{q})$  no depende de  $\varphi$ . Por esta razón se obtuvo la constante de movimiento  $\partial G / \partial \dot{q} = \text{constante}$ . Ahora tenemos algo más general.

#### Teorema de Noether.

Sea  $\Omega \subset \mathbb{R}^n$  un abierto y sea  $G(q, \dot{q})$  un lagrangiano definido en  $T\Omega = \Omega \times \mathbb{R}^n$ . Si  $X$  es un campo vectorial definido en  $\Omega$  tal que  $G$  es constante a lo largo de las curvas integrales de  $X$ , entonces hay una constante de movimiento:

$$P_x = \partial G / \partial \dot{q} \cdot X$$

es constante a lo largo de las curvas extremales de  $\int G$

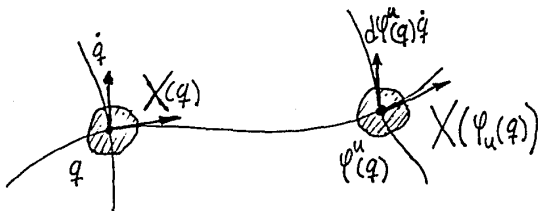
#### Demostración.

Denotemos con  $\varphi$  el flujo de  $X$ , entonces

$$\partial \varphi^u / \partial u|_{u=0}(q) = X(q)$$

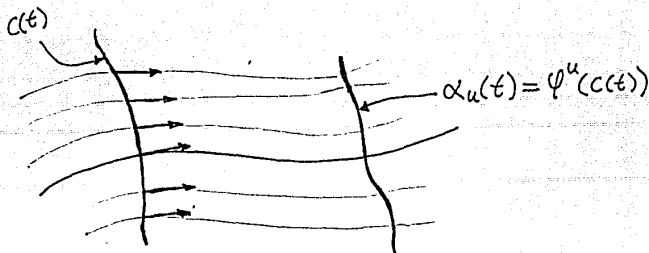
$$G(\varphi^u(q), d\varphi^u(q)\dot{q}) = G(q, \dot{q})$$

y



Sea  $C(t)$  una curva extremal de  $\int G$ . Si llevamos esta curva por el flujo  $\varphi$  durante un tiempo  $u$ , obtenemos una nueva curva  $\alpha_u(t) = \varphi^u(C(t))$  que también es extremal, ya que  $G$  no cambia de flujo de  $X$ .

$$\begin{aligned} G(\alpha_u(t), \dot{\alpha}_u(t)) &= G(\varphi^u(C(t)), d\varphi^u(C(t))\dot{C}(t)) \\ &= G(C(t), \dot{C}(t)); \end{aligned}$$



Como  $\frac{d}{dt} G(C(t), \dot{C}(t)) = 0$ , entonces

$$\begin{aligned}
 0 &= \frac{d}{dt} G(\alpha_u(t), \dot{\alpha}_u(t)) \\
 &= \frac{\partial}{\partial \alpha} G(\alpha_u(t), \dot{\alpha}_u(t)) \cdot \partial \alpha_u(t) / \partial u + \frac{\partial}{\partial \dot{\alpha}} G(\alpha_u(t), \dot{\alpha}_u(t)) \cdot \partial \dot{\alpha}_u(t) / \partial u \\
 &= \frac{\partial}{\partial \alpha} G(\alpha_u(t), \dot{\alpha}_u(t)) \cdot \partial \varphi^u(C(t)) / \partial u + \frac{\partial}{\partial \dot{\alpha}} G(\alpha_u(t), \dot{\alpha}_u(t)) \cdot \frac{\partial}{\partial u} \frac{d}{dt} (\varphi^u(C(t))) \\
 &= \frac{\partial}{\partial \alpha} G(\alpha_u(t), \dot{\alpha}_u(t)) \cdot \partial \varphi^u(C(t)) / \partial u + \frac{\partial}{\partial \dot{\alpha}} G(\alpha_u(t), \dot{\alpha}_u(t)) \cdot \frac{d}{dt} \frac{\partial}{\partial u} (\varphi^u(C(t)))
 \end{aligned}$$

En  $u=0$  se tiene

$$\frac{d}{dt} G(\alpha_u(t), \dot{\alpha}_u(t))|_{u=0} = \frac{\partial}{\partial \alpha} G(C(t), \dot{C}(t)) \cdot X(C(t)) + \frac{\partial}{\partial \dot{\alpha}} G(C(t), \dot{C}(t)) \cdot d(X(C(t))) / dt$$

Como  $P_x(C(t)) = \frac{\partial}{\partial \dot{q}} G(C(t), \dot{C}(t)) \cdot X(C(t))$ , entonces

$$\begin{aligned}
 \frac{d}{dt} P_x(C(t)) &= \frac{d}{dt} \left( \frac{\partial}{\partial \dot{q}} G(C(t), \dot{C}(t)) \right) \cdot X(C(t)) + \frac{\partial}{\partial \dot{q}} G(C(t), \dot{C}(t)) \cdot \frac{d}{dt} (X(C(t))) \\
 \text{por (E-L)} &= \frac{\partial}{\partial \dot{q}} G(C(t), \dot{C}(t)) \cdot X(C(t)) + G(C(t), \dot{C}(t)) \cdot X(C(t)) \\
 &= \frac{d}{dt} G(\alpha_u(t), \dot{\alpha}_u(t))|_{u=0} = 0 \quad \blacklozenge
 \end{aligned}$$

A  $P_x = X \cdot \partial G / \partial \dot{q}$  se le puede llamar el momento en la dirección  $X$ ; y a  $P = \partial G / \partial \dot{q}$  se le puede llamar el momento generalizado.

Como caso particular, considere el campo vectorial  $X(q) = e_i$ . Para este campo se tiene que  $G$  es invariante a lo largo de las soluciones de  $e_i$

si sólo si  $G$  no depende de  $q_i$ .

si y sólo si,  $\frac{d}{dt} (\partial G / \partial \dot{q}_i) = 0$

si y sólo si  $P_i = \partial G / \partial \dot{q}_i$  es constante de movimiento:

Este es el caso del toro de revolución. Como  $G$  no depende de  $\varphi$ , entonces  $P_\varphi$  es constante de movimiento:

$$\begin{aligned}
 P_\varphi &= (1 + r \cos \psi)^2 \dot{\varphi} \\
 &= (1 + r \cos \psi) \cos \theta = \text{constante}
 \end{aligned}$$

Si  $G$  depende en forma cuadrática de  $\dot{q}$ , se tiene que

$$\begin{aligned}
 G &= 1/2 g(q) \dot{q}^2, \quad \text{por lo cual} \\
 P &= \partial G / \partial \dot{q} = g(q) \dot{q}
 \end{aligned}$$

Ejemplo.

Toro de revolución  $q_1 = \varphi$ ;  $q_2 = \psi$ ;  $X = e_\varphi = (1, 0)$

Entonces

$$P = \partial G / \partial \dot{q} = g(q) \dot{q} = \begin{pmatrix} (1+r\cos\psi)^2 & 0 \\ 0 & r^2 \end{pmatrix} \begin{pmatrix} \dot{\psi} \\ \dot{\psi} \end{pmatrix}$$

$$= \begin{pmatrix} (1+r\cos\psi)^2 \dot{\psi} \\ r^2 \dot{\psi} \end{pmatrix}$$

y  $P_x = P_\psi = (1+r\cos\psi)^2 \dot{\psi}$

Ahora veamos una interpretación de P. Sea  $\Omega = \mathbb{R}^n$  un medio óptico homogéneo, pero anisótropo. En este caso se tiene que  $g(q)$  es constante, y por lo tanto

$$G(q, \dot{q}) = (1/2) g \dot{q} \cdot \dot{q};$$

de este modo la ecuación de (E-L) es

$$\frac{d}{dt}(g\dot{q}) = 0$$

es decir,

$$g\dot{q} = \text{constante};$$

de modo que

$$\dot{q} = \text{constante}.$$

Por lo tanto, las trayectorias que sigue la luz son rectas. Ahora, considerensen todas las trayectorias que salen del origen; o supóngase que se ha colocado en el origen una fuente puntual de luz que se enciende un instante al tiempo  $t=0$ .

En tal caso se tiene que

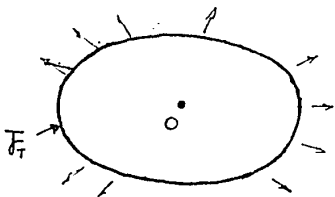
$$q = T\dot{q}, \text{ con } T \in \mathbb{R}, T \geq 0;$$

de modo que la longitud óptica, al tiempo  $T$ , de la trayectoria de un rayo que sale del origen con velocidad  $\dot{q}$  es

$$\|T\dot{q}\| = (gT\dot{q} \cdot T\dot{q})^{1/2} = T(g\dot{q} \cdot \dot{q})^{1/2} = T;$$

esto quiere decir que el frente de onda, al tiempo  $T$ , es el conjunto de puntos en  $\Omega$ , cuya distancia óptica al origen es  $T$ :

$$\mathcal{F}_T = \{q \in \mathbb{R}^n : (gq \cdot q)^{1/2} = T\}$$



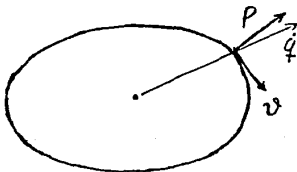
onda luminosa  
a tiempo  $t=T$

Considérese la función  $F: \Omega \rightarrow \mathbb{R}$ , definida por  $F(q) = (gq \cdot q)^{1/2}$ ; entonces, para cada  $T \in \mathbb{R}^+$  el frente de onda al tiempo  $T$ , es un conjunto de nivel de  $F$ . Entonces,  $\nabla F(q) = \nabla F(T\dot{q})$  es perpendicular al frente de onda al tiempo  $T$ . Ahora calculemos

$$\begin{aligned} \nabla F(q) &= gq / (gq \cdot q)^{1/2} \text{ en } q = T\dot{q} \\ \nabla F(q) &= gq / (gq \cdot q)^{1/2} = gT\dot{q} / (gT\dot{q} \cdot T\dot{q})^{1/2} = Tg\dot{q} / T(g\dot{q} \cdot \dot{q})^{1/2} = g\dot{q}, \end{aligned}$$

pero  $g\dot{q} = \partial G / \partial \dot{q} = P$

Por lo tanto,  $P$  es un vector perpendicular al frente de onda, es decir, en cada punto,  $P$  indica la dirección en la que se mueve el frente de onda



El vector  $\vartheta$  es tangente al frente de onda

un observador parado en  $q$ , ve pasar el fotón con velocidad  $\dot{q}$ , pero ve pasar la onda con "lentitud"  $P$ . Además, para nosotros  $P$  es perpendicular al frente de onda, pero para la luz  $\dot{q}$  es perpendicular al frente de onda, ya que, si  $\vartheta$  es un vector tangente al frente de onda entonces

$$P \cdot \vartheta = 0$$

pero

$$\begin{aligned} P \cdot \vartheta &= gq \cdot \vartheta \\ &= \langle \dot{q}, \vartheta \rangle_g; \end{aligned}$$

es decir,

$$\langle \dot{q}, \vartheta \rangle_g = 0.$$

Se puede hacer un cambio de coordenadas, tomando como nuevas coordenadas  $(q, p)$  en lugar de  $(q, \dot{q})$ . Como  $p = g(q)\dot{q}$ , entonces  $\dot{q} = g^{-1}(q)p$ ; esto establece un isomorfismo

$$\begin{aligned} \Omega \times \mathbb{R}^n &\rightarrow \Omega \times \mathbb{R}^n \\ (q, \dot{q}) &\rightarrow (q, p); \end{aligned}$$

y se obtiene un sistema de ecuaciones diferenciales.

$$\begin{cases} \dot{q} = g^{-1}(q)p \\ \dot{p} = \frac{d}{dt} \left( \frac{\partial G}{\partial \dot{q}} \right) = \frac{\partial}{\partial q} G(q, g^{-1}(q)p) \end{cases} \quad (1)$$

Si  $G = G(q, \dot{q})$  se escribe tomando en cuenta que  $\dot{q} = g^{-1}(q)p$ , entonces  $G(q, g^{-1}(q)p)$  es una función de  $q$  y  $p$  que vamos a denotar por  $H(q, p)$ ; ahora, la idea es escribir el sistema (1) en términos de  $H$ . Entonces veamos quien es  $H$ :

$$\begin{aligned} H(q, p) &= G(q, g^{-1}(q)p) = 1/2 g(q)g^{-1}(q)p \cdot g^{-1}(q)p \\ &= 1/2 g^{-1}(q)p \cdot p \end{aligned}$$

Entonces,

$$\partial H / \partial p = g^{-1}(q)p = \dot{q}$$

Para calcular  $\partial H / \partial q$  observemos que, como  $g(q)$  es simétrica, entonces  $g^{-1}(q)$  también lo es; además

$$\frac{\partial}{\partial q} (g(q)g^{-1}(q)) = 0$$

lo cual implica que

$$\begin{aligned} g(q) \frac{\partial g^{-1}(q) / \partial q + (\partial g(q) / \partial q) g^{-1}(q)}{\partial g^{-1}(q) / \partial q} &= 0, \\ \text{ó} \quad \frac{\partial g^{-1}(q) / \partial q}{\partial g^{-1}(q) / \partial q} &= -g^{-1}(q) (\partial g(q) / \partial q) g^{-1}(q). \end{aligned}$$

Entonces

$$\begin{aligned} \partial H / \partial q &= 1/2 (\partial g^{-1}(q) / \partial q) p \cdot p \\ &= -1/2 g^{-1}(q) (\partial g(q) / \partial q) g^{-1}(q) p \cdot p \\ &= -1/2 p' g^{-1}(q) (\partial g(q) / \partial q) g^{-1}(q) p \\ g^{-1} = (g^{-1})': &= -1/2 (g^{-1}(q) p)' (\partial g(q) / \partial q) g^{-1}(q) p \\ &= -1/2 \dot{q}' (\partial g(q) / \partial q) \dot{q} \\ &= -1/2 (\partial g(q) / \partial q) \dot{q}' \dot{q} \\ &= -\partial G(q, \dot{q}) / \partial q \\ \text{(E-L):} &= -\frac{d}{dt} (\partial G(q, \dot{q}) / \partial \dot{q}) \\ &= -\dot{p} \end{aligned}$$

## 5.2 Ecuaciones de Hamilton

Ahora tenemos la función  $H: \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}$ , llamada función de Hamilton o hamiltoniano y tenemos el sistema de  $2n$  ecuaciones diferenciales de primer orden

$$\begin{aligned} \dot{q} &= \partial H / \partial p \\ \dot{p} &= -\partial H / \partial q \end{aligned} \quad (2)$$

llamadas ecuaciones de Hamilton.

**Advertencia.**

$H(q, P)$  es  $G(q, g^{-1}(q)P)$  en las nuevas coordenadas, porque  $G$  es cuadrática en  $\dot{q}$ . Si no es así, entonces el paso de el lagrangiano  $G$  al hamiltoniano  $H$ , se hace através de la transformación de Legendre.

**Proposición.**

$H$  es constante a lo largo de las soluciones de (2).

**Demostración.**

Sea  $q(t)$ ,  $p(t)$  solución de (2), y sea

$$\begin{aligned} h(t) &= H(q(t), p(t)), \text{ entonces} \\ h'(t) &= \partial H / \partial q \cdot \dot{q}(t) + \partial H / \partial p \cdot \dot{p}(t) \\ &= \partial H / \partial q \cdot \partial H / \partial p - \partial H / \partial p \cdot \partial H / \partial q = 0. \end{aligned}$$

Ahora vamos a ver algo acerca de la conservación del volumen.

Sea  $U \subset \mathbb{R}^n$  un abierto, y sea  $F$  un campo vectorial suave en  $U$ . Suponiendo que todas las soluciones  $\sigma_x(t)$ , del campo vectorial  $F$  con la condición inicial  $\sigma_x(0) = x$  están definidas  $\forall t$ , hay un flujo  $\varphi: U \times \mathbb{R} \rightarrow U$  asociado al campo vectorial  $F$ , en la siguiente forma.



**ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA**

Se define  $\varphi$  mediante

$$\varphi(x,t) = \sigma_x(t); \quad \text{de esta forma,}$$

$$\varphi(x, 0) = x \quad \text{y} \quad \partial\varphi(x,t)/\partial t = F(\varphi(x,t)).$$

Para  $t$  fijo  $\varphi^t(x) = \varphi(x,t)$  es una función suave de  $U$  en  $U$ . De este modo el flujo define una familia de funciones a un parámetro

$$\varphi^t: U \rightarrow U, \quad t \in \mathbb{R};$$

esta familia satisface

$$\begin{aligned} \text{a) } & \varphi(x,0) = x \\ \text{b) } & \varphi(\varphi(x,t),s) = \varphi(x,t+s), \end{aligned}$$

es decir,

$$\begin{aligned} \text{a) } & \varphi^0 = \text{id} \\ \text{b) } & \varphi^s \circ \varphi^t = \varphi^{s+t}. \end{aligned}$$

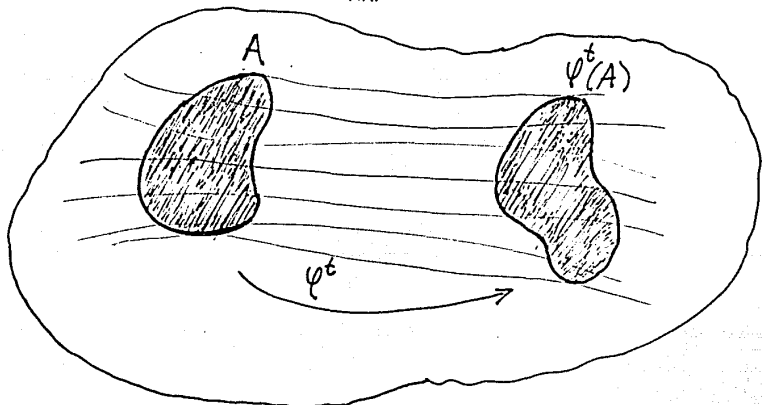
Ejercicio.

Probar que la familia  $\{\varphi^t\}_{t \in \mathbb{R}}$  es un grupo conmutativo.

Sea  $A \subset U$  un compacto.

Definimos  $v(t) = \text{vol}(\varphi^t(A))$  entonces

$$v(t) = \int_{\varphi^t(A)} 1 = \int_A J\varphi^t,$$



$$v'(t) = \frac{d}{dt} \int_A J\varphi^t = \int_A \frac{\partial}{\partial t} J\varphi^t,$$

pero, ¿cuánto vale  $\partial(J\varphi^t)/\partial t$ ?

$J\varphi^t = \det(D\varphi^t)$ , es decir, si  $\varphi^t = (\varphi_1^t, \varphi_2^t, \dots, \varphi_m^t)$

$$J\varphi^t = \begin{vmatrix} \frac{\partial \varphi_1^t}{\partial x_1} & \dots & \frac{\partial \varphi_1^t}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial \varphi_m^t}{\partial x_1} & \dots & \frac{\partial \varphi_m^t}{\partial x_m} \end{vmatrix};$$

entonces

$$\frac{\partial(J\varphi^t)/\partial t} = \sum_{i=1}^m \begin{vmatrix} \frac{\partial \varphi_1^t}{\partial x_1} & \dots & \frac{\partial}{\partial t} \left( \frac{\partial \varphi_1^t}{\partial x_i} \right) & \dots & \frac{\partial \varphi_1^t}{\partial x_m} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial \varphi_m^t}{\partial x_1} & \dots & \frac{\partial}{\partial t} \left( \frac{\partial \varphi_m^t}{\partial x_i} \right) & \dots & \frac{\partial \varphi_m^t}{\partial x_m} \end{vmatrix},$$

pero  $\frac{\partial}{\partial t} \left( \frac{\partial \varphi_i^t}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left( \frac{\partial \varphi_i^t}{\partial t} \right) = \frac{\partial}{\partial x_j} (F_i(\varphi^t))$ ,

de modo que

$$\frac{\partial}{\partial t} (J\varphi^t) = \sum_{i=1}^m \begin{vmatrix} \frac{\partial \varphi_1^t}{\partial x_1} & \dots & \frac{\partial F_i}{\partial x_i}(\varphi^t) & \dots & \frac{\partial \varphi_1^t}{\partial x_m} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial \varphi_m^t}{\partial x_1} & \dots & \frac{\partial F_m}{\partial x_i}(\varphi^t) & \dots & \frac{\partial \varphi_m^t}{\partial x_m} \end{vmatrix}$$

en  $t = 0$  se tiene que  $\partial \varphi_i^t / \partial x_i|_{t=0} = \delta_{ij}$ , es decir,

$$\frac{\partial}{\partial t} J\varphi^t|_{t=0} = \sum_{i=1}^m \begin{vmatrix} 1 & 0 & \dots & \frac{\partial F_i}{\partial x_i} & \dots & 0 \\ 0 & 1 & & \vdots & & \vdots \\ 0 & 0 & & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & & \frac{\partial F_m}{\partial x_i} & \dots & 1 \end{vmatrix}$$

desarrollando el i-ésimo sumando por la i-ésima columna se obtiene que

$$\frac{\partial}{\partial t} J\varphi|_{t=0} = \sum_{i=1}^m \partial F_i / \partial x_i = \operatorname{div}(F);$$

entonces,

$$v'(0) = \int_A \operatorname{div}(F)$$

Si  $\int_A \operatorname{div}(F) \neq 0$  el proceso "empieza mal", es decir, el volumen de A, crece (o decrece) de entrada; si  $\int_A \operatorname{div}(F) = 0$ , se tiene que el "cambio inicial" del volumen A es cero. Pero, si  $\vartheta'(0)=0$ , podemos concebir la cosa como que en cada tiempo t se vuelve a empezar, y cada vez que empezamos de nuevo  $\vartheta'(0)=0$ . Veamos como es esto, usando las propiedades del flujo. Ahora hay que saber cuanto vale

$$\partial(J\varphi^t)/\partial t|_{t=0};$$

para esto hagamos  $t=s+t_0$ ; así,

$$\varphi^t = \varphi^s \circ \varphi^{t_0},$$

por lo que

$$D\varphi^t(x) = D\varphi^s(\varphi^{t_0}(x)) \cdot D\varphi^{t_0}(x);$$

por lo tanto

$$y \quad \frac{\partial}{\partial t} J\varphi^t|_{t=0} = \frac{\partial}{\partial t} (J\varphi^s(\varphi^{t_0}(x)))|_{t=0} \cdot J\varphi^{t_0}(x) = \operatorname{div} F(\varphi^{t_0}(x)) J\varphi^{t_0}(x)$$

De esta manera

$$v'(t_0) = \int_A \operatorname{div} F(\varphi^{t_0}(x)) J\varphi^{t_0}(x)$$

**Teorema:**

$\varphi^t$  preserva volumen, es decir,  $\operatorname{vol}(\varphi^t(A)) = \operatorname{vol}(A)$  si y sólo si  $\operatorname{div} F = 0$

**Demostración.**

Si  $\operatorname{div} F = 0$  entonces  $v'(t) = 0 \quad \forall t \in \mathbb{R}$ , por tanto  $v(t)$  es constante, es decir,  $\varphi^t$  preserva volumen. Recíprocamente, si  $\varphi^t$  preserva volumen, entonces  $v(t)$  es constante para toda A, por lo tanto  $v'(0) = 0$  para toda A, lo que implica que  $\int_A \operatorname{div} F = 0$  para todo conjunto A, y  $\therefore \operatorname{div} F = 0$

Ahora volvamos al flujo hamiltoniano. Tenemos la función de Hamilton

$$H: \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad H = H(q, p).$$

Las ecuaciones de Hamilton definen un campo vectorial en  $\Omega \times \mathbb{R}^n$

$$X_H = (\partial H / \partial p, -\partial H / \partial q)$$

### Teorema de Liouville.

El flujo de  $X_H$  (llamado flujo hamiltoniano) preserva el volumen.

Demostración.

$$X_H = (\partial H / \partial p_1, \dots, \partial H / \partial p_n, -\partial H / \partial q_1, \dots, -\partial H / \partial q_n)$$

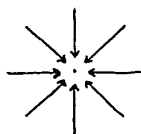
$$\operatorname{div}(X_H) = \frac{\partial^2 H}{\partial q_1 \partial p_1} + \frac{\partial^2 H}{\partial q_2 \partial p_2} + \dots + \frac{\partial^2 H}{\partial q_n \partial p_n} - \frac{\partial^2 H}{\partial p_1 \partial q_1} - \dots - \frac{\partial^2 H}{\partial p_n \partial q_n}$$

$$= 0$$

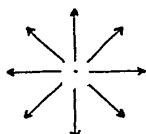
ya que estamos suponiendo, que todas las parciales son continuas.

En el flujo Hamiltoniano

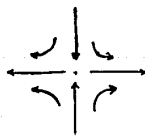
- i) No puede haber atractores ni repulsores, ni ciclos límite, ya que  $\phi$  preservan volumen, sin embargo,
- ii) Si puede haber singularidades, por ejemplo punto silla



no se puede



no se puede



si puede ser un  
flujo hamiltoniano

Supongamos que  $H^{-1}(a)$  es compacto,  $a$  es valor regular y  $dH \neq 0$  en todo punto de  $H^{-1}(a)$ . Vamos a suponer también, que el conjunto

$$D = \{(q,p) \in \Omega \times \mathbb{R}^n : H(q,p) \leq a\}$$

también es compacto. El flujo se queda adentro de  $H^{-1}(a)$  y no puede cruzar la "cáscara", porque el campo es tangente, ya que a lo largo de las soluciones del campo  $X_H$ ,  $H$  es constante.

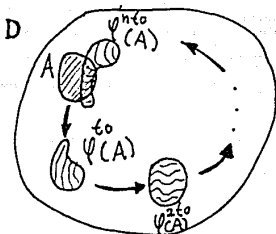
Una de las consecuencias de las propiedades del flujo hamiltoniano, es el siguiente teorema.

### Teorema de Recurrencia de Poincaré. (o del eterno retorno)

En cualquier vecindad  $A$  de cualquier punto en  $D$ , hay un punto  $x$  en  $A$  que regresa a  $A$ , es decir,  $\phi^t(x) \in A$  para algún  $t > 0$  o, un  $\phi^t(A) \cap A \neq \emptyset$ .

Demostración.

Si  $\phi^t(A) \cap A \neq \emptyset \quad \forall t > 0$ , no hay nada que demostrar. Supongamos que para algún  $t_0 > 0$ ,  $\phi^{t_0}(A) \cap A = \emptyset$



Si  $\varphi^{nt_0}(A) \cap A = \emptyset \quad \forall n \in \mathbb{N}$ , entonces los conjuntos  $A, \varphi^{t_0}(A), \dots, \varphi^{nt_0}(A), \dots$  son ajenos, es decir,  $\varphi^{nt_0}(A) \cap \varphi^{mt_0}(A) = \emptyset \quad \forall n, m \in \mathbb{N}$ , pero entonces

$$\text{vol}(\cup \varphi^n(A)) = \sum \text{vol}(\varphi^{nt_0}(A)) = \infty,$$

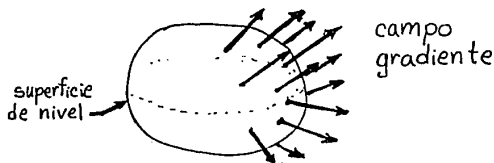
lo cual contradice la hipótesis de que D es compacto. En particular, dado cualquier  $x \in H^{-1}(a)$  al cabo de cierto tiempo, la órbita de  $x$  pasa tan cerca como se quiera de  $x$  (Tomamos A con  $\text{vol}(A) < \epsilon$  y  $x \in A$ ).

Ahora, se verá una comparación entre los campos vectoriales hamiltonianos ( $X_H$ ) y los campos gradientes. Sea  $U \subset \mathbb{R}^m$  un abierto, y  $F: U \rightarrow \mathbb{R}$  una función suave.

$\nabla F$  tiene las siguientes propiedades:

- $\nabla F$  es perpendicular a los conjuntos de nivel de  $F$ ;
- $\nabla F \cdot \xi = dF(\xi), \quad \xi \in \mathbb{R}^m$ ;
- $\nabla F$  no preserva volumen y por lo general no preserva medidas.

Como las soluciones de la ecuación  $\dot{x} = \nabla F$  son perpendiculares a los conjuntos de nivel de  $F$ ,  $\dot{x} = \nabla F$  no tiene soluciones periódicas (se escapan).



Una propiedad importante es que  $\nabla F$  queda determinado por completo, por la ecuación del inciso b); observemos que si  $B: V \times V \rightarrow \mathbb{R}$  ( $V$  espacio vectorial) es una función bilineal, simétrica y no-degenerada; y a cada  $x \in V$  le asociamos una función lineal  $f_x(y) = B(x, y)$ , entonces tenemos un isomorfismo

$$\tilde{B}: V \rightarrow V'$$

Definido por  $\tilde{B}(x) = f_x$ , donde

$V^* = \{f: V \rightarrow R: \text{lineal}\}$  es el espacio dual de  $V$ .

Si  $f: V \rightarrow R$  es lineal, entonces existe un único  $x$  en  $V$  tal que  $\tilde{B}(x) = f$ , es decir,

$$f(y) = f_x(y) = B(x, y)$$

En particular, consideremos  $B(x, y) = x \cdot y$ ; para  $x$  en  $U$  fijo, la diferencial  $dF_x: R^n \rightarrow R$ , es una función lineal, es decir  $(dF_x) \in (R^n)^*$ , por lo tanto existe un único vector  $(\nabla F_x)$  tal que

$$dF_x(\xi) = \nabla F_x \cdot \xi.$$

En el campo vectorial que determinan las ecuaciones de Hamilton, la cosa va a ser "un poco" distinta, en este caso, el papel de la forma bilineal no-degenerada (el producto punto), lo va a jugar una forma bilineal antisimétrica no-degenerada (forma simpléctica). Veamos como es esto.

Ahora estamos en nuestro espacio con coordenadas  $(q, p)$  en  $R^{2n}$ ; vamos a considerar vectores  $u = (\xi, \eta) \in R^{2n}$ . Definimos la forma bilineal  $\omega(u_1, u_2) = \xi_1 \cdot \eta_2 - \xi_2 \cdot \eta_1$ ;  $\omega$  es una forma bilineal, antisimétrica y no-degenerada, es decir, es una forma simpléctica. La base usual de  $R^{2n}$ ,  $e_1, \dots, e_n, f_1, \dots, f_n$  es una base simpléctica:

$$\omega(e_i, e_j) = \omega(f_i, f_j) = 0,$$

$$\omega(e_i, f_j) = \delta_{ij};$$

y la matriz de  $\omega$  es:

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

Estamos considerando un campo vectorial  $X_H$  definido por las ecuaciones de Hamilton

$$X_H = (\partial H / \partial p, -\partial H / \partial q),$$

decimos que este campo es el "gradiente simpléctico" de  $H$ .

Si  $u = (\xi, \eta)$ , entonces

$$\begin{aligned} \omega(X_{H(q,p)}, u) &= \partial H / \partial p \cdot \eta - \xi \cdot (-\partial H / \partial q) \\ &= \partial H / \partial q \cdot \xi + \partial H / \partial p \cdot \eta = dH_{(q,p)}(u) \end{aligned}$$

Para el gradiente simpléctico  $X_H$  tenemos las siguientes propiedades:

- $X_H$  es tangente a las curvas de nivel de  $H$ ;
- $\omega(X_{H(q,p)}, u) = dH_{(q,p)}(u)$ ,  $u \in R^{2n}$ ;
- $X_H$  preserva volumen.

$X_H$  es tangente a los conjuntos de nivel de  $H$ , debido a que, a lo largo de las trayectorias de  $X_H$ ,  $H$  es constante (algo que ya demostramos). Podemos verlo así:

$$\frac{d}{dt} H(q(t), p(t)) = dH(\dot{q}(t), \dot{p}(t)) = dH(X_H) = \omega(X_H, X_H) = 0$$

En particular, si la superficie de nivel (conjunto de nivel) es compacta, las soluciones de  $X_H$  no se escapan, sino que se quedan dando vueltas ahí. Consideremos un hamiltoniano  $H: \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}$  y su gradiente simpléctico  $X_H = (\partial H / \partial p, \partial H / \partial q)$ .

El flujo de  $\varphi^t$  asociado a  $X_H$ , tiene las siguientes propiedades (que ya vimos):

1)  $H$  es constante a lo largo de  $\varphi^t$

2)  $\varphi^t$  preserva volúmen;

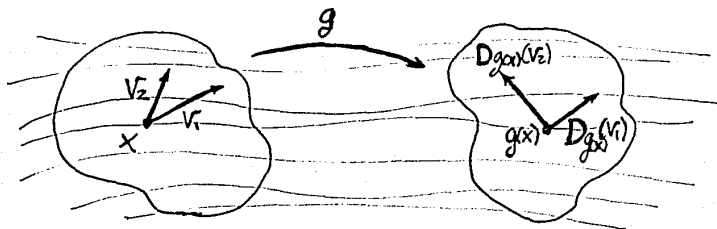
ahora veamos una tercera propiedad.

**Teorema:**

$\varphi^t$  preserva  $\omega$  (o bien,  $\varphi^t$  es una transformación simpléctica) (para los físicos  $\varphi^t$  es una transformación canónica).

En general se dice que una transformación  $g: (X, \omega) \rightarrow (X, \omega)$  es canónica, o simpléctica o preserva  $\omega$  si

$$\omega(v_1, v_2) = \omega(Dg_{(x)}(v_1), Dg_{(x)}(v_2))$$



### Demostración.

Vamos a considerar  $X \in \Omega \times \mathbb{R}^n$ ,  $t \in \mathbb{R}$ ,  $u, v \in \mathbb{R}^{2n}$

$$\text{P.D. } \omega(u, v) = \omega(D\varphi^t_{(x)}(u), D\varphi^t_{(x)}(v)) \\ = a(t);$$

es decir, P.D.  $a(t) = \text{constante para toda } t$ , es decir  $a'(t) = 0$

$$a'(t) = \frac{d}{dt} \omega(D\varphi^t(u), D\varphi^t(v)) \\ = \omega\left(\frac{d}{dt}(D\varphi^t)(u), D\varphi^t(v)\right) + \omega(D\varphi^t(u), \frac{d}{dt}(D\varphi^t)(v));$$

pero  $\frac{d}{dt}(D\varphi^t) = D(d\varphi^t/dt) = DX_H(\varphi^t)$ ;

así,  $\frac{d}{dt}(D\varphi^t)|_{t=0}(u) = DX_H(u)$ , por lo tanto

$$a'(0) = \omega(DX_H(u), v) + \omega(u, DX_H(v)).$$

La demostración consiste en desarrollar  $\omega(DX_H(u), v)$  y mostrar que es simétrico en  $u$  y  $v$ , es decir, hay que mostrar que

$$\omega(DX_H(u), v) = \omega(DX_H(v), u).$$

Se tiene

$$X_H = (\partial H / \partial p_1, \dots, \partial H / \partial p_n, -\partial H / \partial q_1, \dots, -\partial H / \partial q_n),$$

$$u = (u_1, \dots, u_n, u_1', \dots, u_n') \text{ y } v = (v_1, \dots, v_n, v_1', \dots, v_n')$$

entonces

$$DX_H(u) = \begin{pmatrix} \frac{\partial^2 H}{\partial q_1 \partial p_1} & \dots & \frac{\partial^2 H}{\partial q_n \partial p_1} & \frac{\partial^2 H}{\partial p_1 \partial p_1} & \dots & \frac{\partial^2 H}{\partial p_n \partial p_1} \\ \vdots & & & & & \\ \frac{\partial^2 H}{\partial q_1 \partial p_n} & & & & & \\ -\frac{\partial^2 H}{\partial q_1 \partial q_1} & & & & & \\ \vdots & & & & & \\ -\frac{\partial^2 H}{\partial q_1 \partial q_n} & \dots & -\frac{\partial^2 H}{\partial q_n \partial q_n} & -\frac{\partial^2 H}{\partial p_1 \partial q_n} & \dots & -\frac{\partial^2 H}{\partial p_n \partial q_n} \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \\ u_1' \\ \vdots \\ u_n' \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^n \left( \frac{\partial^2 H}{\partial q_i \partial p_1} u_i + \frac{\partial^2 H}{\partial p_1 \partial p_1} u_i' \right) \\ \vdots \\ \sum_{i=1}^n \left( \frac{\partial^2 H}{\partial q_i \partial p_n} u_i + \frac{\partial^2 H}{\partial p_n \partial p_n} u_i' \right) \\ - \sum_{i=1}^n \left( \frac{\partial^2 H}{\partial q_i \partial q_1} u_i + \frac{\partial^2 H}{\partial p_1 \partial q_1} u_i' \right) \\ \vdots \\ - \sum_{i=1}^n \left( \frac{\partial^2 H}{\partial q_i \partial q_n} u_i + \frac{\partial^2 H}{\partial p_1 \partial q_n} u_i' \right) \end{pmatrix}$$



Entonces,

$$\begin{aligned} \omega(DX_H(u), v) &= \sum_{j=1}^n \left( \sum_{i=1}^n \left[ \frac{\partial^2 H}{\partial q_i \partial p_j} u_i + \frac{\partial^2 H}{\partial p_i \partial p_j} u_i^j \right] v_j^i - \sum_{i=1}^n \left( -\sum_{j=1}^n \left[ \frac{\partial^2 H}{\partial q_i \partial q_j} u_i + \frac{\partial^2 H}{\partial p_i \partial q_j} u_i^j \right] v_j^i \right) \right) \\ &= \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 H}{\partial q_i \partial p_j} u_i v_j^i + \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 H}{\partial p_i \partial p_j} u_i^j v_j^i + \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 H}{\partial p_i \partial p_j} u_i^j v_j^i + \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 H}{\partial q_i \partial q_j} u_i v_j^i \end{aligned}$$

como los índices son mudos podemos cambiar i por j y j por i y donde haga falta para obtener:

$$= \sum_{i,j} \frac{\partial^2 H}{\partial q_i \partial p_j} (u_i v_j^i + u_j^i v_i) + \sum_{i,j} \frac{\partial^2 H}{\partial p_i \partial p_j} u_i^j v_j^i + \sum_{i,j} \frac{\partial^2 H}{\partial q_i \partial q_j} u_i v_j^i$$

es evidente que el primer término es simétrico en  $u$  y  $v$ ; veamos que los otros dos términos también lo son, usando que los índices i y j pueden intercambiarse :

$$\sum_{i,j} \frac{\partial^2 H}{\partial p_i \partial p_j} u_i^j v_j^i = \sum_{i,j} \frac{\partial^2 H}{\partial p_j \partial p_i} u_j^i v_i^j = \sum_{i,j} \frac{\partial^2 H}{\partial q_j \partial p_i} v_i^j u_j^i ;$$

$$\sum_{i,j} \frac{\partial^2 H}{\partial q_i \partial q_j} u_i v_j^i = \sum_{i,j} \frac{\partial^2 H}{\partial q_j \partial q_i} u_j^i v_i = \sum_{i,j} \frac{\partial^2 H}{\partial q_j \partial q_i} v_i u_j^i .$$

Hemos demostrado que

$$\omega(DX_H(u), v) = \omega(DX_H(v), u),$$

por lo tanto

$$a'(0) = \omega(DX_H(v), u) + \omega(u, DX_H(v)) = 0$$

Ahora, veamos que  $a'(t_0) = 0$ , usando la misma idea que cuando demostramos que  $\Psi$  preserva volumen.

$$a'(t_0) = \left. \frac{d}{ds} (a(t_0 + s)) \right|_{s=0} ,$$

pero

$$\begin{aligned} a(t_0+s) &= \omega \left( D\varphi_{(x)}^{t_0+s}(u), D\varphi_{(x)}^{t_0+s}(v) \right) \\ &= \omega \left( D\varphi^S(\varphi_{(x)}^{t_0}) D\varphi_{(x)}^{t_0}(u), D\varphi^S(\varphi_{(x)}^{t_0}) D\varphi_{(x)}^{t_0}(v) \right) \end{aligned}$$

Como  $a(t) = \omega(D\varphi_{(x)}^t(u), D\varphi_{(x)}^t(v))$  en  $(x, u, v)$ ,

entonces podemos poner  $\tilde{x} = \varphi_{(x)}^t$

$$\tilde{u} = D\varphi_{(x)}^t(u)$$

$$\tilde{v} = D\varphi_{(x)}^t(v)$$

y entonces se puede escribir

$$a(t_0+s) = \omega(D\varphi_{(\tilde{x})}^s(\tilde{u}), D\varphi_{(\tilde{x})}^s(\tilde{v})) = \tilde{a}(s)$$

entonces

$$a'(t_0) = \tilde{a}'(0) = 0$$

por lo tanto  $a(t)$  es constante, es decir,  $\varphi^t$  es una transformación simpléctica, o decimos que el flujo hamiltoniano consiste de transformaciones simplécticas.

## Bibliografía

1. Victor Guillemin, *Symplectic Technics in Physics*, Cambridge University.
2. Alain Chenciner, *Notas de Mecánica Celeste* (versión preliminar) Université. de Paris VII.
3. V.I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer Verlag, N.Y. 1978.
4. Lius ternik, *Líneas más cortas* (Lecciones populares de matemáticas) Mir, Moscú.