



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y  
DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

Una mirada a la física a través de la geometría de contacto  
y la geometría simpléctica

Tesis  
QUE PARA OPTAR POR EL GRADO DE:  
MAESTRO (A) EN CIENCIAS

PRESENTA:  
Luis Felipe López Reyes

DIRECTOR DE LA TESINA O TESIS  
Ana Rechtman Bulajich  
[Facultad de Ciencias](#)

Ciudad Universitaria, junio de 2018.  
Cd. Mx.



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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 esta 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 general

<b>1. Introducción</b>	<b>3</b>
<b>2. Geometría Simplética</b>	<b>5</b>
2.1. Geometría simpléctica . . . . .	5
2.2. El álgebra de Lie de las funciones suaves . . . . .	8
<b>3. Geometría de Contacto</b>	<b>11</b>
3.1. Hamiltonianos de Contacto . . . . .	14
3.2. Ecuaciones Diferenciales Parciales de Primer Orden . . . . .	16
<b>4. Mecánica Clásica</b>	<b>27</b>
4.1. Las leyes de Newton . . . . .	27
4.2. Leyes de conservación . . . . .	31
4.3. El principio de D'Alembert . . . . .	32
4.4. Las ecuaciones de Euler-Lagrange . . . . .	35
4.5. Las ecuaciones de Hamilton . . . . .	38
4.6. El método de Hamilton-Jacobi . . . . .	44
<b>5. Relatividad Especial</b>	<b>53</b>
5.1. El principio de relatividad . . . . .	54
5.2. Fundamentos de la relatividad especial . . . . .	55
5.3. La transformación de Lorentz . . . . .	58
5.4. Las ecuaciones de la Relatividad Especial . . . . .	62
5.5. Ondas electromagnéticas . . . . .	66
<b>6. Conclusiones</b>	<b>73</b>



# Capítulo 1

## Introducción

Desde que las matemáticas aparecen en la historia de la humanidad, el ser humano las ha utilizado para resolver muchísimos problemas prácticos tales como la repartición de terrenos, el cobro de tributos, la construcción de ciudades o el estudio de los cuerpos celestes. Debido a la belleza intrínseca de las matemáticas y al gran papel que han jugado en el desarrollo de las civilizaciones, los seres humanos se han dado cuenta de lo valioso que es estudiar y crear nuevas matemáticas para lograr entender problemas cada vez más complejos.

Históricamente la física siempre ha sido una mina de oro para las matemáticas. El ejemplo más conocido de esto sin duda es la teoría del cálculo diferencial cuyos métodos fueron utilizados por Isaac Newton para estudiar el movimiento de los cuerpos en el espacio. Newton descubrió una ecuación sencilla que relaciona la física del movimiento de cuerpos con el cálculo diferencial y en particular con las ecuaciones diferenciales. Cabe mencionar que Leibniz también desarrollo una teoría del cálculo diferencial.

A pesar de que en el siglo XVIII grandes científicos como Euler, Lagrange y Laplace aplicaron con éxito las ideas de Newton y Leibniz a diversos problemas, siempre hubo duda y escepticismo acerca de los fundamentos lógicos, físicos y filosóficos del cálculo diferencial. Por ejemplo, en 1734 el filósofo Irlandés publico El Analista, en este trabajo Berkeley no critica la utilidad del cálculo diferencial sino la falta de fundamentos en sus ideas y pruebas. Berkeley argumenta que el concepto físico de velocidad depende de intervalos espaciales y temporales, así, la velocidad instantánea, para la cual la longitud de dichos intervalos es cero, no es una noción física [9]. Aún así, siempre prevaleció la idea de que las matemáticas eran el lenguaje con el cual habían sido escritas las leyes de la naturaleza y por lo tanto ciertos conceptos matemáticos tenían una estrecha relación con la física de la naturaleza.

Con los trabajos para fundamentar el cálculo diferencial y en general a las matemáticas, la visión de ellas comenzó a cambiar. Se empezó a concebir a las matemáticas como un conjunto de proposiciones que pueden ser derivadas, independientemente de la física, a partir de un sistema axiomático. En su problema número 6 David Hilbert se pregunta acerca de la posibilidad de crear un

sistema axiomático para la física. Finalmente, la visión de las matemáticas cambió radicalmente cuando Godel demostró que cualquier sistema axiomático lo suficientemente poderoso para describir la aritmética de los números naturales, o bien tiene contradicciones, o tiene proposiciones que no son demostrables a partir de los axiomas.

Lo anterior aunado a que la física estudia problemas cada vez más complejos, los cuales requieren de un mayor conocimiento previo y simplificaciones para estudiarse de manera matemática, ha causado que por fin las matemáticas y la física tomen caminos separados.

Los objetivos de este trabajo son: describir las nociones físicas básicas, tanto de la mecánica clásica como de la relatividad especial y, ver como se pueden usar la geometría de contacto y simpléctica para estudiar las ecuaciones de movimiento en estas dos ramas de la física. Se tiene particular cuidado en el proceso que hay que seguir para, a partir de la segunda ley de Newton, obtener unas ecuaciones equivalentes a esta ley. En caso de que las fuerzas que aparecen en esta ley sean conservativas estas ecuaciones describen un campo vectorial en una variedad simpléctica, si a estas fuerzas añadimos fuerzas de fricción proporcional a la velocidad de la partícula las ecuaciones describen un campo vectorial en una variedad de contacto. En el caso de la relatividad especial también se hace énfasis en la importancia que adquiere el campo de fuerzas para la descripción de las ecuaciones de movimiento. Esta importancia se traduce en que el campo de fuerzas tiene que satisfacer las ecuaciones de Maxwell. Bajo ciertas condiciones sobre este campo veremos que se puede usar a la geometría de contacto para estudiar la manera en la que este campo se propaga en el espacio. Las ideas básicas de la geometría simpléctica y de contacto se describen en los capítulos 2 y 3 respectivamente, en los capítulos 4 y 5 veremos como usar estas teorías en la mecánica clásica y la relatividad especial respectivamente. En este trabajo no se asume ningún prerrequisito de física y tampoco se hace un análisis profundo de estas ideas. Respecto a la parte matemática sólo se requieren los resultados básicos de formas diferenciables.

## Capítulo 2

# Geometría Simpléctica

### 2.1. Geometría simpléctica

En este capítulo discutimos las bases de la geometría y simpléctica y los sistemas Hamiltonianos. Los sistemas Hamiltonianos sirven para describir sistemas mecánicos en los cuales la energía se conserva. Estos sistemas están definidos en variedades simplécticas. El espacio de funciones de una variedad simpléctica es un álgebra de Lie cuyo corchete está relacionado con la razón de cambio de funciones a lo largo de sistemas Hamiltonianos. Veremos como ciertas propiedades de esta álgebra de Lie se traducen en propiedades geométricas de los sistemas Hamiltonianos.

**Definición 1.** Sean  $M$  una variedad diferenciable y  $\omega \in \Lambda^2 M$  una 2-forma diferencial. Diremos que  $(M, \omega)$  es una *variedad simpléctica* si,  $\omega$  es cerrada ( $d\omega = 0$ ) y no degenerada, esto último quiere decir que para todo punto  $p \in M$  y todo vector  $\xi \in T_p M$  no nulo, existe un vector  $\eta \in T_p M$  tal que  $\omega(\xi, \eta) \neq 0$ .

Recordemos que para cada  $p \in M$ , una 2-forma  $\omega$  define una aplicación alternante y bilineal  $\omega_p: T_p M \times T_p M \rightarrow \mathbb{R}$ . Como primera observación tenemos que si  $(M, \omega)$  es una variedad simpléctica entonces, para cada  $p \in M$  podemos definir una transformación lineal

$$I: T_p M \rightarrow (T_p M)^*$$

como  $I(\xi) = \omega(\cdot, \xi)$  el hecho de que  $\omega$  sea no degenerada implica que  $I$  es un isomorfismo.

Si escogemos un producto interior en  $T_p M$ , como  $\omega$  es alternante, la matriz de  $I$  respecto a una base ortogonal es antisimétrica, así,  $\det(I) = \det I^t = \det(-I) = (-1)^n \det(I)$ . Por lo tanto, no puede haber variedades simplécticas de dimensión impar.

**Ejemplo 1.** El espacio vectorial  $\mathbb{R}^{2n}$  con las coordenadas usuales  $(p_1, \dots, p_n,$

$q_1, \dots, q_n$ ) equipado de la 2-forma

$$\omega = \sum_{i=1}^n dp_i \wedge dq_i, \quad (2.1)$$

es una variedad simpléctica. En efecto, tenemos que  $\omega = d(\sum p_i dq_i)$  y en consecuencia  $d\omega = 0$ . Además,

$$\omega \left( \frac{\partial}{\partial p_j}, \frac{\partial}{\partial q_j} \right) = \det \begin{pmatrix} dp_j \left( \frac{\partial}{\partial p_j} \right) & dp_j \left( \frac{\partial}{\partial q_j} \right) \\ dq_j \left( \frac{\partial}{\partial p_j} \right) & dq_j \left( \frac{\partial}{\partial q_j} \right) \end{pmatrix} = 1$$

como la condición de ser no degenerada es válida en una base concluimos que  $\omega$  es no degenerada.

**Ejemplo 2.** El espacio cotangente  $T^*M$  de una variedad diferenciable  $M$  es una variedad simpléctica. Para ver esto consideremos el siguiente diagrama,

$$\begin{array}{ccc} T(T^*M) & \xrightarrow{\pi} & T^*M \\ D\Pi \downarrow & & \downarrow \\ TM & \longrightarrow & M \end{array}$$

si  $v \in TM$  y  $w^* \in T^*M$  tienen el mismo punto base en  $M$  podemos tomar su producto  $\langle v, w^* \rangle$ , así, definimos  $\theta(\xi) := \langle D\Pi(\xi), \pi(\xi) \rangle$ . Sean  $\mathbf{q} = (q_1, \dots, q_n)$  coordenadas locales de  $M$  y  $(\mathbf{q}, \mathbf{p}) = (q_1, \dots, q_n, p_1, \dots, p_n)$  las coordenadas de  $T^*M$  asociadas a las coordenadas  $\mathbf{q}$  de  $M$ . Un vector  $\xi$  con punto base  $(\mathbf{q}, \mathbf{p})$  se escribe de la forma

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

Por un lado,  $D\Pi(\xi) = (\mathbf{q}, \sum A_i \frac{\partial}{\partial q_i})$  mientras que, si  $\mathbf{p} = \sum p_i dq_i$ ,  $\pi(\xi) = (\mathbf{q}, \mathbf{p})$  de forma tal que,

$$\theta(\xi) = \sum_{i=1}^n A_i p_i.$$

Notemos que  $\theta(\frac{\partial}{\partial p_i}) = 0$  y  $\theta(\frac{\partial}{\partial q_i}) = p_i$  y por lo tanto, en las coordenadas  $(\mathbf{q}, \mathbf{p})$  la 1-forma  $\theta$  se escribe,

$$\theta = \sum_{i=1}^n p_i dq_i =: \mathbf{p} d\mathbf{q}.$$

La expresión anterior si define un elemento en  $T^*(T^*M)$  pues este espacio, de manera local, se descompone como  $T^*(U) \times T^*(\mathbb{R}^n)$  y arriba hemos visto que  $\theta$  se anula en la parte correspondiente a  $T^*(\mathbb{R}^n)$ .

El ejemplo anterior muestra que si definimos  $\omega = d(\theta)$  entonces  $\omega$  es cerrada y como en cada punto es no degenerada concluimos que  $(M, \omega)$  es una variedad simpléctica.

**Definición 2.** Sean  $(M, \omega)$  y  $(N, \tilde{\omega})$  dos variedades simplécticas, una función diferenciable  $f: M \rightarrow N$  es un *simplectomorfismo* si  $f^*(\tilde{\omega}) = \omega$ .

El teorema de Darboux asegura que alrededor de cualquier punto en una variedad simpléctica, existen coordenadas  $(\mathbf{p}, \mathbf{q})$  tales que la 2-forma  $\omega$  se escribe  $\omega = d(\sum_{i=1}^n p_i dq_i)$ , a unas coordenadas con estas características se les llama *coordenadas simplécticas*, para ver un prueba de este resultado ver [1]. El resultado anterior implica que cualesquiera dos variedades simplécticas son localmente simplectomorfas, esto es, hay un difeomorfismo local que lleva la 2-forma de una variedad en la 2-forma de la otra.

El isomorfismo  $I^{-1}: T^*M \rightarrow TM$  nos permite asociar a cada función suave  $H: M \rightarrow \mathbb{R}$  un campo vectorial  $X_H$  definido como  $X_H := I^{-1}(dH)$ . Vamos a calcular la matriz de  $I$  en coordenadas simplécticas,

$$I \left( \frac{\partial}{\partial p_i} \right) \left( \frac{\partial}{\partial p_j} \right) = \omega \left( \frac{\partial}{\partial p_j}, \frac{\partial}{\partial p_i} \right) = \sum_{i=1}^n dp_i \wedge dq_i \left( \frac{\partial}{\partial p_j}, \frac{\partial}{\partial p_i} \right) = 0,$$

y por otro lado,

$$\begin{aligned} I \left( \frac{\partial}{\partial p_i} \right) &= \omega \left( -, \frac{\partial}{\partial p_i} \right) = dp_i \wedge dq_i \left( -, \frac{\partial}{\partial p_i} \right) \\ &= dp_i dq_i \left( \frac{\partial}{\partial p_i} \right) - dp_i \left( \frac{\partial}{\partial p_i} \right) dq_i = -dq_i. \end{aligned}$$

Lo anterior nos dice que,

$$I \left( \frac{\partial}{\partial p_i} \right) = -dq_i,$$

por la antisimetría de  $\omega$  tenemos que,

$$I \left( \frac{\partial}{\partial q_i} \right) = dp_i.$$

Por lo tanto,

$$X_H = I^{-1} \left( \sum_{i=1}^n \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial q_i} dq_i \right) = \sum_{i=1}^n -\frac{\partial H}{\partial q_i} \frac{\partial}{\partial p_i} + \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i}$$

Así, la ecuación diferencial asociada al campo vectorial  $X_H$  es

$$\begin{aligned} \dot{p}_i &= -\frac{\partial H}{\partial q_i} \\ \dot{q}_i &= \frac{\partial H}{\partial p_i}. \end{aligned} \tag{2.2}$$

La ecuación diferencial anterior es muy importante pues, como veremos más adelante, muchos sistemas físicos se pueden describir mediante esta ecuación en donde  $H$  es la energía del sistema.

**Definición 3.** Sea  $(M, \omega)$  una variedad simpléctica y  $H: M \rightarrow \mathbb{R}$  una función diferenciable. A la tripleta  $(M, \omega, H)$  le llamaremos *sistema Hamiltoniano* y al flujo del campo vectorial  $X_H$ , donde  $dH = \omega(\cdot, X_H)$ , le llamaremos *flujo hamiltoniano*. En todo este capítulo vamos a asumir que el flujo  $g_H^t$  del campo vectorial  $X_H$  es completo, esto es  $g_H^t$  existe para todo  $t \in \mathbb{R}$ .

El campo vectorial  $X_H$  tiene algunas propiedades importantes, primero que nada,  $dH = I(X_H) = \omega(\cdot, X_H)$  y por lo tanto  $dH(X_H) = 0$ . Esto quiere decir que la función  $H$  permanece constante a lo largo de las curvas que son solución a la ecuación (2.2) o dicho de otro modo, estas curvas están contenidas en las superficies de nivel de la función  $H$ .

Usando que  $\omega$  es cerrada, la definición de  $X_H$  y la fórmula de Cartan (ver [6]),

$$\mathcal{L}_{X_H}\omega = i_{X_H}(d\omega) + d(i_{X_H}\omega) = -ddH,$$

donde  $i_X\omega = \omega(X, \cdot)$ , y así obtenemos,

**Proposición 2.1.** Sea  $(M, \omega, H)$  un sistema Hamiltoniano, entonces  $\mathcal{L}_{X_H}\omega = 0$ .

## 2.2. El álgebra de Lie de las funciones suaves

En muchos problemas físicos se tiene interés por varias funciones que dependen de la posición que ocupa el sistema, por ejemplo, la distancia a un punto, el ángulo que se forma con cierta dirección, la energía, el trabajo o las funciones coordenadas de un sistema de referencia particular. Así pues, dado un sistema Hamiltoniano  $(M, \omega, H)$  y  $f \in C^\infty(M)$  una función diferenciable, definimos la *derivada total* de  $f$  como,

$$\frac{df(x)}{dt} = \left. \frac{d(f \circ g_H^t(x))}{dt} \right|_{t=0} = df(X_H),$$

donde  $g_H^t: M \rightarrow M$  es el flujo del campo vectorial  $X_H$ . La expresión anterior nos dice como cambia la cantidad  $f$  a lo largo de las soluciones del sistema Hamiltoniano. Podemos escribir la derivada total como:

$$\frac{df}{dt} = X_H f = df(X_H) = \omega(X_H, X_f) = -\omega(X_f, X_H),$$

y por lo tanto, en coordenadas simplécticas,

$$\frac{df}{dt} = \sum_{i=1}^n \frac{\partial H}{\partial p_i} \frac{\partial f}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial f}{\partial p_i}.$$

La expresión anterior es una ecuación diferencial (parcial) para la función  $f$  y en particular, si  $f = q_i$  o  $f = p_i$  obtenemos las ecuaciones (2.2).

**Definición 4.** Sea  $(M, \omega, H)$  un sistema Hamiltoniano y  $f, g \in C^\infty(M)$  dos funciones diferenciables. Definimos el *corchete de Poisson* de  $f$  y  $g$  como  $\{f, g\} = \omega(X_f, X_g)$ . En coordenadas simplécticas  $(\mathbf{p}, \mathbf{q})$  tenemos que:

$$\{f, g\} = \sum_{i=1}^n \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i}. \quad (2.3)$$

Como  $X_f = I^{-1}(df)$  y tanto el operador  $d$  como  $I$  son  $\mathbb{R}$ -lineales, el corchete de Poisson  $\{, \}$ :  $C^\infty(M) \times C^\infty(M) \rightarrow C^\infty(M)$  es  $\mathbb{R}$ -lineal. El operador  $d$  no es  $C^\infty(M)$ -lineal pero, como  $d(fg) = f dg + g df$  obtenemos que el corchete de Poisson satisface la regla de Leibniz:

$$\{h, fg\} = f\{h, g\} + \{h, f\}g.$$

Esta propiedad quiere decir que la expresión  $\{h, \cdot\}$  es un campo vectorial tangente a  $M$ , en coordenadas simplécticas

$$\begin{aligned} \{h, p_i\} &= -\frac{\partial h}{\partial q_i} \\ \{h, q_i\} &= \frac{\partial h}{\partial p_i}, \end{aligned}$$

y por lo tanto,  $\{h, \cdot\} = X_h$ . Claramente  $\{, \}$  es antisimétrica, la última propiedad que cumple el corchete de Poisson es la identidad de Jacobi (ver [1]):

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0. \quad (2.4)$$

Una consecuencia de la identidad de Jacobi es la siguiente, la asociación  $f \mapsto X_f = I^{-1}(df)$  es  $\mathbb{R}$ -lineal, el conjunto de campos vectoriales tiene un corchete de Lie  $[X, Y]$ , la identidad de Jacobi implica que la asociación anterior es un morfismo de álgebras de Lie. Para ver esto vamos a mostrar que  $[X_f, X_g]h = \{\{f, g\}, h\}$  y en consecuencia  $\{\{f, g\}, \cdot\} = X_{\{f, g\}} = [X_f, X_g]$ . Usando que  $X_g = dh(X_g) = \{g, h\}$  obtenemos:

$$\begin{aligned} [X_f, X_g]h &= X_f X_g h - X_g X_f h \\ &= X_f \{g, h\} - X_g \{f, h\} \\ &= \{f, \{g, h\}\} + \{g, \{h, f\}\}, \end{aligned}$$

por lo tanto, usando la identidad de Jacobi,  $[X_f, X_g]h = -\{h, \{f, g\}\} = \{\{f, g\}, h\}$ , así los vectores  $[X_f, X_g]$  y  $\{\{f, g\}, \cdot\}$  son el mismo vector.

Antes vimos que el Hamiltoniano  $H$  es constante a lo largo del flujo del campo vectorial  $X_H$ . Si  $f$  es una función suave,

$$\frac{df}{dt} = \omega(X_H, X_f) = \{H, f\},$$

y por ende, la cantidad  $f$  se conserva si y sólo si  $\{H, f\} = 0$ .

**Definición 5.** Sea  $(M, \omega, H)$  un sistema Hamiltoniano, a una función  $f$  tal que  $\{H, f\} = 0$  la llamaremos *constante de movimiento* y de manera más general diremos que dos funciones  $f, g \in C^\infty(M)$  están en *involución* si  $\{f, g\} = 0$ .

Otra consecuencia importante de la identidad de Jacobi es que si  $f$  y  $g$  son constantes de movimiento, entonces  $\{f, g\}$  también es una constante de movimiento, en efecto por la identidad de Jacobi,

$$-\{H, \{f, g\}\} = \{f, \{g, H\}\} + \{g, \{H, f\}\} = 0.$$

El hecho de que la asociación  $f \mapsto X_f$  es un morfismo de álgebras de Lie tiene una importante consecuencia geométrica, si  $\{f, g\} = 0$  entonces los flujos de los campos vectoriales  $X_f$  y  $X_g$  conmutan.

**Teorema 2.1.** Sean  $(M^{2n}, \omega, H)$  un sistema Hamiltoniano y  $g_H^t =$  el flujo de este sistema. Si existen  $n$  constantes del movimiento  $H = f_1, \dots, f_n \in C^\infty(M)$  tales que  $\{f_i, f_j\}$  entonces, si  $c \in \mathbb{R}$  es un valor regular de  $F = (f_1, \dots, f_n)$ , para todo  $x \in M_c := F^{-1}(c)$  tenemos que  $g_H^t(x) \in M_c$  y existen coordenadas  $(u_1, \dots, u_n)$  de  $M_c$  alrededor del punto  $x$  tales que  $X_H = \frac{\partial}{\partial u_1}$ .

*Demostración.* Como cada  $f_i$  es constante a lo largo del flujo de  $X_H$  entonces  $F = (f_1, \dots, f_n)$  es constante a lo largo del flujo de  $X_H$ . Esto implica que si  $x \in M_c$  entonces  $g_H^t(x) \in M_c$ .

Definimos una parametrización  $g: U \rightarrow V \subset M_c$  de  $M_c$  alrededor de  $x$  de la siguiente manera,  $g(u_1, \dots, u_n) = g_{f_n}^{u_n} \circ \dots \circ g_{f_1}^{u_1}$  con  $U$  una vecindad suficientemente pequeña del  $0 \in \mathbb{R}^n$ . El hecho que  $0 = \{f_i, f_j\} = [X_{f_i}, X_{f_j}]$  implica que  $g$  está bien definida. Si pensamos a cada  $f_i$  como otro Hamiltoniano, la misma hipótesis  $\{f_i, f_j\}$  nos dice que los flujos  $g_{f_i}^s$  preservan la curva de nivel  $M_c$  y por lo tanto  $g(u) \in M_c$  para toda  $u \in U$ . Por la definición de la parametrización tenemos que  $\frac{\partial g}{\partial u_i} = X_{f_i}$ . Como  $c$  es un valor regular de  $F$  entonces  $df_1, \dots, df_n$  son linealmente independientes y así  $g$  es efectivamente una parametrización de  $M_c$  alrededor del punto  $x$ . Bajo esta parametrización  $\frac{\partial}{\partial u_1} \mapsto X_{f_1} = X_H$ .  $\square$

Si añadimos la hipótesis de que alguna de las hipersuperficies definidas por  $f_i = c$  es compacta entonces se puede mostrar que  $M_c$  es homeomorfo a un toro. Más aún, se pueden construir coordenadas simplécticas  $(\varphi_1, \dots, \varphi_n, J_1, \dots, J_n)$  tales que las coordenadas  $J_i$  parametrizan a los distintos toros y las coordenadas  $\varphi_i$  son coordenadas angulares en los toros. Para ver una prueba de esto ver [1].

## Capítulo 3

# Geometría de Contacto

Hemos visto que no pueden existir variedades simplécticas de dimensión impar. Las variedades de contacto son el análogo a las variedades simplécticas pero en dimensión impar. En este capítulo estudiaremos sistemas Hamiltonianos para variedades de contacto. Estos sistemas son el análogo de los sistemas Hamiltonianos para variedades simplécticas. También veremos la utilidad de la geometría de contacto para resolver ecuaciones diferenciales parciales de primer orden. Para definir a una variedad de contacto necesitamos de algunos preliminares.

**Definición 6.** Sea  $M$  una variedad diferenciable. Una *distribución de hiperplanos* es un subhaz  $\xi$  de codimensión 1 del haz tangente  $TM$ .

El haz vectorial  $TM/\xi$  es de dimensión 1 y, para cada  $p \in M$ , si  $U \subseteq M$  es una vecindad de  $p$  en la cual el haz  $TM/\xi$  es trivial, tomemos  $\alpha'_U \in \Gamma((TM/\xi)_U^*)$  una sección local que no se anula (el argumento de  $\alpha'_U$  son secciones del haz  $(TM/\xi)_U$ ) y definimos la 1-forma  $\alpha_U$  como  $\alpha_U(v) := \alpha'_U \circ \Pi(v)$ , donde  $\Pi: TM \rightarrow TM/\xi$  es la proyección canónica. Como  $TM/\xi$  es de dimensión 1 punto a punto  $\alpha'_U$  define un isomorfismo y por lo tanto  $\ker(\alpha_U) = \ker(\Pi|_U) = \xi_U$ . Así, de manera local, toda distribución de hiperplanos está dada por el kernel de una 1-forma.

La 1-forma  $\alpha$  que define de manera local a la distribución de hiperplanos no es única y, la distribución está definida por una 1-forma global si y solamente si el haz  $TM/\xi$  es trivial. En efecto, si dicha 1-forma global  $\alpha$  existe, como  $\ker(\alpha_p) = \xi_p$ , la función de  $TM$  en  $M \times \mathbb{R}$  dada por  $(p, v) \mapsto (p, \alpha_p(v))$  induce un isomorfismo de  $TM/\xi$  en  $M \times \mathbb{R}$ . Para la afirmación recíproca notemos que en la construcción anterior, si  $TM/\xi$  es trivial podemos tomar  $U = M$ . Por último, notemos que si en un espacio vectorial  $V$  tomamos  $\alpha, \beta \in V^*$  tales que  $\ker(\alpha) = \ker(\beta) \neq V$  entonces, tomando  $x \notin \ker(\beta)$  tenemos que  $V = \ker(\beta) \oplus \langle x \rangle$  y haciendo  $\lambda = \frac{\alpha(x)}{\beta(x)}$  obtenemos  $(\alpha + \lambda\beta)x = 0$  y por lo tanto  $\alpha = \lambda\beta$ .

Supongamos que en una carta local  $U \subseteq M$  la distribución de hiperplanos

queda definida por las 1-formas  $\alpha$  y  $\beta$ , por lo anterior, existe una función diferenciable  $\lambda: U \rightarrow \mathbb{R}$  que no se anula y es tal que  $\alpha = \lambda\beta$ . Así,  $d\alpha = d\lambda \wedge \beta + \lambda \wedge d\beta$  y por lo tanto,

$$d\alpha|_{\xi} = \lambda d\beta|_{\xi}.$$

En particular,  $\text{ran}(d\alpha) = \text{ran}(d\beta)$ .

**Definición 7.** Sea  $M$  una variedad diferenciable y  $\xi$  una distribución de hiperplanos. Diremos que  $(M, \xi)$  es una *variedad de contacto* si para cualquier carta  $U \subseteq M$  en la cual la distribución  $\xi$  está definida por la 1-forma  $\alpha$  tenemos que  $d\alpha|_{\xi}$  tiene rango máximo.

Si la variedad  $M$  tiene dimensión par entonces los hiperplanos  $\xi$  son de dimensión impar y por lo tanto la 2-forma  $d\alpha$  no puede tener rango máximo, así, las variedades de contacto son de dimensión impar.

**Proposición 3.1.** *Sea  $M$  una variedad diferenciable y  $\xi$  una distribución de hiperplanos definida por el kernel de la 1-forma  $\alpha$ . Son equivalentes:*

1.  $d\alpha|_{\xi}$  tiene rango máximo.
2.  $\alpha \wedge (d\alpha)^n$  es una forma de volumen.

*Demostración.* Primero que nada, si  $v_1, \dots, v_{2n+1}$  son vectores tenemos que

$$\begin{aligned} \alpha \wedge (d\alpha)^n(v_1, \dots, v_{2n+1}) &= \\ &= \sum_{\sigma \in S_{2n+1}} \text{sgn } \sigma \alpha(v_{\sigma(1)}) d\alpha(v_{\sigma(2)}, v_{\sigma(3)}) \cdots d\alpha(v_{\sigma(2n)}, v_{\sigma(2n+1)}). \end{aligned}$$

Supongamos que  $d\alpha|_{\xi}$  es no degenerada y sea  $v_1, v_2, \dots, v_{2n-1}, v_{2n}$  una base simpléctica de forma tal que para  $k$  un número impar  $d\alpha(v_k, v_{k+1}) = 1$  y sea  $v_{2n+1}$  un vector que completa al conjunto anterior para formar una base. Como  $v_1, \dots, v_{2n} \in \xi = \ker(\alpha)$  tenemos que

$$\begin{aligned} \alpha \wedge (d\alpha)^n(v_1, \dots, v_{2n+1}) &= \\ &= \sum_{\sigma(1)=2n+1} \text{sgn } \sigma \alpha(v_{2n+1}) d\alpha(v_{\sigma(2)}, v_{\sigma(3)}) \cdots d\alpha(v_{\sigma(2n)}, v_{\sigma(2n+1)}). \end{aligned}$$

Como los vectores  $v_1, \dots, v_{2n}$  son una base simpléctica de  $\xi$ , para que uno de los sumandos anteriores no se anule, cada uno de los factores  $d\alpha(v_{\sigma(i)}, v_{\sigma(i+1)})$  tiene que ser de la forma  $d\alpha(v_{\sigma(k)}, v_{\sigma(k)+1})$  si  $\sigma(k)$  es impar o de la forma  $d\alpha(v_{\sigma(k)}, v_{\sigma(k)-1})$  si  $\sigma(k)$  es par. En caso de que  $\sigma(k)$  sea par al cambiar el orden de los argumentos de  $d\alpha$  obtenemos un signo menos pero la nueva permutación obtenida al hacer este cambio también arroja un signo menos así, podemos llevar a cada uno de los sumandos a la forma

$$\alpha(v_{2n-1}) d\alpha(v_1, v_2) \cdots d\alpha(v_{\sigma(2n-1)}, v_{\sigma(2n)}),$$

y por lo tanto  $\alpha \wedge (d\alpha)^n(v_1, \dots, v_{2n+1}) \neq 0$  y así  $\alpha \wedge (d\alpha)^n$  es una forma de volumen. Si ahora suponemos que  $\alpha \wedge (d\alpha)^n$  es una forma de volumen, sean  $v_1, \dots, v_{2n}$  una base de  $\xi$  y  $v_{2n+1}$  un vector que completa a este conjunto para formar una base de todo el espacio. Nuevamente

$$0 \neq \alpha \wedge (d\alpha)^n(v_1, \dots, v_{2n+1}) = \sum_{\sigma(1)=2n+1} \text{sgn } \sigma \alpha(v_{2n+1}) d\alpha(v_{\sigma(2)}, v_{\sigma(3)}) \cdots d\alpha(v_{\sigma(2n)}, v_{\sigma(2n+1)}),$$

y entonces al menos uno de los sumandos no se anula. Para que esto suceda cada uno de los factores  $d\alpha(v_{\sigma(k)}, v_{\sigma(k+1)})$  no se debe de anular. Así, para cada vector  $v_i$  con  $i = 1 \dots, 2n$  existe un vector  $v_l$  tal que  $d\alpha(v_k, v_l) \neq 0$  y por lo tanto  $d\alpha|_{\xi}$  es no degenerada.  $\square$

Así, una manera alternativa de definir una variedad de contacto es pidiendo que para cualquier 1-forma  $\alpha$  que defina a la distribución de hiperplanos de manera local, se tenga que  $\alpha \wedge (d\alpha)^n \neq 0$ .

**Ejemplo 3.** Sea  $M$  una variedad diferenciable, a un hiperplano en un espacio tangente  $T_q M$  se le conoce como un *elemento de contacto*. Afirmamos que el conjunto de todos los elementos de contacto es una variedad de contacto. Primero veamos que este conjunto coincide con la proyectivización del espacio cotangente de  $M$ ,  $\mathbb{P}T^*M$ . En efecto, si  $V$  es un hiperplano en  $T_q M$  entonces  $V = \ker(\alpha)$  con  $\alpha \in T_q^* M$  y  $\lambda\alpha$  tiene el mismo kernel que  $\alpha$  si  $\lambda \neq 0$ . Así, podemos identificar a cada hiperplano en  $T_q M$  con un elemento de  $\mathbb{P}T_q^* M$  y claramente cada elemento de este conjunto define un hiperplano en  $T_q M$ . Al proyectivizar cada uno de los espacios  $T_q^* M$  obtenemos el haz fibrado  $\mathbb{P}T^* M$ . Este espacio viene equipado de manera natural con una distribución de hiperplanos. Si  $q_1, \dots, q_n$  son coordenadas locales en  $U \subseteq M$  y  $q_1, \dots, q_n, p_1, \dots, p_n$  son las coordenadas inducidas en  $T^* M$  consideremos los abiertos  $U_i$  en los cuales  $p_i \neq 0$ . Si  $q: \mathbb{P}T^* M \rightarrow M$  es la aplicación cociente entonces podemos definir las cartas usuales  $\varphi_i: q(U_i) \rightarrow \mathbb{R}^{2n-1}$  del proyectivo mediante

$$(q_1, \dots, q_n, [p_1 : \dots : p_n]) \mapsto \left( q_1, \dots, q_n, \frac{p_1}{p_i}, \dots, \frac{p_n}{p_i} \right).$$

Cuando  $p_i \neq 0 \neq p_j$  entonces el cambio de coordenadas viene dado por  $v_l = \frac{u_l}{u_j}$  si  $l \neq i$  y  $v_i = \frac{1}{u_j}$ . Para cada carta  $U_j$  consideremos la 1-forma

$$\alpha_j = dq_j + \sum_{l=1, l \neq j}^n v_l dq_l.$$

La derivada exterior de esta 1-forma tiene rango maximal, por lo tanto, si en las distintas cartas  $q(U_i)$  el kernel de estas 1-formas coinciden, entonces ellas definen una estructura de contacto. Usando la expresión del cambio de coordenadas en la intersección  $q(U_i \cap U_j)$  tenemos que,

$$u_j \alpha_j = \alpha_i.$$

Veamos ahora que esta estructura de contacto no puede ser definida por una 1-forma global. Si este fuera el caso entonces tendríamos una 1-forma global  $\alpha$  y funciones diferenciables nunca nulas  $g_i$  definidas en los abiertos  $q(U_i)$  tales que  $g_i\alpha|_{U_i} = \alpha_i$  y  $g_j\alpha|_{U_j} = \alpha_j$ . Sustituyendo esto en la igualdad anterior obtenemos  $u_jg_j = g_i$ . Aplicando esto mismo pero ahora a la igualdad  $u_i\alpha_i = \alpha_j$  obtenemos  $u_i g_i = g_j$  y por lo tanto  $u_i u_j = 1$  lo cual no puede ser. Así pues vemos que no todas las estructuras de contacto pueden ser definidas por una 1-forma global. Al conjunto  $\mathbb{P}T^*M$  se le conoce como la *la variedad de elementos de contacto*.

Cabe mencionar que, al igual que en la geometría simpléctica, en la geometría de contacto no hay invariantes locales. Esto quiere decir que cualesquiera dos variedades de contacto son localmente difeomorfas mediante un difeomorfismo que manda la estructura de contacto de una en la estructura de contacto de la otra. Para ver esto consultar [1].

### 3.1. Hamiltonianos de Contacto

En el capítulo anterior, para cada sistema Hamiltoniano  $(M, \omega, H)$  definimos un campo vectorial  $X_H$  tal que la función  $H: M \rightarrow \mathbb{R}$  es una constante de movimiento ( $H$  se conserva a lo largo del flujo de  $X_H$ ). En la mecánica clásica los sistemas Hamiltonianos se utilizan para modelar sistemas de partículas en los cuales la energía  $H$  se conserva. En esta sección vamos a hacer algo análogo para una terna  $(M, \alpha, H)$  en donde  $M$  es una variedad de contacto cuya estructura de contacto está globalmente definida por el kernel de la 1-forma  $\alpha$ . En el siguiente capítulo vamos a ver que este tipo de sistemas se pueden utilizar para modelar sistemas con fuerzas de fricción que son proporcionales a la velocidad. Como las fuerzas de fricción disminuyen la velocidad y, si no aplicamos nuevas fuerzas la fricción eventualmente detiene a los objetos, la energía  $H$  no se puede conservar.

**Definición 8.** Sea  $(M, \xi)$  una variedad de contacto con estructura de contacto  $\xi = \ker \alpha$  con  $\alpha$  una 1-forma definida globalmente. Si  $H: M \rightarrow \mathbb{R}$  es una función diferenciable llamaremos *sistema Hamiltoniano de contacto* a la terna  $(M, \alpha, H)$ . Diremos que un campo vectorial  $X$  es *un campo vectorial de contacto* si  $\mathcal{L}_X \alpha = f\alpha$  para alguna función no nula  $f: M \rightarrow \mathbb{R}$ . Esto es equivalente a que  $(g_X^t)_*(\xi) = \xi$ , es decir, el flujo  $g_X^t$  del campo vectorial  $X$  preserva a la estructura de contacto (ver [3]).

En un sistema Hamiltoniano  $(M, \omega, H)$  el hecho de que  $\omega$  es una 2-forma no degenerada nos permite asociar a este sistema un campo vectorial  $X_H$  mediante la igualdad  $dH = \omega(\cdot, X_H)$ . En este caso tenemos:

**Definición 9.** Sea  $(M, \alpha, H)$  un sistema Hamiltoniano de contacto, definimos un campo vectorial  $X_H$  por las igualdades

$$\mathcal{L}_{X_H} \alpha = f_H \alpha \quad -H = \alpha(X_H), \quad (3.1)$$

en donde  $f_H: M \rightarrow \mathbb{R}$  es una función diferenciable.

En lo siguiente vamos a ver que estas igualdades efectivamente definen un campo vectorial. Usando la fórmula de Cartan tenemos que  $\mathcal{L}_{X_H}\alpha = d\alpha(X_H, \cdot) + d(\alpha(X_H)) = d\alpha(X_H, \cdot) - dH$  y por lo tanto,

$$d\alpha(X_H, \cdot) = f_H\alpha + dH \quad (3.2)$$

Esta es la igualdad análoga a  $dH = \omega(\cdot, X_H)$ . En este caso  $dH(X_H) = Hf_H$  y así la función  $H$  no es una constante de movimiento. Para determinar a la función  $f_H$  vamos a usar el campo de Reeb  $R_\alpha$  asociado a la 1-forma  $\alpha$ . Este campo vectorial está definido por las igualdades,

$$d\alpha(R_\alpha) \equiv 0 \quad \alpha(R_\alpha) = 1.$$

La primera igualdad no dice que  $R_\alpha \in \ker\alpha$  que por definición de variedad de contacto tiene dimensión 1. La segunda igualdad nos dice que elemento de  $\ker\alpha$  es el campo vectorial  $R_\alpha$ . De la igualdad (3.2) tenemos que  $f_H\alpha = d\alpha(X_H, \cdot) - dH$  y al evaluar en el campo de Reeb  $R_\alpha$  obtenemos  $f_H = -dH(R_\alpha)$ . Así, la función  $f_H$  está determinada de manera única por la 1-forma  $\alpha$  y la función  $H$ .

Veamos ahora si que las condiciones (3.1) determinan al campo vectorial  $X_H$ . Supongamos primero que  $f_H\alpha + dH \equiv 0$ , en este caso de la igualdad (3.2) obtenemos  $d\alpha(X_H, \cdot) \equiv 0$ . Por lo tanto  $X_H = \lambda R_\alpha$  con  $\lambda: M \rightarrow \mathbb{R}$  una función diferenciable y además  $-H = \alpha(X_H) = \alpha(\lambda R_\alpha) = \lambda$ . Supongamos ahora que  $f_H\alpha + dH \neq 0$ , vamos a obtener una expresión local para el campo  $X_H$ . Para esto vamos a hacer uso de un sistema coordenado análogo a las coordenadas simplécticas. En estas coordenadas  $(\mathbf{q}, \mathbf{p}, y)$ , llamadas *coordenadas de Darboux*, la 1-forma  $\alpha$  que define a la estructura de contacto se escribe como  $\alpha = dy - \sum_{j=1}^n p_j dq_j$  (ver [1]). Así,  $d\alpha = -\sum dp_j \wedge dq_j = \sum dq_j \wedge dp_j$ . Si en estas coordenadas

$$X_H = \sum_{i=1}^n \left( \dot{q}_i \frac{\partial}{\partial q_i} + \dot{p}_i \frac{\partial}{\partial p_i} \right) + \dot{y} \frac{\partial}{\partial y},$$

entonces el lado izquierdo de la igualdad (3.2) es

$$\begin{aligned} d\alpha(X_H, \cdot) &= \sum_{j=1}^n \left[ dq_j \wedge dp_j \left( \sum_{i=1}^n \dot{q}_i \frac{\partial}{\partial q_i}, \cdot \right) + dq_j \wedge dp_j \left( \sum_{i=1}^n \dot{p}_i \frac{\partial}{\partial p_i}, \cdot \right) \right] \\ &= \sum_{i=1}^n \sum_{j=1}^n \left( \dot{q}_i \left| dq_j \left( \frac{\partial}{\partial q_i} \right) \quad dp_j \left( \frac{\partial}{\partial q_i} \right) \right| \dot{p}_i \left| dq_j \left( \frac{\partial}{\partial p_i} \right) \quad dp_j \left( \frac{\partial}{\partial p_i} \right) \right| \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \dot{q}_i \delta_{ij} dp_j - \dot{p}_i \delta_{ij} dq_j = \sum_{i=1}^n \dot{q}_i dp_i - \dot{p}_i dq_i. \end{aligned}$$

El lado derecho de la igualdad (3.2) es,

$$f_H\alpha + dH = \sum_{i=1}^n \left( \left( \frac{\partial H}{\partial p_i} \right) dp_i + \left( \frac{\partial H}{\partial q_i} - f_H p_i \right) dq_i \right) + \left( \frac{\partial H}{\partial y} + f_H \right) dy.$$

Comparando ambos lados obtenemos  $f_H = -\frac{\partial H}{\partial y}$  y,

$$\begin{aligned}\dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\left(\frac{\partial H}{\partial q_i} + p_i \frac{\partial H}{\partial y}\right).\end{aligned}\tag{3.3}$$

Para obtener una expresión para  $\dot{y}$  notemos que  $dy = \alpha + \sum_i p_i dq_i$  y en consecuencia,

$$\dot{y} = dy(X_H) = \alpha(X_H) + \sum_{i=1}^n p_i dq_i = \sum_{i=1}^n p_i \frac{\partial H}{\partial p_i} - H.\tag{3.4}$$

Como en coordenadas de Darboux el kernel de la 2-forma  $\sum_i dq_i \wedge dp_i = d\alpha$  está generado por el vector  $\frac{\partial}{\partial y}$ , en estas coordenadas el campo de vector de Reeb  $R_\alpha$  coincide con el vector  $\frac{\partial}{\partial y}$ . Así, en el caso en que  $f_H \alpha + dH \equiv 0$  tenemos que las coordenadas del campo vectorial  $X_H$  son  $\dot{q}_i = \dot{p}_i = 0$  y  $\dot{y} = -H$ . En el caso simpléctico teníamos que el Hamiltoniano  $H$  se conserva a lo largo del flujo del campo vectorial  $X_H$ , en este caso de (3.2) tenemos que

$$dH(X_H) = -f_H \alpha(X_H) = -\frac{\partial H}{\partial y} H.\tag{3.5}$$

## 3.2. Ecuaciones Diferenciales Parciales de Primer Orden

En esta sección vamos a estudiar ecuaciones diferenciales para funciones  $u: M \rightarrow \mathbb{R}$ . Para esto vamos a considerar el espacio  $J^1(M) := T^*M \times \mathbb{R}$ , este espacio resulta ser una variedad de contacto y podemos pensar a una ecuación diferencial parcial como una función  $F: J^1(M) \rightarrow \mathbb{R}$ . Usando la estructura de contacto a cada ecuación parcial le vamos a asociar un campo de direcciones que va a ser muy importante en la construcción de soluciones a esta ecuación. En esta sección seguimos a [2].

**Definición 10.** Sea  $M$  una variedad diferenciable, *el espacio de 1-jets de  $M$*  es la variedad  $J^1(M) := T^*M \times \mathbb{R}$ . Una *ecuación diferencial parcial en  $M$*  es una función  $F: J^1(M) \rightarrow \mathbb{R}$ . Diremos que una función  $u: M \rightarrow \mathbb{R}$  es una solución de la ecuación diferencial determinada por  $F$  si para todo  $\mathbf{q} \in M$  tenemos que  $F(du_{\mathbf{q}}, u(\mathbf{q})) = 0$ . Si  $q_1, \dots, q_n$  son coordenadas locales de  $M$  entonces  $u$  es una solución si y sólo si  $F(q_1, \dots, q_n, \frac{\partial u}{\partial q_1}, \dots, \frac{\partial u}{\partial q_n}, u(q_1, \dots, q_n)) = 0$ .

A toda función diferenciable  $u: M \rightarrow \mathbb{R}$  le podemos asociar su *1-gráfica* que, en las coordenadas locales  $(\mathbf{q}, \mathbf{p})$  de  $T^*M$  inducidas por las coordenadas  $\mathbf{q} = (q_1, \dots, q_n)$  de  $M$  está definida por  $J^1(u) := \{(\mathbf{q}, \mathbf{p}, y) | y = u(\mathbf{q}), p_i = \frac{\partial u}{\partial q_i}\}$ . De manera global  $J^1(u) = \{du_{\mathbf{q}}, u(\mathbf{q})\}$ .

Tenemos que  $u: M \rightarrow \mathbb{R}$  es una solución de la ecuación diferencial determinada por  $F$  si y sólo si  $J^1(u) \subset \{F = 0\}$ . Usando las coordenadas locales

### 3.2. ECUACIONES DIFERENCIALES PARCIALES DE PRIMER ORDEN 17

$\mathbf{q}$  podemos parametrizar a  $J^1(u)$  mediante  $\mathbf{q} \mapsto^{\Phi} (\mathbf{q}, \frac{\partial u}{\partial q_1}, \dots, \frac{\partial u}{\partial q_n}, u(\mathbf{q}))$ . Si con esta carta local de  $M$  identificamos a un abierto de  $J^1(M)$  con  $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$  (si  $\Pi: T^*M \rightarrow M$  es la proyección y las coordenadas  $q_1, \dots, q_n$  están definidas en  $V \subset M$  entonces  $\Pi^{-1}(V) \cong \mathbb{R}^n \times \mathbb{R}^n$ ) entonces,

$$\frac{\partial \Phi}{\partial q_i} = \frac{\partial}{\partial q_i} + \sum_{j=1}^n \frac{\partial^2 u}{\partial q_i \partial q_j} \frac{\partial}{\partial p_j} + \frac{\partial u}{\partial q_i} \frac{\partial}{\partial y}. \quad (3.6)$$

Con esto vemos que los primeros  $n$  renglones de la matriz  $D\Phi$  forman a la matriz identidad y por lo tanto  $D\Phi$  tiene rango máximo, esto quiere decir que  $J^1(u)$  es una subvariedad diferenciable de  $J^1(M)$  de dimensión  $n$ .

En coordenadas locales  $(\mathbf{q}, \mathbf{p}, y)$  de  $J^1(M)$  definimos la 1-forma,

$$\alpha = dy - \sum_{i=1}^n p_i dq_i,$$

La importancia de esta 1-forma reside en el hecho de que si  $u: M \rightarrow \mathbb{R}$  es una función suave entonces  $du = \sum_{i=1}^n \frac{\partial u}{\partial q_i} dq_i$  y como a lo largo de la 1-gráfica  $J^1(u)$  tenemos que  $p_i = \frac{\partial u}{\partial q_i}$  entonces  $\alpha$  se anula a lo largo de  $J^1(u)$ , esto quiere decir que para todo  $x \in J^1(u)$ ,  $T_x J^1(u) \subset \ker(\alpha_x)$ .

Como el coeficiente de  $dy$  en la 1-forma  $\alpha$  es 1, entonces  $\ker(\alpha_x)$  se proyecta de manera biyectiva al espacio  $(\mathbf{q}, \mathbf{p})$ , usando a estas últimas coordenadas como coordenadas de  $\ker(\alpha_x)$  tenemos que  $d\alpha = -\sum_{i=1}^n dp_i \wedge dq_i$  y por lo tanto  $d\alpha$  tiene rango máximo. Esto quiere decir que en el abierto  $V \subset M$  en el cual están definidas las coordenadas  $q_i$ , la 1-forma  $\alpha$  define una estructura de contacto en  $\Pi^{-1}(V) \times \mathbb{R}$ .

A continuación vamos a ver que la construcción anterior define una forma de contacto en todo  $J^1(M)$ . La afirmación es que la distribución de hiperplanos  $\xi := \ker(\alpha)$  en  $\Pi^{-1}(V) \times \mathbb{R}$  es la cerradura de la unión de todos los espacios tangentes de 1-gráficas de funciones  $u: V \rightarrow \mathbb{R}$ . Ya hemos visto que  $TJ^1(u) \subset \xi$ , así

$$\bigcup_{u \in C^\infty(V)} TJ^1(u) \subset \xi,$$

como  $\xi$  es cerrado,

$$\overline{\bigcup_{u \in C^\infty(V)} TJ^1(u)} \subset \xi.$$

Para la otra contención, en coordenadas locales tenemos que el espacio tangente a  $J^1(u)$  está generado por los vectores  $\frac{\partial \Phi}{\partial q_i}$  de la expresión (3.6). Cada uno de estos  $n$  vectores está en la distribución  $\xi$ , notemos además que los vectores  $\frac{\partial}{\partial p_i}$  satisfacen  $\alpha(\frac{\partial}{\partial p_i}) = 0$ . Como la intersección del espacio tangente a  $J^1(u)$  y este último subespacio es el vector cero y ambos tienen dimensión  $n$  su suma es la distribución  $\xi$ . Así pues basta ver que podemos aproximar arbitrariamente bien

vectores de la forma  $\sum_{i=1}^n a_i \frac{\partial}{\partial p_i x}$  mediante vectores en  $T_x J^1(u)$  para alguna función  $u: V \rightarrow \mathbb{R}$ .

Si  $u, v: M \rightarrow \mathbb{R}$  son funciones diferenciables entonces  $T_x J^1(u), T_x J^1(v) \subset \xi_x$  y por la estructura lineal de este último espacio y la linealidad de la derivada, basta ver que podemos aproximar arbitrariamente bien a los vectores  $\frac{\partial}{\partial p_i x}$ . Supongamos que en nuestras coordenadas  $q_1, \dots, q_n$  el punto  $x$  está representado por el punto  $\bar{0} \in \mathbb{R}^n$  y consideremos la función  $u(q_1, \dots, q_n) = \frac{a}{2} q_i^2$ , de acuerdo a la expresión (3.6) el vector  $\frac{\partial}{\partial q_i x} + a \frac{\partial}{\partial p_i x}$  pertenece a  $T_x J^1(u)$ , si tomamos  $a \in \mathbb{R}$  lo suficientemente grande podemos hacer que el vector  $\frac{1}{a} \frac{\partial}{\partial q_i} + \frac{\partial}{\partial p_i}$  esté tan cercano como queramos al vector  $\frac{\partial}{\partial p_i}$ . Con lo anterior tenemos la

**Proposición 3.2.** *Sea  $M$  una variedad diferenciable y sea  $J^1(M)$  el espacio de 1-jets de  $M$ , entonces el conjunto*

$$\xi := \overline{\bigcup_{u \in C^\infty(M)} T J^1(u)}$$

define una estructura de contacto en  $J^1(M)$ .

**Definición 11.** Sea  $F: J^1(M) \rightarrow \mathbb{R}$  una ecuación diferencial parcial en  $M$ , diremos que  $F$  es no singular si para todo  $x \in \{F = 0\}$  tenemos que  $\dim(T_x \{F = 0\} \cap \xi_x) = 2n - 1$ . Esto es lo mismo que  $\{F = 0\}$  sea transversal a la distribución  $\xi$ . Si  $F$  es una ecuación parcial no singular entonces la estructura simpléctica de  $\xi$  nos permite asociar un campo de direcciones  $X$  definido como el complemento ortogonal de  $T\{F = 0\} \cap \xi$  respecto a la estructura simpléctica  $\omega$  de  $\xi$ . Llamaremos *características de la ecuación  $F$*  a las curvas integrales del campo de direcciones  $X$ .

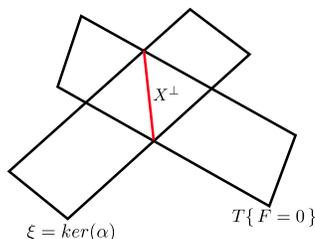


Figura 3.1: El espacio ortogonal a las direcciones de las características.

La importancia de las características queda ilustrada en la siguiente,

**Proposición 3.3.** *Sea  $u: M \rightarrow \mathbb{R}$  una solución a la ecuación diferencial parcial no singular determinada por  $F: J^1(M) \rightarrow \mathbb{R}$ . Entonces las características de la ecuación  $F$  son tangentes a la 1-gráfica  $J^1(u)$ .*

*Demostración.* Primero que nada, el campo de direcciones  $X$  de las características satisface  $\omega(X, X) = 0$ , esto quiere decir que  $X \subset X^\perp := T\{F = 0\} \cap \xi$ .

Segundo, como la derivada exterior conmuta con funciones diferenciables, si  $i: J^1(u) \rightarrow J^1(M)$  es la inclusión,  $i^*(\omega) = i^*(d\alpha) = d(i^*(\alpha)) = 0$  pues  $TJ^1(u) \subset \xi = \ker(\alpha)$ . A las subvariedades  $N$  tal que  $i^*(\omega) = 0$  se les conoce como subvariedades isotrópicas, veamos que una subvariedad isotrópica no puede tener dimensión mayor a  $n$ , en efecto, si  $W := TN^\perp$  entonces  $2n = \dim(N) + \dim(W)$  pero por la isotropía de  $N$  tenemos que  $TN \subset W$  y en consecuencia la dimensión de la subvariedad  $N$  no puede exceder a  $n$ . Por último tenemos que  $TJ^1(u) \subset \xi$  y  $TJ^1(u) \subset T\{F = 0\}$  y por ende  $TJ^1(u) \subset T\{F = 0\} \cap \xi = X^\perp$ , esto nos dice que  $TJ^1(u)$  es ortogonal al campo de direcciones  $X$ . Con todo lo anterior  $X$  está forzado a estar en  $TJ^1(u)$  pues en caso contrario  $X \oplus TJ^1(u)$  sería un subespacio isotrópico de dimensión  $n + 1$ .  $\square$

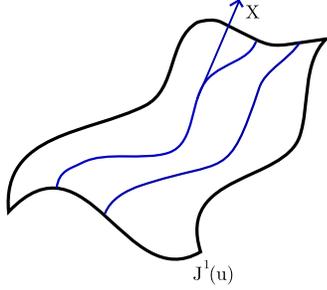


Figura 3.2: Las características son tangentes a la 1-gráfica  $J^1(u)$

**Definición 12.** Si  $W \subset M$  es un abierto tal que  $\partial W$  es una subvariedad suave, dados un dominio  $\Gamma \subseteq \partial W$ , una función suave  $u_0: \Gamma \rightarrow \mathbb{R}$  y una ecuación diferencial parcial  $F: J^1(M) \rightarrow \mathbb{R}$ , el *problema de Cauchy con condición inicial*  $u_0$  consiste en encontrar una solución  $u: \overline{W} \rightarrow \mathbb{R}$  a la ecuación diferencial determinada por  $F$  tal que  $u|_\Gamma = u_0$ .

En lo siguiente vamos a usar a las características de la ecuación  $F$  para construir una solución local al problema de Cauchy. Como primer paso vamos a calcular, en coordenadas locales, al campo vectorial que genera el campo de direcciones  $X$  de las características.

**Lema 3.1.** *En las coordenadas  $(\mathbf{q}, \mathbf{p}, y)$  de  $J^1(M)$  inducidas por las coordenadas locales  $\mathbf{q}$  de  $M$ , el campo de direcciones  $X$  está generado por el campo vectorial*

$$\sum_{i=1}^n \frac{\partial F}{\partial p_i} \frac{\partial}{\partial q_i} - \left( \frac{\partial F}{\partial q_i} + p_i \frac{\partial F}{\partial y} \right) \frac{\partial}{\partial p_i} + p_i \frac{\partial F}{\partial p_i} \frac{\partial}{\partial y}.$$

*Demostración.* Recordemos que como cada hiperplano de la distribución  $\xi$  se proyecta de manera biyectiva en el plano  $(\mathbf{q}, \mathbf{p})$ , las funciones  $\mathbf{q}, \mathbf{p}$  sirven como coordenadas de  $\xi$ , en ellas vamos a calcular  $T\{F = 0\} \cap \xi = X^\perp$ . Tomemos

un vector  $\sum_{i=1}^n \left( a_i \frac{\partial}{\partial q_i} + b_i \frac{\partial}{\partial p_i} \right) + c \frac{\partial}{\partial y}$  si este vector está en la distribución  $\xi = \ker(dy - \sum p_i dq_i)$  entonces

$$\sum p_i a_i = c.$$

Si este mismo vector está en  $T\{F = 0\}$  entonces

$$\sum_{i=1}^n \left( \frac{\partial F}{\partial q_i} a_i + \frac{\partial F}{\partial p_i} b_i \right) + \frac{\partial F}{\partial y} c = 0.$$

De estas dos ecuaciones obtenemos que en las coordenadas  $\mathbf{q}, \mathbf{p}$  el espacio ortogonal a las características  $X^\perp$  queda descrito por

$$\sum_{i=1}^n \left[ \left( \frac{\partial F}{\partial q_i} + \frac{\partial F}{\partial y} p_i \right) a_i + \frac{\partial F}{\partial p_i} b_i \right] = 0. \quad (3.7)$$

En estas mismas coordenadas la 2-forma  $\omega$  se escribe como  $\omega = -\sum dp_i \wedge dq_i$  y la matriz asociada en estas coordenadas es la matriz

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

así el producto de dos vectores  $v = \sum \dot{q}_i \frac{\partial}{\partial q_i} + \dot{p}_i \frac{\partial}{\partial p_i}$ ,  $w = \sum a_i \frac{\partial}{\partial q_i} + b_i \frac{\partial}{\partial p_i}$  es igual a

$$\omega(v, w) = \sum_{i=1}^n \dot{q}_i b_i - \dot{p}_i a_i.$$

Comparando esta expresión con la ecuación (3.7) que describe a  $X^\perp$  obtenemos las primeras  $2n$  ecuaciones para  $X$ :

$$\begin{aligned} \dot{q}_i &= \frac{\partial F}{\partial p_i} \\ \dot{p}_i &= - \left( \frac{\partial F}{\partial q_i} + \frac{\partial F}{\partial y} p_i \right), \end{aligned}$$

la última ecuación claramente es  $\dot{y} = \sum p_i \frac{\partial F}{\partial p_i}$ . □

Tomemos ahora un punto  $\mathbf{q}_0 \in \partial W$ , como  $W$  es una variedad con frontera podemos encontrar coordenadas  $q_1, \dots, q_n$  definidas en un subconjunto  $V$  tales que  $\Gamma \cap V$  queda descrito por la ecuación  $q_n = 0$ , supongamos que en las coordenadas que  $\mathbf{q}$  induce en  $J^1(M)$  tenemos que  $\left. \frac{\partial F}{\partial p_n} \right|_{\mathbf{q}_0} \neq 0$ . Bajo estas condiciones podemos construir una subvariedad  $N \subset J^1(V)$  de dimensión  $n-1$  e isotrópica (la 2-forma se anula en esta subvariedad) de la siguiente manera: Por nuestra hipótesis la ecuación

$$F \left( q_1, \dots, q_{n-1}, 0, \left. \frac{\partial u_0}{\partial q_1} \right|_{\mathbf{q}}, \dots, \left. \frac{\partial u_0}{\partial q_{n-1}} \right|_{\mathbf{q}}, \phi, u_0(\mathbf{q}) \right) = 0$$

### 3.2. ECUACIONES DIFERENCIALES PARCIALES DE PRIMER ORDEN 21

admite una solución  $\phi(q_1, \dots, q_{n-1}, 0, \frac{\partial u_0}{\partial q_1} \Big|_{\mathbf{q}}, \dots, \frac{\partial u_0}{\partial q_{n-1}} \Big|_{\mathbf{q}}, u(\mathbf{q}))$  para puntos  $\mathbf{q}$  cercanos a  $\mathbf{q}_0$ . Así, definimos a  $N \subset J^1(M)$  como los puntos de la forma

$$N = \left\{ \left( q_1, \dots, q_{n-1}, 0, \frac{\partial u_0}{\partial q_1}, \dots, \frac{\partial u_0}{\partial q_{n-1}}, \phi, u_0 \right) \right\},$$

donde  $u_0$  y sus derivadas están evaluadas en el punto  $(q_1, \dots, q_{n-1}, 0)$  y la función  $\phi$  está evaluada en este mismo punto, en el valor de  $u_0$  en este punto y las derivadas de  $u_0$  en este punto. Llamaremos a  $N$  la *variedad inicial*. El hecho de que este conjunto sea una subvariedad y que  $\omega$  se anule en ella es idéntico al caso de 1-gráficas de funciones. La proposición 3.3 sugiere que si tomamos puntos en la  $n-1$  variedad  $N$  y los movemos usando a las características obtenemos una solución a la ecuación diferencial determinada por  $F$ . Para formalizar lo anterior llamemos  $g^s = (\mathbf{q}(s), \mathbf{p}(s), y(s))$  al flujo del campo vectorial que genera a las direcciones de las características.

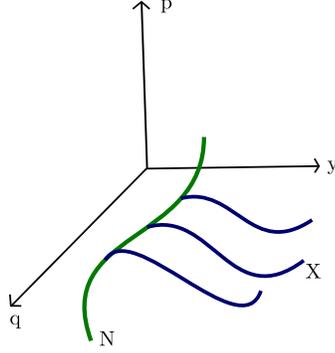


Figura 3.3: Construcción de una solución al problema de Cauchy.

**Lema 3.2.** Si en el punto  $\mathbf{q}_0 \in \Gamma$  tenemos que  $\frac{\partial F}{\partial p_n} \neq 0$ , entonces la asociación  $(q_1, \dots, q_{n-1}, s) \xrightarrow{G} (q_1(s), \dots, q_n(s))$ , donde las funciones  $q_i(s)$  son las proyecciones de las características al espacio  $\mathbf{q}$  y la condición inicial del flujo de las características es tomada en la variedad inicial  $N$ , es un difeomorfismo local alrededor de  $(\mathbf{q}_0, 0)$ .

*Demostración.* Para  $s = 0$  la función  $G$  es la identidad en  $\Gamma$ , así pues basta calcular  $\frac{\partial G}{\partial s}$ . Si  $\Pi$  denota la proyección de  $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$  en el primer factor,

$$\frac{\partial G}{\partial s} \Big|_{s=0} = \frac{d}{ds} \Big|_{s=0} (\Pi(g^s)) = \left( \frac{\partial F}{\partial p_1}, \dots, \frac{\partial F}{\partial p_n} \right),$$

por el lema 3.1. Así pues  $\det \left( DG \Big|_{(\mathbf{q}_0, 0)} \right) = \frac{\partial F}{\partial p_n} \neq 0$ . □

El lema anterior nos permite construir una función  $u$  para puntos cercanos a  $\mathbf{q}_0$ , en efecto, si  $\mathbf{q}$  está lo suficientemente cerca a  $\mathbf{q}_0$  podemos escribir  $\mathbf{q} =$

$G(\tilde{\mathbf{q}}, s)$  para algunos  $\tilde{\mathbf{q}} \in \Gamma$  y  $s \in \mathbb{R}$ , definimos  $u(\mathbf{q}) = y(\tilde{\mathbf{q}}, s)$  donde  $g^s = (\mathbf{q}(s), \mathbf{p}(s), y(s))$  es el flujo de las características de la ecuación. A continuación vamos a probar que la función que acabamos de construir es una solución local al problema de Cauchy, para esto vamos a probar que la  $n$  variedad  $Z$  que consiste en las características con condición inicial en la  $n-1$  variedad inicial  $N$  coincide con la 1-gráfica de la función  $u$ . En el siguiente lema no suponemos que  $d\alpha$  es no degenerada.

**Lema 3.3.** *Sea  $V$  una variedad diferencial con una distribución de hiperplanos  $\xi$  dada por una 1-forma no nula  $\alpha$ . Consideremos un campo vectorial no nulo  $X$  en la distribución ( $\alpha(X) = 0$ ) y sea  $g^s$  su flujo, entonces  $g_*^s(\xi) = \xi$  si y solamente si  $d\alpha(X, v) = 0$  para todo campo vectorial  $v \in \xi$ .*

*Demostración.* Tenemos que  $g_*^s(\xi) = \xi$  si y sólo si  $\mathcal{L}_X\alpha = \lambda\alpha$  para alguna función  $\lambda: V \rightarrow \mathbb{R}$ . Si en un punto  $p$  ocurre que  $\lambda(p) = 0$  entonces en este punto tenemos que  $\mathcal{L}_X\alpha = \alpha$  con la fórmula de Cartan obtenemos  $0 = \mathcal{L}_X\alpha = (i_X \circ d)\alpha + (d \circ i_X)\alpha = d\alpha(X, \cdot)$ . Si  $\lambda(p) \neq 0$  entonces  $\mathcal{L}_X\alpha = (i_X \circ d)\alpha + (d \circ i_X)\alpha$  si y sólo si  $\ker(\mathcal{L}_X\alpha) = \ker(\alpha) = \xi$ , nuevamente con la fórmula de Cartan tenemos que  $\mathcal{L}_X\alpha = (i_X \circ d)\alpha + (d \circ i_X)\alpha$ , pero como  $X \in \ker(\alpha)$  entonces  $\mathcal{L}_X\alpha = (i_X \circ d)\alpha$ , así  $\ker(\mathcal{L}_X\alpha) = \xi$  si sólo si  $d(\alpha, v) = 0$  para todo  $v \in \xi$ .  $\square$

**Proposición 3.4.** *Sean  $W \subset M$  un abierto con frontera suave y  $F: J^1(M) \rightarrow \mathbb{R}$  una ecuación diferencial parcial no singular. Dada una condición inicial  $u_0: \Gamma \subseteq \partial W \rightarrow \mathbb{R}$  consideremos la subvariedad isotrópica  $N \subset J^1(M)$  asociada a esta condición inicial. Si para un punto  $q_0 \in \Gamma \subseteq \partial W$ , en coordenadas locales  $\frac{\partial F}{\partial p_n} \neq 0$  entonces la función  $u(\mathbf{q}) = y(\tilde{\mathbf{q}}, s)$ , donde  $\mathbf{q} = G(\tilde{\mathbf{q}}, s)$  y  $g^s = (\mathbf{q}(s), \mathbf{p}(s), y(s))$  es el flujo característico con condición inicial en el punto de  $N$  correspondiente  $\tilde{\mathbf{q}} \in \Gamma$ , es una solución local al problema de Cauchy.*

*Demostración.* Por construcción la subvariedad inicial  $N \subset \{F = 0\}$  y el campo de direcciones  $X$  de las características de la ecuación  $F$  satisface  $X \in X^\perp = T\{F = 0\} \cap \xi \subset T\{F = 0\}$  así, las características con condición inicial en  $N$  se mantienen en  $\{F = 0\}$ . Sea  $Z$  la  $n$  variedad de las características con condición inicial en la variedad inicial  $N$ , hemos visto que  $Z \subset \{F = 0\}$  y ahora veremos que también es tangente a la distribución  $\xi$ . Para todo punto  $x \in N$  tenemos que  $T_x Z = T_x N \oplus X$ , este espacio está contenido en la distribución  $\xi$ , consideremos la inclusión  $i: \{F = 0\} \rightarrow J^1(M)$  y la 1-forma  $i^*(\alpha)$ , donde  $\xi = \ker(\alpha)$ , tenemos que  $X \in \ker(i^*\alpha)$  y si  $v \in \ker(i^*\alpha)$  entonces  $v \in \xi \cap \{F = 0\} = X^\perp$  y por lo tanto  $d\alpha(X, v) = 0$ . Por el lema anterior y como  $T_x Z \subset \xi$  para todo  $x \in N$  obtenemos  $TZ \subset \xi = \ker(\alpha)$ , de esto se sigue que  $dy - \sum p_i dq_i = 0$  en  $Z$  y por lo tanto  $\frac{\partial u}{\partial q_i} = p_i$ . Esto último quiere decir que  $Z$  es efectivamente la 1-gráfica de la función  $u$ , como  $Z \subset \{F = 0\}$ ,  $u$  es una solución local al problema de Cauchy.  $\square$

**Ejemplo 4.** En este ejemplo vamos a tomar por  $M$  a  $\tilde{M} \times \mathbb{R}_t$ , (el subíndice  $t$  indica que esta coordenada va a funcionar como tiempo y a veces la denotaremos por  $q_0$  y a veces con  $t$ ) en este caso  $J^1(M) \cong T^*\tilde{M} \times \mathbb{R}_t \times \mathbb{R}^2$  (una copia de  $\mathbb{R}$  es

### 3.2. ECUACIONES DIFERENCIALES PARCIALES DE PRIMER ORDEN 23

para la coordenada  $p_0$  de  $T^*M$  correspondiente a  $q_0 = t$  y la otra corresponde a la coordenada que registra los valores de funciones  $u: M \rightarrow \mathbb{R}$  y entonces la proyección canónica de este espacio a  $T^*M \times \mathbb{R}_t$  nos permite jalar funciones  $H: T^*\tilde{M} \times \mathbb{R} \rightarrow \mathbb{R}$  a  $J^1(M)$ . Para dicha función consideremos la siguiente función que está definida en  $J^1(M)$ ,

$$F = p_0 + H(x, t),$$

donde  $x$  es un punto de  $T^*\tilde{M}$ , notemos que  $\frac{\partial F}{\partial p_0} = 1$ , además  $\frac{\partial F}{\partial y} = 0$ , así el vector  $\frac{\partial}{\partial y}$  está en  $T\{F = 0\}$ , pero este vector nunca está en la distribución  $\xi$  y por lo tanto podemos aplicar la proposición anterior. Si para una función  $S: M = \tilde{M} \times \mathbb{R} \rightarrow \mathbb{R}$  denotamos por  $dS_{\mathbf{q}}$  a la diferencial respecto a  $\tilde{M}$  entonces la función anterior representa la ecuación diferencial parcial

$$\frac{\partial S}{\partial t} + H(dS_{\mathbf{q}}, t) = 0,$$

En coordenadas locales  $\mathbf{q}$  de  $\tilde{M}$  esto se escribe como

$$\frac{\partial S}{\partial t} + H(\mathbf{q}, \frac{\partial S}{\partial \mathbf{q}}, t) = 0,$$

esta es la famosa *ecuación de Hamilton-Jacobi*. El campo vectorial que genera las direcciones de las características está dado por las ecuaciones

$$\begin{aligned} \dot{p}_0 &= -\frac{\partial H}{\partial t} & \dot{p}_i &= -\frac{\partial H}{\partial q_i} \\ \dot{q}_0 &= 1 & \dot{q}_i &= \frac{\partial H}{\partial p_i}. \end{aligned}$$

Notemos que las dos ecuaciones de la derecha son idénticas a las ecuaciones diferenciales de un sistema Hamiltoniano sólo que en este caso el Hamiltoniano puede depender del tiempo. Si  $H$  no depende del tiempo  $t$ , puesto que las 1-gráficas de las soluciones están en el subconjunto  $\{F = 0\}$ , entonces  $p_0 = -H$  y así obtenemos que el Hamiltoniano  $H$  es invariante a través del tiempo.

**Ejemplo 5.** Consideremos la ecuación diferencial parcial para una función  $S(\mathbf{q}, t)$  dada por

$$\frac{\partial S}{\partial t} + H\left(\mathbf{q}, \frac{\partial S}{\partial \mathbf{q}}, \mathbf{q}, t, S(\mathbf{q})\right) = 0.$$

Con la notación ( $q_0 = t$ ,  $\mathbf{p} = \frac{\partial S}{\partial \mathbf{q}}$  y  $p_0 = -H$  para la 1-gráfica del cualquier solución) del ejemplo anterior y usando a  $y$  como una coordenada que registra los valores de las funciones  $S$ , las características de esta ecuación están dadas por

$$\begin{aligned} \dot{p}_0 &= -\left(\frac{\partial H}{\partial t} - H \frac{\partial H}{\partial y}\right) & \dot{p}_i &= -\left(\frac{\partial H}{\partial q_i} + p_i \frac{\partial H}{\partial y}\right) \\ \dot{q}_0 &= 1 & \dot{q}_i &= \frac{\partial H}{\partial p_i}. \end{aligned}$$

$$\dot{y} = \sum_{i=1}^n p_i \frac{\partial H}{\partial p_i} - H.$$

Estas son las ecuaciones (3.3) y (3.4). A diferencia de la igualdad (3.5) en este caso tenemos que el cambio en el Hamiltoniano  $H = -p_0$  está dado por

$$\dot{H} = \frac{\partial H}{\partial t} - H \frac{\partial H}{\partial y}.$$

Esto es análogo al caso simplético en el cual el Hamiltoniano se conserva y el ejemplo anterior en el cuál el cambio en el Hamiltoniano está dado por  $\frac{\partial H}{\partial t}$ .

Hemos visto que, bajo ciertas condiciones podemos usar a las características para construir soluciones locales de una ecuación diferencial parcial. Al final del siguiente capítulo veremos que podemos hacer lo opuesto, conociendo suficientes soluciones de la ecuación diferencial parcial podemos encontrar a las características. La idea es la siguiente: cada solución de la ecuación parcial determina una 1-gráfica en  $J^1(M)$  que es de dimensión  $n$ . Si conseguimos  $n$  de ellas cuyas 1-gráficas se intersectan de manera transversal entonces puesto que cada 1-gráfica está formada por características (lema 3.1), su intersección tiene que ser una característica.

**Ejemplo 6.** Consideremos una variedad Riemanniana  $(M, g)$ , el isomorfismo que define  $g$  entre  $TM$  y  $T^*M$  nos permite considerar a la métrica  $g$  como una función en  $T^*M$ , al igual que en el ejemplo anterior, para cada  $c \in \mathbb{R}_{>0}$  podemos considerar la función  $F = \frac{1}{2}(g - c^2) : J^1(M) \rightarrow \mathbb{R}$ . Si en un abierto coordenado tenemos que  $g = \sum_{i=1}^n dx_i^2$  entonces en esta carta  $F = \frac{1}{2}(\sum p_i^2 - c^2)$  y así las características de la ecuación  $\{F = 0\}$  están dadas por:

$$\dot{q}_i = p_i \quad \dot{p}_i = 0 \quad \dot{y} = \sum_{i=1}^n p_i^2.$$

Si para  $\Gamma = \{q_n = 0\}$  consideramos el problema de Cauchy con condición inicial  $u_0 \equiv 0$ , entonces para construir la subvariedad inicial  $N$ , tenemos que  $p_i = 0$  para  $i = 1 \dots, n-1$  y en consecuencia  $p_n = \pm c$ . Así, las características con condición inicial en la subvariedad  $N$  quedan descritas por  $q_i(s) = q_{i0}, p_i(s) = 0$  para  $i = 1 \dots, n$ ,  $q_n(s) = \pm cs$  y  $y(s) = c^2 s$ . Si tomamos unidades en las cuales  $c = 1$  podemos pensar que las soluciones a los problemas de Cauchy de esta ecuación, con  $u_0 \equiv 0$ , nos indican el tiempo que tarda en llegar a un punto  $q$  una señal emitida desde la subvariedad  $\Gamma$ . A la ecuación diferencial parcial

$$\sum_{i=1}^n \left( \frac{\partial u}{\partial q_i} \right)^2 - 1 = 0,$$

se le conoce como *la ecuación Eikonal*.

Si  $\Gamma$  no es globalmente un plano, puede ocurrir que dos características con condiciones iniciales en dos puntos distintos de la subvariedad inicial  $N$  se proyecten

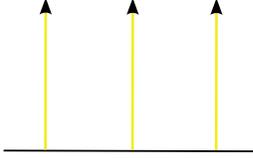


Figura 3.4: Propagación local de señales.

al mismo punto en  $M$ . Cuando lo anterior ocurre el lema 3.2 deja de funcionar. Este fenómeno se puede apreciar muy bien si reflejamos una señal de luz en un vaso o superficie convexa. A una cierta distancia de la superficie podemos ver puntos donde la intensidad de la luz es mucho mayor que en otras partes. Estos puntos se corresponden a características que se proyectan a un mismo punto.

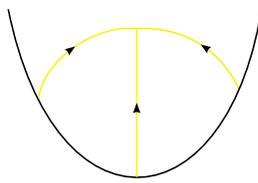


Figura 3.5: Propagación global de señales.

Otra manera de estudiar la propagación de señales en una variedad Riemanniana  $(B, g)$  es la siguiente: Si la señal se propaga con velocidad  $c$  consideremos la forma cuadrática en  $M = B \times \mathbb{R}$  definida por  $g - c^2 dt^2$ . Esta forma define un polinomio homogéneo en cada  $T_p^*M$  y por lo tanto define una hipersuperficie en  $\mathbb{P}T^*M$  que es la variedad de elementos de contacto de  $M$ . Si  $q_1, \dots, q_n$  son coordenadas de  $B$  y  $q_0, q_1, \dots, q_n$  coordenadas de  $M = B \times \mathbb{R}$  entonces en la carta afín dada por  $p_0 = -1$  este polinomio se convierte en  $g - c^2$  y en esta carta afín la 1-forma que define la estructura de contacto de  $\mathbb{P}T^*M$  se vuelve  $\alpha = dq_0 - \sum_{i=1}^n p_i dq_i$  (ver ejemplo 3). Con la hipersuperficie definida por  $g - c^2 = 0$  y la estructura de contacto podemos definir a las características de manera idéntica a como lo hicimos antes y es claro que estas características coinciden con las del ejemplo anterior. Este hecho también va de acuerdo a que las variedades de contacto no tienen invariantes locales.



# Capítulo 4

## Mecánica Clásica

Desde que en 1687 Isaac Newton publicara sus Principia Mathematica las ecuaciones de movimiento han servido para modelar y estudiar, desde un punto de vista matemático, diversos problemas físicos. A pesar de su sencillez, casi nunca es posible obtener soluciones explícitas de estas ecuaciones. A lo largo de la historia se han creado diversos métodos para estudiar casos particulares (pero lo suficientemente generales) de las ecuaciones de movimiento. En este capítulo estudiaremos las ecuaciones de Euler-Lagrange, las de Hamilton y la ecuación de Hamilton-Jacobi. Estas ecuaciones son equivalentes a las ecuaciones de movimiento de los sistemas conservativos.

### 4.1. Las leyes de Newton

Si tenemos un sistema mecánico conformado por una partícula de masa  $m$  en  $\mathbb{R}^3$ , el objetivo de la mecánica clásica es describir la trayectoria que sigue la partícula. Esto quiere decir que para cada tiempo  $t \in \mathbb{R}$  debemos encontrar una función  $x = x(t): \mathbb{R} \rightarrow \mathbb{R}^3$  tal que  $x(t)$  es la posición de la partícula al tiempo  $t$ . Para lograr esta descripción necesitamos usar algún sistema de referencia (coordenadas) de  $\mathbb{R}^3$ . En su famoso libro Principia Mathematica, Newton postula la existencia de una clase especial de sistemas de referencia llamados sistemas de referencia inercial, en ellos, una partícula en la cual no actúa ninguna fuerza se mueve a velocidad constante y las leyes de la naturaleza son las mismas en distintos marcos de referencia inercial. Lo anterior es conocido como la primera ley de Newton, en su segunda ley se postula que, para cada sistema mecánico, en un sistema de referencia inercial existe un campo de vectores  $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  tal que

$$m\ddot{x} = F(x),$$

donde  $\ddot{x}$  es la aceleración de la partícula ( $\ddot{x} = \frac{d^2x}{dt^2}$ ) y  $m$  es la masa, si llamamos a  $p := m\frac{dx}{dt}$  el *momento lineal* podemos escribir lo anterior como,

$$\frac{dp}{dt} = F. \tag{4.1}$$

La igualdad anterior es llamada la *ecuación de Newton* y es una ecuación diferencial de segundo orden en  $x$  y por lo tanto nos proporciona un método para encontrar la trayectoria que sigue el sistema mecánico. Notemos que la segunda ley parece implicar la primera, si  $F(x) = 0$  para toda  $x \in \mathbb{R}^3$  entonces  $\dot{x} = \frac{dx}{dt} = c$  con  $c$  un vector constante, pero no olvidemos que para que la igualdad  $m\ddot{x} = F(x)$  sea cierta debemos estar en un sistema de referencia inercial.

Así pues, para encontrar la trayectoria que sigue la partícula basta calcular la fuerza  $F$  que actúa sobre la partícula y resolver la ecuación de Newton (4.1). Como primer paso, la primera ley de Newton nos dice que si en un sistema de referencia inercial, la partícula no se queda en reposo ni sigue una línea recta a velocidad constante, entonces hay una fuerza que está actuando en la partícula. Además, Newton también postula que la fuerza total que actúa en una partícula es la suma de la fuerzas que actúan en la partícula.

En los siguientes ejemplos vamos a tomar como sistema de referencia inercial a las coordenadas rectangulares de  $\mathbb{R}^n$ . Para muchos fenómenos mecánicos en la Tierra un sistema de coordenadas rectangular fijo es, hasta un muy buen grado de aproximación, un sistema de referencia inercial. Lo mismo ocurre con un sistema rectangular en el sistema solar al estudiar el movimiento de cuerpos celestes.

**Ejemplo 7.** Cuando vamos en una autopista y tomamos una curva muy pronunciada sentimos una fuerza (fuerza centrífuga) que nos jala hacia la parte exterior de la curva. Esto se debe a que, si  $\alpha: \mathbb{R} \rightarrow \mathbb{R}^2$  representa la trayectoria que sigue el carro entonces  $\ddot{\alpha} = AT + BN$ , donde  $T$  es el vector tangente a la autopista y  $N$  es el vector normal. El motor del carro sólo puede hacer fuerza en la dirección tangente (la dirección de las llantas), es por eso que algunas autopistas están ligeramente inclinadas. Esta inclinación ayuda a mantener el carro dentro de la autopista aunque con la velocidad adecuada, los pasajeros siempre pueden experimentar la fuerza centrífuga. Esto se debe a que la componente de la fuerza en la dirección normal tiene norma igual a la rapidez con la que el carro se desplaza.

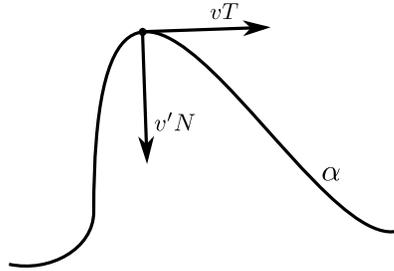


Figura 4.1: Un carro desplazándose en una carretera con velocidad  $v$ .

**Ejemplo 8.** La fuerza de gravedad que actúa sobre una partícula en la Tierra

puede ser considerada constante  $F = g$ . Así, una partícula bajo el campo gravitatorio de la Tierra sigue una trayectoria de acuerdo a  $m\ddot{x} = g$ . De aquí es inmediato que las trayectorias son líneas rectas (velocidad inicial paralela a  $g$ ) o parábolas. Consideremos ahora el caso en que la partícula está restringida a moverse en un péndulo circular, nuevamente la fuerza de gravedad actúa en la partícula pero además, la cuerda que une a la partícula con el centro del péndulo hace una fuerza  $F^{res}$  sobre la partícula para mantenerla en el péndulo.

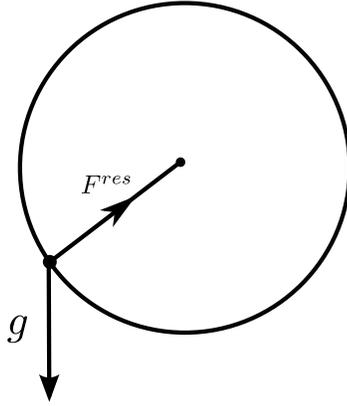


Figura 4.2: Péndulo bajo la fuerza de gravedad.

Si tenemos un sistema mecánico conformado por  $n$  partículas de masas  $m_i$  entonces, en un sistema de referencia inercial, la  $i$ -ésima partícula sigue una trayectoria de acuerdo a

$$\frac{dp_i}{dt} = F_i,$$

con  $p_i = m_i \dot{x}_i$  el momento lineal de la  $i$ -ésima partícula y  $F_i$  la fuerza total que actúa en esta partícula. Podemos escribir

$$F_i = \sum_{j=1, j \neq i}^n F_{ij} + F_i^{ext},$$

donde  $F_{ij}$  es la fuerza debida a la interacción entre la  $i$ -ésima y  $j$ -ésima partícula y  $F_i^{ext}$  es la fuerza exterior que actúa en la  $i$ -ésima partícula. Por ejemplo, si consideramos la fuerza gravitatoria que ejercen los planetas y el Sol entre ellos entonces en nuestro sistema no hay fuerzas exteriores, por otro lado si sólo analizamos la órbita de un planeta alrededor del Sol, la aceleración  $F/m_{sol}$  del Sol debida a la interacción gravitatoria es tan pequeña que la podemos despreciar y entonces, en el planeta sólo actúa una fuerza gravitatoria externa debida al Sol. La tercera ley de Newton asegura que

$$F_{ij} = -F_{ji},$$

y además la dirección de estos vectores coincide con la dirección de la recta que une a estas dos partículas. Podemos pensar que las funciones  $F_{ij}$  forman una matriz antisimétrica y con esto vemos que

$$\sum_{i=1}^n \sum_{j=1, j \neq i}^n F_{ij} = \sum_{i < j} F_{ij} + \sum_{i > j} F_{ij} = 0.$$

Así,

$$\sum_{i=1}^n \frac{dp_i}{dt} = \sum_{i=1}^n F_i^{ext} = F^{ext}.$$

Si  $M := \sum m_i$  es la masa total del sistema, al vector

$$x_{cm} = \sum_{i=1}^n \frac{m_i x_i}{M}$$

se le conoce como *centro de masa del sistema* y a  $p_{cm} = M \dot{x}_{cm}$  como el *momento lineal total*. Notemos que,

$$\frac{dp_{cm}}{dt} = \sum_{i=1}^n \frac{dp_i}{dt} = F^{ext}, \quad (4.2)$$

esto quiere decir que el centro de masa del sistema se mueve como si únicamente la fuerza total exterior actuara en él.

Ahora vamos a usar la tercera ley de Newton para obtener un resultado análogo a la igualdad (4.2) pero para la cantidad

$$L := \sum_{i=1}^n m_i x_i \times \dot{x}_i$$

a esta cantidad se le conoce como el *momento angular total*. Derivando respecto al tiempo

$$\frac{dL}{dt} = \sum_{i=1}^n m_i (\dot{x}_i \times \dot{x}_i + x_i \times \ddot{x}_i) = \sum_{i=1}^n x_i \times F_i,$$

sustituyendo,

$$\sum_{i=1}^n x_i \times F_i = \sum_{i < j} x_i \times F_{ij} + \sum_{i > j} x_i \times F_{ij} + \sum_{i=1}^n x_i \times F_i^{ext},$$

el segundo sumando es igual a

$$\sum_{i > j} x_i \times F_{ij} = - \sum_{i > j} x_i \times F_{ji} = - \sum_{i < j} x_j \times F_{ij}$$

así,

$$\frac{dL}{dt} = \sum_{i < j} (x_i - x_j) \times F_{ij} + \sum_{i=1}^n x_i \times F_i^{ext}.$$

La tercera ley de Newton dice que la fuerza  $F_{ij}$  tiene la dirección del vector  $x_i - x_j$  y por lo tanto:

$$\frac{dL}{dt} = \sum_{i=1}^n x_i \times F_i^{ext}. \quad (4.3)$$

## 4.2. Leyes de conservación

**Definición 13.** Consideremos un sistema mecánico conformado por  $n$  partículas. Diremos que este sistema es un *sistema cerrado* si es libre de fuerzas exteriores, esto quiere decir que para cada  $i = 1, \dots, n$  tenemos que  $F_i^{ext} = 0$ .

De esta definición, las igualdades (4.2) y (4.3) obtenemos de manera inmediata el siguiente teorema:

**Teorema 4.1.** *En cualquier sistema cerrado el momento lineal total y el momento angular total se conservan.*

A continuación vamos a imponer una condición más al campo de fuerzas. Si el sistema está conformado por una sola partícula diremos que el sistema es *conservativo* si existe una función  $U: \mathbb{R}^3 \rightarrow \mathbb{R}$  tal que  $F = -\nabla U$ . En este caso definimos la *energía* del sistema como la cantidad

$$H := \frac{1}{2} m \dot{x} + U(x).$$

La importancia de las definiciones anteriores reside en que la energía de este sistema se conserva, en efecto,

$$\frac{dH}{dt} = m \langle \dot{x}, \ddot{x} \rangle + \langle \nabla U(x), \dot{x} \rangle = \langle \dot{x}, F \rangle - \langle F, \dot{x} \rangle = 0.$$

Para un sistema de  $n$  partículas podemos pensar que el espacio de todas las posibles configuraciones espaciales es  $\mathbb{R}^{3n}$ .

**Definición 14.** Consideremos un sistema mecánico conformado por  $n$  partículas de masas  $m_i$ . Diremos que el sistema es conservativo si existe una función  $U: \mathbb{R}^{3n} \rightarrow \mathbb{R}$  tal que  $F_i = \nabla_{x_i} U$ . A  $U$  se le conoce como la *energía potencial*, además definimos la *energía cinética*  $T$ , y la *energía*  $H$  del sistema como

$$T = \frac{1}{2} \sum_{i=1}^n m_i \langle \dot{x}_i, \dot{x}_i \rangle$$

$$H = T(\dot{x}_1, \dots, \dot{x}_n) + U(x_1, \dots, x_n).$$

De manera análoga a lo que si hizo anteriormente obtenemos:

**Teorema 4.2.** *La energía de un sistema mecánico conservativo se conserva.*

**Ejemplo 9.** Una partícula en  $\mathbb{R}^3$  bajo un campo de fuerzas centrales es un sistema mecánico en el cual, la fuerza sólo depende de la distancia  $r$  de la partícula a un punto fijo que tomaremos por el origen. Así, las curvas de nivel de  $U$  son esferas concéntricas y por lo tanto  $F = -\nabla U$  tiene la misma dirección que el vector de posición. Con la anterior vemos que  $\frac{d}{dt}(x \times \dot{x}) = 0$  y por lo tanto el momento angular de este sistema se conserva. Como el vector  $x \times \dot{x}$  es constante la partícula se desplaza en el plano normal a este vector.

### 4.3. El principio de D'Alembert

Si tenemos una partícula de masa  $m$  en  $\mathbb{R}^3$  la cual está restringida a moverse en una subvariedad  $M \subset \mathbb{R}^3$ , podemos descomponer al vector de fuerza  $F$  como una suma  $F = F^{ext} + F^{res}$ , donde  $F^{ext}$  es la fuerza exterior que actúa sobre la partícula y  $F^{res}$  es una fuerza que mantiene a la partícula en la subvariedad  $M$ . Si la fuerza  $F^{res}$  está en la dirección normal a la subvariedad entonces, para todo vector tangente  $\tau$  a la subvariedad  $M$ .

$$\langle F - F^{ext}, \tau \rangle = 0.$$

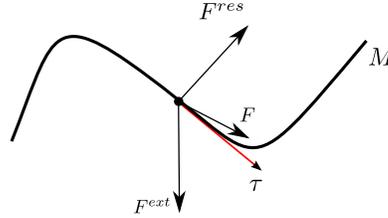


Figura 4.3: Las fuerzas de restricción.

Usando la ecuación de Newton  $\frac{dp}{dt} = F$  y como  $F^{res}$  es normal a la subvariedad tenemos que,

$$\left\langle \frac{dp}{dt} - F, \tau \right\rangle = 0. \quad (4.4)$$

a la igualdad anterior se le conoce como el principio *D'Alembert*, notemos que cuando la subvariedad  $M$  es  $\mathbb{R}^3$  recuperamos la ecuación de Newton. A continuación vamos a usar este principio para derivar la ecuación de Euler-Lagrange que satisface este sistema mecánico cuando la fuerza  $F^{ext}$  es conservativa. En este caso,  $U(x) = -\int_{x_0}^x F^{ext} dl - U(x_0)$  donde la integral de línea se calcula sobre cualquier camino que una a  $x_0$  con  $x$ . Si nos restringimos a caminos que estén contenidos en la subvariedad  $M$ , el tangente a la subvariedad es ortogonal a  $F^{res}$  y entonces  $U(x) = -\int_{x_0}^x (F^{ext} + F^{res}) dl - U(x_0)$  y por lo tanto la fuerza total  $F = F^{ext} + F^{res}$  admite un potencial  $U$ .

Sean  $q_1, \dots, q_m$  coordenadas locales de  $M$ ,  $x_1, x_2, x_3$  coordenadas rectangulares de algún marco de referencia inercial de  $\mathbb{R}^3$  y  $q(t)$  una solución a la ecuación  $m\ddot{x} = F$ . Consideremos una variación pequeña  $q(t, \epsilon)$ , ( $q(t, 0) = q(t)$ ), de la trayectoria  $q(t)$  que sigue la partícula. El vector  $\tau := \left. \frac{\partial x}{\partial q} \frac{\partial q(t)}{\partial \epsilon} \right|_{\epsilon=0}$  es tangente a  $M$  en el punto  $q(t)$  y las coordenadas rectangulares de este vector son

$$\tau_i = \sum_{j=1}^m \frac{\partial x_i}{\partial q_j} \frac{\partial q_j}{\partial \epsilon} \Big|_{\epsilon=0}.$$

En lo siguiente todas las derivadas respecto a  $\epsilon$  estarán evaluadas en  $\epsilon = 0$  y por lo tanto vamos a omitir el símbolo de evaluación. Tenemos que:

$$\begin{aligned} \left\langle \frac{dp}{dt}, \tau \right\rangle &= \sum_{i=1}^3 \frac{dp_i}{dt} \sum_{j=1}^m \frac{\partial x_i}{\partial q_j} \frac{\partial q_j}{\partial \epsilon} \\ &= \sum_{j=1}^m \sum_{i=1}^3 \left( m \frac{d^2 x_i}{dt^2} \frac{\partial x_i}{q_j} \right) \frac{\partial q_j}{\partial \epsilon}. \end{aligned}$$

Si usamos la regla para derivar un producto obtenemos,

$$\left\langle \frac{dp}{dt}, \tau \right\rangle = \sum_{j=1}^m \left( \sum_{i=1}^3 m \frac{d}{dt} \left( \frac{dx_i}{dt} \frac{\partial x_i}{\partial q_j} \right) - m \frac{dx_i}{dt} \frac{d}{dt} \left( \frac{\partial x_i}{\partial q_j} \right) \right) \frac{\partial q_j}{\partial \epsilon}. \quad (4.5)$$

Por otro lado, en las coordenadas cartesianas de la subvariedad  $x(t) = x(q(t))$  la  $i$ -ésima componente de la velocidad es

$$\frac{dx_i}{dt} = \sum_{j=1}^m \frac{\partial x_i}{\partial q_j} \frac{dq_j}{dt} = \sum_{j=1}^m \frac{\partial x_i}{\partial q_j} \dot{q}_j,$$

al derivar respecto a  $\dot{q}_j$  obtenemos,

$$\frac{\partial}{\partial \dot{q}_j} \left( \frac{dx_i}{dt} \right) = \frac{\partial x_i}{\partial q_j}.$$

Si sustituimos esto en la igualdad (4.5) y cambiamos el orden de las derivadas con respecto a  $q_j$  y  $t$  tenemos que:

$$\left\langle \frac{dp}{dt}, \tau \right\rangle = \sum_{j=1}^m \left( \sum_{i=1}^3 m \frac{d}{dt} \left( \frac{dx_i}{dt} \frac{\partial}{\partial \dot{q}_j} \left( \frac{dx_i}{dt} \right) \right) - m \frac{dx_i}{dt} \frac{\partial}{\partial q_j} \left( \frac{dx_i}{dt} \right) \right) \frac{\partial q_j}{\partial \epsilon}.$$

De la energía cinética  $T = \frac{m}{2} \sum_{i=1}^3 \left( \frac{dx_i}{dt} \right)^2$  obtenemos

$$\begin{aligned} \frac{\partial T}{\partial \dot{q}_j} &= \sum_{i=1}^3 m \frac{dx_i}{dt} \frac{\partial}{\partial \dot{q}_j} \left( \frac{dx_i}{dt} \right) \\ \frac{\partial T}{\partial q_j} &= \sum_{i=1}^3 m \frac{dx_i}{dt} \frac{\partial}{\partial q_j} \left( \frac{dx_i}{dt} \right), \end{aligned}$$

y podemos escribir la igualdad anterior como:

$$\left\langle \frac{dp}{dt}, \tau \right\rangle = \sum_{j=1}^m \left( \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right) \frac{\partial q_j}{\partial \epsilon}. \quad (4.6)$$

Como la fuerza  $F$  es conservativa ( $F = -\frac{\partial U}{\partial x}$ ) el término  $\langle F, \tau \rangle$  es simplemente la derivada de  $-U$  en la dirección de  $\tau$  y esto es igual a  $\sum -\frac{\partial U}{\partial q_j} \frac{\partial q_j}{\partial \epsilon}$ . Si definimos  $L = T - U$  entonces usando la igualdad (4.6) obtenemos:

$$0 = \left\langle \frac{dp}{dt} - F, \tau \right\rangle = \sum_{j=1}^m \left( \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} \right) \frac{\partial q_j}{\partial \epsilon}. \quad (4.7)$$

Como la igualdad anterior debe satisfacerse para cualquier variación entonces para cada  $j = 1, \dots, m$  debe ocurrir que:

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = 0. \quad (4.8)$$

Las ecuaciones anteriores son de gran importancia pues, como veremos en la siguiente sección, se pueden derivar de un principio variacional.

**Ejemplo 10.** Consideremos un péndulo de longitud 1 el cual está restringido a moverse en el plano  $xy$  bajo la acción de la fuerza de gravedad  $(0, -g)$ . Si ponemos el origen de este plano en el pivote del péndulo podemos describir la posición del extremo libre del péndulo mediante el ángulo  $\theta$  que forma con la parte negativa del eje  $y$ . La fuerza de gravedad admite por potencial a la función  $U(x, y) = mgy$ , en coordenadas angulares  $x = \text{sen}(\theta)$ ,  $y = -\text{cos}(\theta)$  y  $\dot{x}^2 + \dot{y}^2 = \dot{\theta}^2$  y por lo tanto el Lagrangiano para este sistema es

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - U = \frac{m}{2}\dot{\theta}^2 + mg \cos(\theta)$$

y la ecuación de Euler-Lagrange es

$$\ddot{\theta} = -g \text{sen}(\theta).$$

En relación con el ejemplo 7 de un carro moviéndose en una autopista, la fricción y calentamiento de las llantas hacen que no podamos encontrar un potencial que describa este sistema. Además los pasajeros de la parte trasera del carro estarán de acuerdo a que no en todas las partes del carro se siente la misma fuerza y por lo tanto no parece muy útil modelar a todo un carro como una sola partícula.

Si ahora suponemos que la fuerza que actúa sobre la partícula se descompone como  $F = F^{ext} + F^{res} = F^{pot} + F^{np} + F^{res}$ , donde  $F^{pot}$  es una fuerza que obtenemos mediante un potencial  $U$  y  $F^{np}$  es una fuerza que no se puede obtener mediante un potencial entonces,

$$\langle F^{np}, \tau \rangle = \sum_{i=1}^4 F_i^{np} \left( \sum_{j=1}^m \frac{\partial x_i}{\partial q_j} \frac{q_j}{\epsilon} \right) = \sum_{j=1}^m \left( \sum_{i=1}^3 F_i^{np} \frac{\partial x_i}{\partial q_j} \right) \frac{\partial q_j}{\partial \epsilon}.$$

Si definimos,

$$D_j := \sum_{i=1}^3 F_i^{np} \frac{\partial x_i}{\partial q_j}, \quad (4.9)$$

entonces podemos escribir la igualdad (4.7) como,

$$0 = \left\langle \frac{dp}{dt} - F, \tau \right\rangle = \sum_{j=1}^m \left( \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} - D_j \right) \frac{\partial q_j}{\partial \epsilon}.$$

Así, obtenemos que

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) = \frac{\partial L}{\partial q_j} + D_j. \quad (4.10)$$

En muchos casos las fuerzas no potenciales son fuerzas de fricción debidas al contacto entre la partícula y la superficie donde se mueve. Una manera de modelar estas fuerzas de fricción es suponer que son proporcionales a la velocidad de la partícula, como estas fuerzas frenan a la partícula la constante de proporcionalidad es negativa.

**Ejemplo 11.** Volviendo al ejemplo 10 si suponemos que  $F^{np} = -m\lambda(\dot{x}, \dot{y})$  entonces, como  $x = \text{sen}(\theta)$  y  $y = \text{cos}(\theta)$ , de la igualdad (4.9) obtenemos:

$$D_\theta = -m\lambda(\dot{x}(-\text{sen}(\theta)) + \dot{y}\text{cos}(\theta)) = -m\lambda(\text{sen}^2(\theta)\dot{\theta} + \text{cos}^2(\theta)\dot{\theta}) = -m\lambda\dot{\theta}.$$

Por lo tanto las ecuación de movimiento es:

$$m\ddot{\theta} = -mg\text{sen}(\theta) - m\lambda\dot{\theta}.$$

## 4.4. Las ecuaciones de Euler-Lagrange

El principio de mínima acción asegura que, dado un sistema mecánico conservativo, de todas las posibles trayectorias que unen a los puntos  $q_1$  y  $q_2$  la trayectoria física que sigue el sistema es tal que la función,

$$S(q(t), q_1, q_2) = \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t) dt,$$

toma un valor crítico. Tomando pequeñas variaciones de la solución física que empieza en el punto  $q_1$  y termina en  $q_2$  podemos llegar a las ecuaciones de Euler-Lagrange (4.8).

Las ecuaciones de Euler-Lagrange sólo dependen de los valores de  $L$  en una vecindad arbitrariamente pequeña de cada punto, en coordenadas locales  $q_1, \dots, q_n$  una variación de la solución física  $\alpha$  puede escribirse como  $\alpha + s\beta$  donde  $\beta(t_1) = \beta(t_2) = 0$ , así,

$$\left. \frac{dS(\alpha + s\beta)}{ds} \right|_{s=0} = \int_{t_1}^{t_2} \sum_{i=1}^n \left( \frac{\partial L(\alpha(t), \dot{\alpha}(t), t)}{\partial q_i} \beta_i(t) + \frac{\partial L(\alpha(t), \dot{\alpha}(t), t)}{\partial \dot{q}_i} \dot{\beta}_i(t) \right) dt,$$

tenemos que,

$$\int_{t_1}^{t_2} \frac{\partial L(\alpha(t), \dot{\alpha}(t), t)}{\partial \dot{q}_i} \dot{\beta}_i(t) dt = \frac{\partial L}{\partial \dot{q}_i} \beta_i(t) \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \beta_i(t) dt,$$

y por lo tanto,

$$\frac{dS(\alpha + s\beta)}{ds} \Big|_{s=0} = \frac{\partial L}{\partial \dot{q}_i} \beta_i(t) \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \sum_{i=1}^n \left( \frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \right) \beta_i(t) dt. \quad (4.11)$$

Como  $\alpha$  es un punto crítico de  $S$ , para cualquier variación  $\beta$  con  $\beta(t_1) = \beta(t_2) = 0$  tenemos que,

$$\int_{t_1}^{t_2} \sum_{i=1}^n \left( \frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \right) \beta_i(t) dt = 0,$$

podemos tomar variaciones tales que  $\beta_i = 0$  salvo para un índice, si el integrando que queda no fuera cero idénticamente podríamos tomar la variación de tal forma que la integral no se anule y por lo tanto, para cada  $i = 1, \dots, n$ ,

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) = 0.$$

Si nuestro sistema se desarrolla en  $\mathbb{R}^m$  y tomamos algún sistema rectangular inercial  $x_1, \dots, x_m$  tenemos que,

$$L = T - U = \frac{m}{2} \dot{x}^2 - U(x),$$

donde  $\dot{x}^2 = \dot{x}_1^2 + \dots + \dot{x}_m^2$  y entonces la ecuación de Euler-Lagrange para este sistema es,

$$\frac{d}{dt} (m\dot{x}) = -\frac{\partial U}{\partial x} = F(x).$$

Si el sistema está restringido a moverse en una subvariedad  $M \subset \mathbb{R}^n$  el principio de mínima acción asegura que, de entre todas las trayectorias que unen a los puntos  $q_1$  y  $q_2$  y que además están contenidas en  $M$ , la trayectoria física es un punto crítico de,

$$S(x(t)) = \int_{t_1}^{t_2} \frac{m}{2} \dot{x}(t)^2 - U(x(t)) dt.$$

Supongamos que en cada punto de la solución física  $\alpha$  tenemos un vector tangente  $\tau$  a la subvariedad  $M$  tal que  $\tau(q_1) = \tau(q_2) = 0$ , en una carta local de  $M$  podemos considerar la variación  $\alpha + s\tau$  (aquí hacemos un abuso de notación identificamos  $\alpha$  y  $\tau$  con sus imágenes bajo la carta coordenada) el principio de mínima acción asegura que,

$$0 = \frac{dS(\alpha + s\tau)}{ds} \Big|_{s=0} = \int_{t_1}^{t_2} \sum_{i=1}^n \left( m\dot{\alpha}(t)\dot{\tau}(t) - \frac{\partial U(\alpha(t))}{\partial x} \tau(t) \right) dt$$

e integrando por partes el primer término,

$$0 = - \int_{t_1}^{t_2} \sum_{i=1}^n \left( m\ddot{\alpha}_i + \frac{\partial U(\alpha)}{\partial x_i} \right) \tau_i dt = - \int_{t_1}^{t_2} \left\langle \frac{dp}{dt} - F, \tau \right\rangle dt.$$

Como la igualdad anterior es válida para cualquier vector tangente a la subvariedad  $M$  a lo largo de la solución  $\alpha$  que cumpla  $\tau(q_1) = \tau(q_2) = 0$ , concluimos que,

$$\left\langle \frac{dp}{dt} - F, \tau \right\rangle = 0. \quad (4.12)$$

La diferencia de esta igualdad con el principio de D'Alembert es que en la igualdad (4.4) no hay ninguna restricción sobre el campo vectorial  $\tau$ . Así pues, cuando un sistema mecánico está restringido a moverse en una subvariedad de  $\mathbb{R}^n$ , las ecuaciones de Euler-Lagrange resultan ser una manera alternativa de escribir las ecuaciones de movimiento.

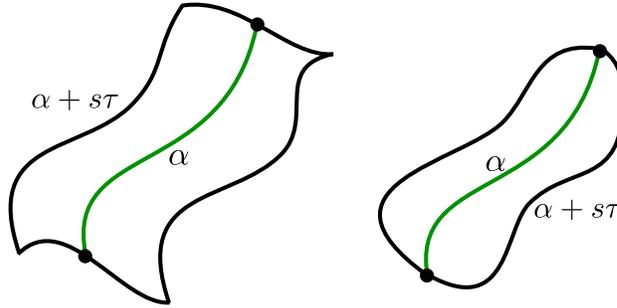


Figura 4.4: Variaciones permitidas en el principio de D'Alembert (izquierda) y en el principio de mínima acción (derecha).

**Ejemplo 12.** El Lagrangiano es una función que está definida en el espacio tangente  $TM$ , en donde  $M$  es la variedad de todas las posibles configuraciones en el espacio de nuestro sistema. La variedad  $TM$  está formada por todos los vectores tangentes a puntos de la variedad  $M$ . Ejemplos de esto son los siguientes. Del espacio  $\mathbb{R}^n$  su espacio tangente es simplemente  $\mathbb{R}^n \times \mathbb{R}^n$ . Para el círculo  $\mathbb{S}^1$ , el hecho de que podamos recorrer el círculo sin parar implica que su espacio tangente es  $T\mathbb{S}^1 \cong \mathbb{S}^1 \times \mathbb{R}$  que es un cilindro. Si el espacio de posiciones es  $T^n$ , un producto de  $n$  círculos, entonces  $TT^n \cong T^n \times \mathbb{R}^n$ . Este hecho al igual que el anterior son consecuencia de que  $T^n$  es un grupo de Lie para todo  $n$ . El grupo de Lie  $SO(3)$  es el espacio de todas las posibles configuraciones en el espacio de un cuerpo sólido que gira en  $\mathbb{R}^3$  alrededor de un punto fijo. Nuevamente  $TSO(3) \cong SO(3) \times \mathbb{R}^3$ .

**Ejemplo 13.** En un campo de fuerzas centrales (ejemplo 14) la conservación del momento angular nos asegura que la partícula está restringida a moverse

en un plano, si tomamos coordenadas en las cuales este plano está descrito por el conjunto  $\{z = 0\}$  y luego usamos coordenadas polares  $(r, \theta)$  en este plano, entonces  $x = r \cos(\theta)$ ,  $y = r \sin(\theta)$  y en consecuencia la energía cinética es  $T = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2)$  y por lo tanto el Lagrangiano de este sistema es

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) - U(r).$$

Las ecuaciones de Euler-Lagrange son:

$$\begin{aligned} m\ddot{r} &= -\frac{\partial U}{\partial r} \\ mr^2\ddot{\theta} &= 0. \end{aligned}$$

En las ecuaciones de Euler-Lagrange (4.8) notemos que si el Lagrangiano  $L$  no depende de la variable  $q_j$  entonces la cantidad  $\frac{\partial L}{\partial \dot{q}_j}$  se conserva. A las funciones  $p_j = \frac{\partial L}{\partial \dot{q}_j}$  se les conoce como momentos generalizados, esto se debe a que en coordenadas rectangulares  $\frac{\partial L}{\partial \dot{x}_j} p_j = m\dot{x}_j$ , que es el momento usual. En la siguiente sección vamos a escribir las ecuaciones de movimiento usando a los momentos generalizados  $p_j$  en lugar de las velocidades  $\dot{q}_j$ . Como preliminar, para cada  $q \in M$  las funciones  $\left. \frac{\partial L}{\partial \dot{q}_i} \right|_q$  definen una función  $T_q M \rightarrow (T_q M)^*$  dada por

$$\sum_{i=1}^n \dot{q}_i \left. \frac{\partial}{\partial \dot{q}_i} \right|_q \mapsto \sum_{i=1}^n \frac{\partial L(q, \dot{q})}{\partial \dot{q}_i} dq_i = \sum_{i=1}^n p_i dq_i.$$

Si la matriz  $\left. \frac{\partial}{\partial \dot{q}} \left( \frac{\partial L}{\partial \dot{q}} \right) \right|_q$  es invertible entonces de manera local podemos escribir a las velocidades  $\dot{q}_i$  en términos de los momentos generalizados.

## 4.5. Las ecuaciones de Hamilton

Para una partícula en  $\mathbb{R}^3$  en la cual actúan ciertas fuerzas debidas a la energía potencial  $U$ , hemos definido la energía de la partícula como la cantidad  $H = T(\dot{x}) + U(x)$ , donde  $T = \frac{m}{2}\dot{x}$  es la energía cinética, el Lagrangiano de este sistema es  $L = T - U$  y por lo tanto  $H = 2T - L$ . Usando la definición de  $T$  vemos que:

$$2T = \frac{\partial L}{\partial \dot{x}} \dot{x} = p\dot{x}. \quad (4.13)$$

La importancia de la energía  $H$  es que es una cantidad conservada. Para un sistema mecánico en general no tenemos una descomposición del Lagrangiano como  $L = T - U$ , sin embargo, la expresión (4.13) sugiere que definamos una cantidad,

$$H := \sum_{i=1}^n p_i \dot{q}_i - L. \quad (4.14)$$

a la función  $H$  le llamaremos el *Hamiltoniano*. Al final de la sección anterior vimos que, si para cada  $q \in M$ , la matriz  $\frac{\partial^2 L}{\partial \dot{q}^2}$  es invertible entonces las funciones

$p_i = \frac{\partial L}{\partial \dot{q}_i}$  definen un difeomorfismo de un abierto  $U \subseteq TM$  en un abierto de  $T^*M$ . Usando la inversa podemos escribir  $\dot{q}_i = \dot{q}_i(\mathbf{p}, \mathbf{q})$  y por lo tanto podemos escribir  $H = H(\mathbf{p}, \mathbf{q})$ . Tenemos que

$$\begin{aligned} dH &= d\left(\sum_{i=1}^n p_i \dot{q}_i - L\right) \\ &= \sum_{i=1}^n \left( \dot{q}_i dp_i + p_i d\dot{q}_i - \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial q_i} dq_i \right) - \frac{\partial L}{\partial t} dt, \end{aligned}$$

como  $p_i = \frac{\partial L}{\partial \dot{q}_i}$  entonces,

$$dH = \sum_{i=1}^n \left( \dot{q}_i dp_i - \frac{\partial L}{\partial q_i} dq_i \right) - \frac{\partial L}{\partial t} dt,$$

usando la ecuación de Euler-Lagrange  $\dot{p}_i = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} \right)$  concluimos que,

$$dH = \sum_{i=1}^n (\dot{q}_i dp_i - \dot{p}_i dq_i) - \frac{\partial L}{\partial t} dt.$$

Por otro lado,

$$dH = \sum_{i=1}^n \left( \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial q_i} dq_i \right) + \frac{\partial H}{\partial t} dt.$$

Comparando esta igualdad con la anterior obtenemos:

$$\begin{aligned} \dot{p}_i &= -\frac{\partial H}{\partial q_i} \\ \dot{q}_i &= \frac{\partial H}{\partial p_i}. \end{aligned} \tag{4.15}$$

Al sistema de ecuaciones diferenciales anterior se les conoce como *ecuaciones de Hamilton*. Si  $H$  no depende del tiempo y  $H$  define una función en  $T^*M$  (la definición de  $H$  es local), las ecuaciones de Hamilton coinciden con el sistema de ecuaciones (2.2) asociadas al sistema Hamiltoniano  $(T^*M, \omega, H)$  y por lo tanto el Hamiltoniano  $H$  es constante a lo largo del flujo asociado a las ecuaciones (4.15). Más aún, en el capítulo 2 vimos que si  $f, g$  eran constantes a lo largo del flujo del sistema Hamiltoniano entonces  $\{f, g\}$  también es constante a lo largo de este flujo. Vamos a probar este mismo resultado pero para el caso en el que  $f, g$  si dependen del tiempo. Primero que nada definimos un corchete para  $f, g: T^*M \times \mathbb{R} \rightarrow \mathbb{R}$ ,

$$\{f, g\} := \omega(d_x f, d_x g),$$

donde  $\omega$  es la 2-forma canónica de  $T^*M$  y  $d_x$  denota la diferencial respecto a  $T^*M$  (el espacio cotangente a  $T^*M \times \mathbb{R}$  se descompone de manera canónica como

$T^*(T^*M) \oplus \mathbb{R}^2$  y  $d_x$  es la proyección a este factor). En coordenadas simplécticas este corchete coincide con la expresión (2.3) y así tenemos que,

$$\frac{df}{dt} = \{H, f\} + \frac{\partial f}{\partial t},$$

por lo tanto  $f$  es una constante del movimiento si y solamente si

$$\{H, f\} = -\frac{\partial f}{\partial t}.$$

Supongamos que  $f, g: T^*M \times \mathbb{R} \rightarrow \mathbb{R}$  son constantes de movimiento, queremos calcular,

$$\frac{d\{f, g\}}{dt} = \{H, \{f, g\}\} + \frac{\partial\{f, g\}}{\partial t},$$

usando la identidad de Jacobi y que  $f$  y  $g$  son constantes del movimiento, el primer sumando es igual a

$$-(\{f, \{g, H\}\} + \{g, \{H, f\}\}) = -(\{f, \frac{\partial g}{\partial t}\} - \{g, \frac{\partial f}{\partial t}\}).$$

Usando la igualdad (2.3) el segundo sumando es igual a,

$$\sum_{i=1}^n \frac{\partial^2 f}{\partial t \partial p_i} \frac{\partial g}{\partial q_i} + \frac{\partial f}{\partial p_i} \frac{\partial^2 g}{\partial t \partial q_i} - (\frac{\partial^2 f}{\partial t \partial q_i} \frac{\partial g}{\partial p_i} + \frac{\partial f}{\partial q_i} \frac{\partial^2 g}{\partial t \partial p_i}) = \{ \frac{\partial f}{\partial t}, g \} + \{ f, \frac{\partial g}{\partial t} \}.$$

Sumando esta última igualdad con la anterior obtenemos que si  $f, g$  son constantes de movimiento entonces  $\{f, g\}$  también es una constante del movimiento.

Consideremos el sistema Hamiltoniano  $(\mathbb{R}^{2n}, \omega, H)$  con  $\omega = \sum dp \wedge dq$  y supongamos que  $\frac{\partial H}{\partial q_1} = 0$ , esto quiere decir que el Hamiltoniano no depende de la coordenada  $q_1$ . Usando las ecuaciones de Hamilton vemos que  $\dot{p}_1 = 0$  y por lo tanto  $p_1$  es una constante del movimiento. Para cada  $c \in \mathbb{R}$  consideremos la función  $\tilde{H}(p_2, \dots, p_n, q_2, \dots, q_n) = H(p_1 = c, p_2, \dots, p_n, q_2, \dots, q_n)$ . Si hacemos  $\tilde{\omega} = \omega - dp_1 \wedge dq_1$ , y logramos encontrar las soluciones  $p_2(t), \dots, p_n(t), q_2(t), \dots, q_n(t)$  al sistema Hamiltoniano  $(\mathbb{R}^{2n-2}, \tilde{\omega}, \tilde{H})$  entonces obtenemos que

$$\dot{q}_1 = \frac{\partial H}{\partial p_1}(c, p_2(t), \dots, p_n(t), q_2(t), \dots, q_n(t)) = f(c, t).$$

Así pues, si el Hamiltoniano no depende de una coordenada el número de grados de libertad del sistema Hamiltoniano desciende en dos. En particular, si el Hamiltoniano sólo depende de los momentos  $H = H(\mathbf{p})$  entonces  $\dot{p}_i = 0$  y por ende,

$$\dot{q}_i = \frac{\partial H}{\partial p_i}(c_1, \dots, c_n) = \alpha_i,$$

por lo tanto  $q_i(t) = \alpha_i t + A_i$ . A los sistemas Hamiltonianos que admiten un Hamiltoniano que sólo depende de los momentos se les conoce como *sistemas totalmente integrables*. En la siguiente sección estudiaremos la ecuación

de Hamilton-Jacobi, esta ecuación, cuando admite una solución que depende de ciertos parámetros, nos ayuda a encontrar un cambio de coordenadas en las cuales el nuevo Hamiltoniano sea una constante y por lo tanto los momentos  $\mathbf{p}$  y las posiciones  $\mathbf{q}$  son cantidades conservadas.

**Ejemplo 14.** En un campo de fuerzas centrales tenemos que el Lagrangiano es  $L = \frac{m}{2}(\dot{r} + r^2\dot{\theta}^2) - U(r)$ . Para este sistema tenemos que,

$$\begin{aligned} p_r &= \frac{\partial L}{\partial \dot{r}} = m\dot{r} & p_\theta &= \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} \\ \dot{r} &= \frac{p_r}{m} & \dot{\theta} &= \frac{p_\theta}{mr^2}. \end{aligned}$$

Por lo tanto el Hamiltoniano  $H$  es igual a

$$\begin{aligned} H &= \frac{p_r^2}{m} + \frac{p_\theta^2}{mr^2} - L = \frac{p_r^2}{m} + \frac{p_\theta^2}{mr^2} - \left( \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} - U(r) \right) \\ &= \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + U(r). \end{aligned}$$

En este caso tenemos que  $\frac{\partial H}{\partial \theta} = 0$  y así  $\dot{p}_\theta = 0$  y con esto basta resolver el sistema

$$\dot{p}_r = \frac{p_\theta^2}{mr^3} - \frac{\partial U}{\partial r} \qquad \dot{r} = \frac{p_r}{m}.$$

Si logramos resolver este sistema obtenemos al radio  $r$  en función del tiempo  $t$ . Por lo tanto sólo requerimos una integración respecto al tiempo para resolver la ecuación

$$\dot{\theta} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{mr^2}. \quad (4.16)$$

Veamos que la conservación del momento angular implica la segunda ley de Kepler. En coordenadas polares, el área delimitada por dos semirectas que forman ángulos  $\theta_0, \theta_1$  y una curva  $r = r(\theta)$  está dada por la integral,

$$A(\theta_0, \theta_1, r(\theta)) = \int_{\theta_0}^{\theta_1} \frac{r^2(\theta)}{2} d\theta = \frac{1}{2} \int_{t_1}^{t_2} r^2(\theta(\tau)) \frac{d\theta}{dt} d\tau,$$

donde  $\theta(t_i) = \theta_i$ . Sustituyendo la ecuación (4.16) obtenemos,

$$A(\theta_0, \theta_1, r(\theta)) = \int_{t_1}^{t_2} \frac{p_\theta}{2m} d\tau = \frac{p_\theta}{2m}(t_1 - t_0).$$

Por lo tanto, obtenemos la segunda ley de Kepler, una partícula desplazándose bajo las fuerzas de un campo central barre áreas iguales en tiempos iguales.

Para obtener las ecuaciones de Hamilton usamos las ecuaciones de Euler-Lagrange. A su vez, usamos la ecuación de Newton y una fuerza conservativa para obtener las ecuaciones de Euler-Lagrange (4.8). En caso de que la fuerza

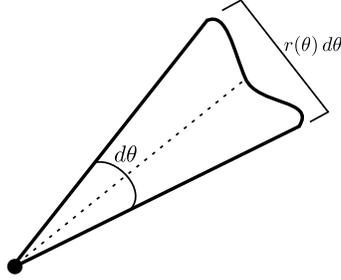


Figura 4.5: Área en coordenadas polares

también conste de una parte no conservativa las ecuaciones de Euler-Lagrange deben cambiarse por  $\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) = \frac{\partial L}{\partial q} + D_q$ , en donde  $D_q = \sum_{i=1}^3 F_i^{np} \frac{\partial x_i}{\partial q}$  (ecuaciones (4.10) y (4.9)). Si en coordenadas rectangulares tenemos que  $F_i^{np} = -m\lambda \dot{x}_i$  y  $L = T - U$ , como  $p_i = m\dot{x}_i = \frac{\partial L}{\partial \dot{x}_i}$  entonces la ecuación (4.10) se vuelve

$$\frac{dp_i}{dt} = - \left( \frac{\partial U}{\partial x_i} + \lambda p_i \right). \quad (4.17)$$

A continuación vamos a obtener estas ecuaciones usando un sistema Hamiltoniano de contacto (sección 3.1), esta parte está basada en [7]. Recordemos que al espacio  $M$  de todas las posibles posiciones de una partícula en  $\mathbb{R}^3$  le podemos asociar su espacio de 1-jets  $J^1(M) := T^*M \times \mathbb{R}$ . Este espacio viene equipado de manera intrínseca de una estructura de contacto  $\xi$  tal que si  $(\mathbf{q}, \mathbf{p})$  son coordenadas simplécticas de  $T^*M$  entonces la estructura de contacto  $\xi$  está definida por el kernel de la 1-forma  $\alpha = dy - \sum_i p_i dq_i$ . Así pues  $(\mathbf{q}, \mathbf{p}, y)$  son coordenadas de Darboux.

Supongamos que tenemos un sistema mecánico en  $T^*M$  con Hamiltoniano  $H = T + U: T^*M \rightarrow \mathbb{R}$ . Si consideramos la función  $\tilde{H}: J^1(M) \rightarrow \mathbb{R}$  definida por  $\tilde{H}(\mathbf{q}, \mathbf{p}, y) = H(\mathbf{p}, \mathbf{q}, ) + \lambda y$ , entonces el campo vectorial de contacto asociado al sistema Hamiltoniano de contacto  $(J^1(M), \alpha, \tilde{H})$  está dado por las ecuaciones:

$$\begin{aligned} \dot{p}_i &= - \left( \frac{\partial H}{\partial q_i} + p_i \lambda \right) \\ \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{y} &= \sum_{i=1}^n p_i \frac{\partial H}{\partial p_i} - H. \end{aligned}$$

En coordenadas rectangulares la primera de estas ecuaciones es la ecuación (4.17).

**Ejemplo 15.** En el ejemplo 11 vimos que si a un péndulo plano de masa  $m$  y longitud 1 le añadimos fuerzas de fricción proporcionales a la velocidad entonces

en coordenadas polares la ecuación de movimiento es:

$$\ddot{\theta} = -g \operatorname{sen}(\theta) - \lambda \dot{\theta}.$$

En este caso el Lagrangiano del sistema sin fricción es  $L = \frac{m}{2}\dot{\theta}^2 + mg \cos(\theta)$  y por lo tanto el Hamiltoniano es  $H = \frac{1}{2m}p_\theta^2 - mg \cos(\theta)$  con  $p_\theta = m\dot{\theta}$ . Así, las ecuaciones del campo vectorial asociado al sistema Hamiltoniano de contacto con función  $\tilde{H} = H + \lambda y$  son,

$$\begin{aligned} \dot{p}_\theta &= -(mg \operatorname{sen}(\theta) + p_\theta \lambda) \\ \dot{\theta} &= \frac{p_\theta}{m} \\ \dot{y} &= \frac{1}{2m}p_\theta^2 + mg \cos(\theta) - \lambda y. \end{aligned}$$

La primera de estas ecuaciones es la ecuación de movimiento de un péndulo con fricción.

**Ejemplo 16.** Si para una partícula en  $\mathbb{R}^3$  usamos coordenadas esféricas  $(r, \theta, \varphi)$ , en donde  $r$  es la distancia al origen,  $\theta$  es el ángulo entre el eje  $x$  y la proyección de la partícula al plano  $xy$ , y  $\varphi$  es el ángulo entre la partícula y el plano  $xy$  (figura 4.6), entonces, bajo las fuerzas debidas a un potencial  $U(r, \theta, \varphi)$  el Lagrangiano para este sistema es:

$$L = \frac{m}{2} \left( \dot{r}^2 + r^2 \cos^2(\varphi) \dot{\theta}^2 + r^2 \dot{\varphi}^2 \right) - U(r, \theta, \varphi).$$

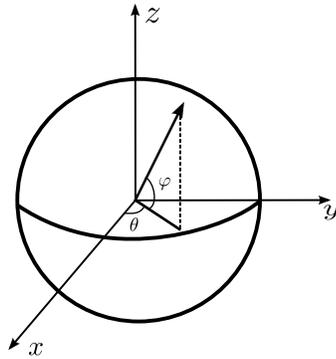


Figura 4.6: Coordenadas esféricas

Con esto las ecuaciones de Euler-Lagrange son,

$$\begin{aligned} \dot{p}_r &= mr(\cos^2(\varphi)\dot{\theta}^2 + \dot{\varphi}^2) - \frac{\partial U}{\partial r} \\ \dot{p}_\theta &= -\frac{\partial U}{\partial \theta} \\ \dot{p}_\varphi &= -mr^2 \cos(\varphi) \operatorname{sen}(\varphi) \dot{\theta}^2 - \frac{\partial U}{\partial \varphi}. \end{aligned}$$

Si añadimos fuerzas de fricciones proporcionales a la velocidad, como  $x = r \cos(\theta) \cos(\varphi)$ ,  $y = r \sin(\theta) \cos(\varphi)$  y  $z = r \sin(\varphi)$ , entonces de acuerdo a (4.9) tenemos que

$$D_r = -m\lambda\dot{r} = -\lambda p_r \quad D_\theta = -m\lambda r^2 \cos(\varphi)^2 \dot{\theta} = -\lambda p_\theta \quad D_\varphi = -m\lambda r^2 \dot{\varphi} = -\lambda p_\varphi.$$

Usando las igualdades,

$$\dot{r} = \frac{p_r}{m} \quad \dot{\theta} = \frac{p_\theta}{mr^2 \cos^2(\varphi)} \quad \dot{\varphi} = \frac{p_\varphi}{mr^2},$$

podemos escribir las ecuaciones de movimiento (4.10) como:

$$\begin{aligned} \dot{p}_r &= \frac{p_\theta^2}{mr^3 \cos^2(\varphi)} + \frac{p_\varphi^2}{mr^3} - \frac{\partial U}{\partial r} - \lambda p_r \\ \dot{p}_\theta &= -\frac{\partial U}{\partial \theta} - \lambda p_\theta \\ \dot{p}_\varphi &= -\frac{\sin(\varphi)}{\cos^3(\varphi)} \frac{p_\theta^2}{mr^2} - \frac{\partial U}{\partial \varphi} - \lambda p_\varphi. \end{aligned}$$

No es difícil ver que estas ecuaciones se obtienen del sistema Hamiltoniano de contacto con función

$$\tilde{H} = \frac{1}{2m} \left( p_r^2 + \frac{p_\theta^2}{r^2 \cos^2(\varphi)} + \frac{p_\varphi^2}{r^2} \right) + U(r, \theta, \varphi) + \lambda y = H + \lambda y,$$

en donde  $H$  es el Hamiltoniano de la partícula sin fuerzas de fricción.

## 4.6. El método de Hamilton-Jacobi

En esta sección presentamos el método de Hamilton-Jacobi, este método consiste en encontrar un cambio de coordenadas en las cuales las ecuaciones de movimiento sean fáciles de resolver. Dado un sistema mecánico con lagrangiano  $L$ , usamos el principio de mínima acción para deducir las ecuaciones de movimiento. Este principio asegura que el sistema físico sigue una trayectoria para la cual  $S = \int L dt$  adquiere un valor crítico. Fijemos ahora un punto  $q_0 \in M$  y un tiempo  $t_0$  y consideremos la función:

$$S_{(q_0, t_0)}(q, t) := \int_{q_0}^q L dt,$$

donde la integral es a lo largo de la trayectoria física que a tiempo  $t_0$  pasa por  $q_0$  y a tiempo  $t$  para por  $q$ . Si en las coordenadas locales  $q_1, \dots, q_n$  hacemos una variación en la  $i$ -ésima dirección del punto final  $q$  entonces, usando la igualdad (4.11) obtenemos,

$$\frac{\partial S}{\partial q_i} = \frac{\partial L}{\partial \dot{q}_i} + \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \right) dt.$$

Como la solución a lo largo de la cual se está calculando la integral del lado derecho satisface las ecuaciones de Euler-Lagrange obtenemos

$$\frac{\partial S}{\partial q_i} = \frac{\partial L}{\partial \dot{q}_i} = p_i. \quad (4.18)$$

Si ahora calculamos la derivada a lo largo de las soluciones,

$$L = \frac{dS}{dt} = \frac{\partial S}{\partial t} + \sum_{i=1}^n p_i \dot{q}_i.$$

De la definición (4.14) del Hamiltoniano  $H$  y sustituyendo la igualdad (4.18) obtenemos,

$$\frac{\partial S}{\partial t}(q_i, t) = -H\left(\frac{\partial S}{\partial q_i}, q_i, t\right). \quad (4.19)$$

A esta ecuación se le conoce como la *ecuación de Hamilton-Jacobi* y su importancia está en el hecho en que usando las soluciones a esta ecuación podemos encontrar coordenadas en las cuales las ecuaciones de movimiento son muy fáciles de resolver. Veremos esto a continuación.

En la proposición 2.1 vimos que el flujo Hamiltoniano preserva la 2-forma canónica  $\omega$  de  $T^*M$ . En este caso el campo vectorial  $X_H$  asociado las ecuaciones de Hamilton (4.15) satisface la identidad  $\omega(\cdot, X_H) = d_x H$  y entonces, aplicando la fórmula de Cartan y el hecho de que  $\omega$  es cerrada vemos que el flujo del campo vectorial  $X_H$  preserva la 2-forma  $\omega$ . A las transformaciones de  $T^*M$  que preservan  $\omega$  se les conoce como *transformaciones canónicas*. Su importancia radica en que si en coordenadas locales  $(\mathbf{p}, \mathbf{q})$  la transformación se expresa como  $\mathbf{P} = \mathbf{P}(\mathbf{p}, \mathbf{q}, t)$ ,  $\mathbf{Q} = \mathbf{Q}(\mathbf{p}, \mathbf{q}, t)$  entonces existe una función  $K$ , definida en el dominio de las nuevas coordenadas, tal que las ecuaciones de movimiento adquieren la forma:

$$\begin{aligned} \dot{P}_i &= -\frac{\partial K}{\partial Q_i} \\ \dot{Q}_i &= \frac{\partial K}{\partial P_i}. \end{aligned}$$

Es decir, en las nuevas coordenadas la función  $K$  funge como Hamiltoniano para las ecuaciones de movimiento.

En las coordenadas dadas por el flujo de las ecuaciones de Hamilton, ¿cuál es el nuevo Hamiltoniano  $K$ ? Denotemos por  $h_t$  a este flujo. Si en las coordenadas  $(\mathbf{p}, \mathbf{q})$  está dado por  $\mathbf{P} = \mathbf{P}(\mathbf{p}, \mathbf{q}, t)$ ,  $\mathbf{Q} = \mathbf{Q}(\mathbf{p}, \mathbf{q}, t)$ , *i.e.*  $(\mathbf{P}, \mathbf{Q}) = h_t(\mathbf{p}, \mathbf{q})$ , entonces, si en las nuevas coordenadas dejamos correr el flujo un tiempo  $\tau$ ,

$$h_\tau(\mathbf{P}, \mathbf{Q}) = h_\tau \circ h_t(\mathbf{p}, \mathbf{q}) = h_{t+\tau}(\mathbf{p}, \mathbf{q}) = (\mathbf{P}, \mathbf{Q}).$$

Esto quiere decir que en las nuevas coordenadas  $\dot{P}_i = \dot{Q}_i = 0$  y por lo tanto el Hamiltoniano  $K$  es una constante. ¿Cómo obtenemos las coordenadas  $(\mathbf{p}, \mathbf{q})$  en

función de las coordenadas  $(\mathbf{P}, \mathbf{Q})$ ? Supongamos que las funciones  $Q_i$  al tiempo  $t_0$  forman un sistema coordenado para un abierto de  $M$ . Si en la acción  $S$  ahora permitimos que el punto inicial  $q_0$  varíe en la carta determinada por las funciones  $Q_i$  al tiempo  $t_0$ . Entonces  $S = S(q, Q, t)$  y de manera análoga a como se dedujo la igualdad (4.18) obtenemos,

$$\frac{\partial S}{\partial Q_i} = -P_i. \quad (4.20)$$

Si  $\frac{\partial^2 S}{\partial q \partial Q}$  es invertible entonces podemos escribir  $\mathbf{q} = \mathbf{q}(\mathbf{P}, \mathbf{Q})$  y sustituyendo en la igualdad (4.18) obtenemos  $\mathbf{p} = \mathbf{p}(\mathbf{P}, \mathbf{Q})$ . Todo lo anterior lo podemos codificar en la 1-forma:

$$dS = \sum_{i=1}^n (p_i dq_i - P_i dQ_i) - H dt.$$

De manera más general, si  $\mathbf{P} = \mathbf{P}(\mathbf{p}, \mathbf{q}, t)$ ,  $\mathbf{Q} = \mathbf{Q}(\mathbf{p}, \mathbf{q}, t)$  es una transformación canónica y en estas nuevas coordenadas el Hamiltoniano es la función  $K$ , entonces existe una función diferenciable  $F$  tal que:

$$dF = \sum_{i=1}^n (p_i dq_i - P_i dQ_i) - (H - K) dt,$$

si la función  $F$  satisface la ecuación de Hamilton-Jacobi (4.19) entonces  $K = 0$  y por lo tanto en las nuevas coordenadas  $\dot{P}_i = \dot{Q}_i = 0$ , para ver más detalles de esto consultar [4]. Así, si encontramos una solución  $S$  de la ecuación de Hamilton-Jacobi que depende de manera diferenciable de  $n$  parámetros independientes  $Q_i$  entonces tomando  $P_i = -\frac{\partial S}{\partial Q_i}$  y aplicando el método anterior, podemos resolver las ecuaciones de movimiento.

Si  $\tilde{S}: M \times \mathbb{R} \rightarrow \mathbb{R}$  es una solución a la ecuación de Hamilton-Jacobi, a partir de ella podemos obtener una solución  $S$  a la misma ecuación pero que dependa de  $n$  parámetros. En efecto, sea  $S_0(\mathbf{Q}) = \tilde{S}(\mathbf{Q}, 0)$  y consideremos la función  $S(\mathbf{q}, t, \mathbf{Q}) = \tilde{S}(\mathbf{q}, t) - S_0(\mathbf{Q})$ , como las derivadas temporales y respecto  $q_i$  de  $S$  y  $\tilde{S}$  coinciden entonces  $\tilde{S}$  es una solución a la ecuación de Hamilton-Jacobi que depende de manera diferenciable de los parámetros  $Q_i$  pues  $\frac{\partial S}{\partial Q_i} = -\frac{\partial S_0}{\partial Q_i} = \frac{\partial \tilde{S}}{\partial q_i} \Big|_{t=0}$ .

Supongamos ahora que el Hamiltoniano  $H$  no depende del tiempo, sabemos que  $H$  se conserva en el tiempo y entonces debe ocurrir que  $H(q_i, \frac{\partial S}{\partial q_i}) = E$  y por lo tanto  $\frac{\partial S}{\partial t} = -E$ . Así, parece razonable buscar una solución de la forma  $S(q, t) = \tilde{S}(q) - Et$ , como la función  $S$  es una solución a la ecuación de Hamilton-Jacobi entonces

$$H \left( q_i, \frac{\partial \tilde{S}}{\partial q_i} \right) = E.$$

A la ecuación anterior se le conoce como la *ecuación de Hamilton-Jacobi libre de tiempo*. Supongamos que encontramos una solución  $\tilde{S}(q_i)$  a esta ecuación que

depende de manera diferenciable de los  $n$  parámetros  $Q_i$  y es tal que  $H(q_i, \frac{\partial \tilde{S}}{\partial q_i})$  es constante para cualesquiera  $Q_i$  fijos. Entonces, si hacemos

$$S(q_i, Q_i, t) = \tilde{S}(q_i, Q_i) - H\left(q_i, \frac{\partial \tilde{S}}{\partial q_i}\right) t = \tilde{S}(q_i, Q_i) - E(Q_i)t, \quad (4.21)$$

tenemos que  $S$  es una solución de la ecuación de Hamilton-Jacobi. Por lo tanto, si el Hamiltoniano  $H$  no depende del tiempo basta resolver la ecuación de Hamilton Jacobi libre de tiempo.

**Ejemplo 17.** Consideremos un sistema mecánico con un grado de libertad y Hamiltoniano  $H = \frac{p^2}{2m} + U(q)$ . La ecuación de Hamilton-Jacobi libre de tiempo es:

$$\frac{1}{2m} \left( \frac{\partial \tilde{S}}{\partial q} \right)^2 + U(q) = E = Q_1,$$

y por lo tanto,

$$\frac{\partial \tilde{S}}{\partial q} = \sqrt{2m(E - U(q))}.$$

Integrando y usando la fórmula (4.21), la solución a la ecuación de Hamilton-Jacobi es

$$S(q, E, t) = \sqrt{2m} \int_{q_0}^q \sqrt{E - U(x)} dx - Et.$$

De aquí obtenemos

$$\begin{aligned} p(q, E) &= \frac{\partial S}{\partial q} = \frac{\partial \tilde{S}}{\partial q} = \sqrt{2m(E - U(q))}. \\ -P(q, E, t) &= \frac{\partial S}{\partial E} = \sqrt{\frac{m}{2}} \int_{q_0}^q \frac{dx}{\sqrt{E - U(x)}} - t. \end{aligned}$$

Notemos que la cantidad  $E - U(q)$  es la energía cinética del sistema y por lo tanto  $\frac{\partial^2 S}{\partial q \partial E} > 0$  excepto en los puntos donde  $U$  adquiere su máximo valor posible  $E$ , así de manera local podemos escribir  $q = q(P, E, t)$ . En la expresión,

$$t - P(q, E, t) = \frac{\partial S}{\partial E} = \sqrt{\frac{m}{2}} \int_{q_0}^q \frac{dx}{\sqrt{E - U(x)}},$$

el lado derecho es la conocida fórmula para el tiempo que le toma a la partícula desplazarse del punto  $q_0$  al punto  $q$ . Así pues, en este caso las cantidades conservadas  $P_1, Q_1$  son el tiempo inicial y la energía respectivamente.

Arriba vimos que si el Hamiltoniano  $H$  no depende del tiempo, para encontrar una solución a la ecuación de Hamilton-Jacobi basta encontrar una función  $\tilde{S}(q_i, Q_i)$  tal que  $H(q_i, \frac{\partial \tilde{S}}{\partial q_i})$  sea constante para cada  $Q_i$  fijos. A continuación vamos a describir el método de separación de variables, cuando este método

funciona siempre es posible encontrar dicha función  $\tilde{S}$ . Supongamos que podemos escribir,

$$H\left(q_1, \frac{\partial \tilde{S}}{\partial q_1}, \dots, q_n, \frac{\partial \tilde{S}}{\partial q_n}\right) = H\left(q_1, \frac{\partial \tilde{S}}{\partial q_1}, \dots, q_{n-1}, \frac{\partial \tilde{S}}{\partial q_{n-1}}, \phi_n\left(q_n, \frac{\partial \tilde{S}}{\partial q_n}\right)\right),$$

si proponemos una solución de la forma  $\tilde{S} = S'(q_1, \dots, q_n) + S_n(q_n)$  entonces la ecuación de Hamilton-Jacobi libre de tiempo se vuelve,

$$H\left(q_1, \frac{\partial S'}{\partial q_1}, \dots, q_{n-1}, \frac{\partial S'}{\partial q_{n-1}}, \phi_n\left(q_n, \frac{\partial S_n}{\partial q_n}\right)\right) = E.$$

Resolviendo la ecuación diferencial ordinaria  $\phi_n\left(q_n, \frac{\partial S_n}{\partial q_n}\right) = Q_n$ . Ahora basta encontrar para cada  $Q_n$ , una solución de

$$H\left(q_1, \frac{\partial S'}{\partial q_1}, \dots, q_{n-1}, \frac{\partial S'}{\partial q_{n-1}}, Q_n\right) = E.$$

Si podemos escribir la expresión anterior como

$$H\left(q_1, \frac{\partial S'}{\partial q_1}, \dots, q_{n-2}, \frac{\partial S'}{\partial q_{n-2}}, \phi_{n-1}\left(q_{n-1}, \frac{\partial S'}{\partial q_{n-1}}, Q_n\right)\right),$$

entonces podemos volver a aplicar el método anterior. Si podemos aplicar  $i$  veces este método entonces obtenemos funciones  $S_n(q_n, Q_n), \dots, S_{n-i+1}(q_{n-i+1}, Q_{n-i+1}, \dots, Q_n)$ . En el  $n$ -ésimo paso y último, tenemos que resolver  $H\left(q_1, \frac{\partial S'}{\partial q_1}, Q_2, \dots, Q_n\right) = E$ .

**Ejemplo 18.** En el ejemplo 14 vimos que el Hamiltoniano para un partícula en  $\mathbb{R}^2$  bajo la influencia de un campo de fuerzas centrales es,

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + U(r).$$

La ecuación de Hamilton-Jacobi libre de tiempo es

$$\frac{1}{2m} \left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{2mr^2} \left(\frac{\partial S}{\partial \theta}\right)^2 + U(r) = E.$$

Para poder separar las variables de esta ecuación multiplicamos ambos lados de la ecuación anterior por  $2mr^2$  para obtener la ecuación equivalente,

$$r^2 \left(\frac{\partial S}{\partial r}\right)^2 - 2mr^2 (U(r) - E) + \left(\frac{\partial S}{\partial \theta}\right)^2 = 0.$$

El hecho de que el Hamiltoniano no dependa de la variable  $\theta$  ahora se traduce en que podemos separar a la variable  $\theta$  en la forma,

$$S_\theta(\theta, Q_2) = Q_2\theta,$$

y por lo tanto  $p_\theta = Q_2$ . Ahora tenemos que resolver la ecuación

$$r^2 \left( \frac{\partial S_r}{\partial r} \right)^2 - 2mr^2 (U(r) - E) + Q_2^2 = 0.$$

La solución a esta ecuación es:

$$S_r(r, Q_2, E) = \int_{r_0}^r \sqrt{2m(U(R) - E) - \frac{Q_2^2}{R^2}} dR.$$

Por lo tanto las cantidades conservadas en este ejemplo son la energía  $E = Q_1$  y el momento angular al cuadrado  $Q_2 = p_\theta^2$ . La solución a la ecuación de Hamilton-Jacobi es:

$$S(r, \theta, E, Q_2, t) = \int_{r_0}^r \sqrt{2m(U(R) - E) - \frac{Q_2^2}{R^2}} dR + Q_2\theta - Et.$$

Tenemos que,

$$P_2(r, \theta, E, Q_2, t) = -\frac{\partial S}{\partial Q_2} = \int_{r_0}^r \frac{Q_2}{R^2} \frac{dR}{\sqrt{2m(U(R) - E) - \frac{Q_2^2}{R^2}}} - \theta,$$

y por lo tanto,

$$\theta(r, E, P_2, Q_2) = \int_{r_0}^r \frac{Q_2}{R^2} \frac{dR}{\sqrt{2m(U(R) - E) - \frac{Q_2^2}{R^2}}} - P_2. \quad (4.22)$$

Si al tiempo  $t_0$  tenemos que  $r(t_0) = r_0$  y  $\theta(t_0) = \theta_0$  entonces  $\theta_0 = \theta(r_0, E, P_2) = -P_2$ . Así,  $-P_2$  es el ángulo que la partícula forma con el eje  $x$  al tiempo  $t_0$ . Por otro lado,

$$P_1(r, E, Q_2, t) = t - m \int_{r_0}^r \frac{dR}{\sqrt{2m(U(R) - E) - \frac{Q_2^2}{R^2}}}, \quad (4.23)$$

así,  $P_1 = P_1(r_0, E, Q_2, t_0) = t_0$  es el tiempo inicial. Además,

$$\frac{\partial P_1}{\partial r} = -\frac{m}{\sqrt{2m(U(r) - E) - \frac{Q_2^2}{r^2}}},$$

y usando el teorema de la función inversa podemos obtener una expresión de la forma  $r = r(P_1, E, Q_2, t)$ . Por lo tanto, si fijamos la energía  $E$ , el momento angular  $Q_2$ , el tiempo inicial  $P_1 = t_0$  y el momento angular  $Q_2$ , la expresión (4.22) y la función inversa respecto a  $r$  de (4.23) nos dan al radio  $r$  y ángulo  $\theta$  como funciones del tiempo.

**Ejemplo 19.** En este ejemplo vamos a estudiar una partícula en  $\mathbb{R}^2$  bajo los efectos de la gravedad debida a dos centros fijos. Para esto vamos a usar coordenadas elípticas, si tomamos coordenadas cartesianas de modo que los dos

centros de gravedad tengan coordenadas  $(\pm c, 0)$ , las coordenadas elípticas están definidas como

$$x(\varphi, \theta) = c \cosh(\varphi) \cos(\theta) \quad y(\varphi, \theta) = c \sinh(\varphi) \sin(\theta).$$

Para cada  $\varphi_0 \in \mathbb{R}$  fijo tenemos que

$$\frac{x^2}{(c \cosh(\varphi_0))^2} + \frac{y^2}{(c \sinh(\varphi_0))^2} = \frac{(c \cosh(\varphi) \cos(\theta))^2}{(c \cosh(\varphi_0))^2} + \frac{(c \sinh(\varphi) \sin(\theta))^2}{(c \sinh(\varphi_0))^2} = 1,$$

y por lo tanto las curvas  $(x(\varphi_0, \theta), y(\varphi_0, \theta))$ , con  $\varphi_0 \in \mathbb{R}$  son elipses. Como  $\cosh(\varphi_0) > \sinh(\varphi_0)$  estas elipses tienen semieje mayor  $a = c \cosh(\varphi_0)$  y semieje menor  $b = c \sinh(\varphi_0)$ . Si los focos de esta elipse tienen coordenadas  $(\pm \tilde{c}, 0)$  entonces  $\tilde{c}^2 = a^2 - b^2 = c^2$  y por ende todas estas elipses tienen a sus focos en los centros de gravedad. Si ahora hacemos  $\theta = \theta_0$  fijo, de manera análoga podemos concluir que las curvas  $(x(\varphi, \theta_0), y(\varphi, \theta_0))$  son hipérbolas cuyos focos están en los centros de gravedad.

Tenemos que,

$$\begin{aligned} \dot{x} &= c(\cos(\theta) \sinh(\varphi) \dot{\varphi} - \sin(\theta) \cosh(\varphi) \dot{\theta}) \\ \dot{y} &= c(\sin(\theta) \cosh(\varphi) \dot{\varphi} + \cos(\theta) \sinh(\varphi) \dot{\theta}), \end{aligned}$$

y con esto

$$\begin{aligned} \dot{x}^2 + \dot{y}^2 &= c^2(\cos^2(\theta) \sinh^2(\varphi) + \sin^2(\theta) \cosh^2(\varphi))(\dot{\varphi}^2 + \dot{\theta}^2) \\ &= c^2(\cosh^2(\varphi) - \cos^2(\theta))(\dot{\varphi}^2 + \dot{\theta}^2). \end{aligned}$$

Así, la energía cinética en coordenadas elípticas es

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) = \frac{m}{2}c^2(\cosh^2(\varphi) - \cos^2(\theta))(\dot{\varphi}^2 + \dot{\theta}^2).$$

La energía potencial debida a los dos centros de gravedad es

$$U(r_1, r_2) = -\left(\frac{k_1}{r_1} + \frac{k_2}{r_2}\right),$$

en donde  $r_1 = \sqrt{(x+c)^2 + y^2}$ ,  $r_2 = \sqrt{(x-c)^2 + y^2}$  son las distancias de la partícula a los centros de gravedad y  $k_1, k_2$  son constantes positivas. Ahora,

$$\begin{aligned} r_1 &= \sqrt{(c \cosh(\varphi) \cos(\theta) + c)^2 + (c \sinh(\varphi) \cos(\theta))^2} \\ &= c\sqrt{\cosh^2(\varphi) \cos^2(\theta) + \sinh^2(\varphi) \sin^2(\varphi) + 2 \cosh(\varphi) \cos(\theta) + 1} \\ &= c\sqrt{\cosh^2(\varphi) - \sin^2(\theta) + 1 + 2 \cosh(\varphi) \cos(\theta)} \\ &= c|\cosh(\varphi) + \cos(\theta)|. \end{aligned}$$

Como el valor mínimo de  $\cosh(\varphi)$  es 1 y el valor máximo de  $\cos(\theta)$  es 1 tenemos que  $r_1 = c(\cosh(\varphi) + \cos(\theta))$ . De manera análoga podemos ver que  $r_2 = c(\cosh(\varphi) - \cos(\theta))$  y con esto obtenemos que

$$U(\varphi, \theta) = -\frac{(k_1 + k_2) \cosh(\varphi) + (k_2 - k_1) \cos(\theta)}{c(\cosh^2(\varphi) - \cos^2(\theta))} = -\frac{k \cosh(\varphi) + \tilde{k} \cos(\theta)}{c(\cosh^2(\varphi) - \cos^2(\theta))}.$$

Así, el Lagrangiano para este sistema es,

$$L = \frac{m}{2} c^2 (\cosh^2(\varphi) - \cos^2(\theta)) (\dot{\varphi}^2 + \dot{\theta}^2) + \frac{k \cosh(\varphi) + \tilde{k} \cos(\theta)}{c(\cosh^2(\varphi) - \cos^2(\theta))}.$$

Con esto el Hamiltoniano es,

$$H = \frac{p_\varphi^2 + p_\theta^2 - 2mc(k \cosh(\varphi) + \tilde{k} \cos(\theta))}{2mc^2(\cosh^2(\varphi) - \cos^2(\theta))}.$$

Así, la ecuación de Hamilton-Jacobi libre de tiempo de este sistema es:

$$\frac{\left(\frac{\partial S}{\partial \varphi}\right)^2 + \left(\frac{\partial S}{\partial \theta}\right)^2 - 2mc(k \cosh(\varphi) + \tilde{k} \cos(\theta))}{2mc^2(\cosh^2(\varphi) - \cos^2(\theta))} = E,$$

que podemos escribir como

$$\begin{aligned} \left(\frac{\partial S}{\partial \varphi}\right)^2 - 2mc \cosh(\varphi)(Ec \cosh(\varphi) + k) + \left(\frac{\partial S}{\partial \theta}\right)^2 + 2mc \cos(\theta)(Ec \cos(\theta) - \tilde{k}) \\ = 0 \end{aligned}$$

De la expresión anterior es claro que las variables se pueden separar. Primero resolvemos

$$\left(\frac{\partial S_\varphi}{\partial \varphi}\right)^2 - 2mc \cosh(\varphi)(Ec \cosh(\varphi) + k) = Q_2,$$

que tiene por solución a

$$S_\varphi(\varphi, E, Q_2) = \int_{\varphi_0}^{\varphi} \sqrt{Q_2 + 2mc \cosh(\phi)(Ec \cosh(\phi) + k)} d\phi.$$

Con esto ahora tenemos que resolver la ecuación

$$Q_2 + \left(\frac{\partial S_\theta}{\partial \theta}\right)^2 + 2mc \cos(\theta)(Ec \cos(\theta) - \tilde{k}) = 0$$

cuya solución es

$$S_\theta(\theta, E, Q_2) = \int_{\theta_0}^{\theta} \sqrt{-\left(Q_2 + 2mc \cos(\alpha)(Ec \cos(\alpha) - \tilde{k})\right)} d\alpha.$$

Para finalizar este capítulo vamos a analizar la ecuación de Hamilton-Jacobi para sistemas Hamiltonianos de contacto. Al final de la sección anterior vimos que este tipo de sistemas pueden modelar algunos sistemas mecánicos en los cuales aparecen fuerzas disipativas que impiden la conservación de la energía. En el ejemplo 5 vimos que, en las coordenadas de Darboux  $(\mathbf{q}, \mathbf{p}, y)$ , las características de la ecuación diferencial parcial,

$$\frac{\partial S}{\partial t} + H\left(\mathbf{q}, \frac{\partial S}{\partial \mathbf{q}}, S, t\right) = 0,$$

coinciden con las ecuaciones que definen al campo vectorial  $X_H$  asociado al Hamiltoniano  $H$ . Veamos ahora que bajo ciertas condiciones una solución  $S(\mathbf{q}, \mathbf{Q}, t)$  a la ecuación de Hamilton-Jacobi nos ayuda a encontrar las curvas integrales del campo vectorial  $X_H$ . Primero, para cada  $\mathbf{Q}$  fija definamos  $p_i := \frac{\partial S}{\partial q_i}$  y  $P_i(\mathbf{q}, t) = \frac{\partial S}{\partial Q_i}$ . Tenemos que,

$$\frac{dP_i}{dt} = \sum_{j=1}^n \frac{\partial^2 S}{\partial q_j \partial Q_i} \frac{dq_j}{dt} + \frac{\partial^2 S}{\partial t \partial Q_i},$$

además

$$\frac{\partial^2 S}{\partial Q_i \partial t} = - \left( \sum_{j=1}^n \frac{\partial H}{\partial p_j} \frac{\partial p_j}{\partial Q_i} + \frac{\partial H}{\partial y} \frac{\partial S}{\partial Q_i} \right) = - \left( \sum_{j=1}^n \frac{\partial H}{\partial p_j} \frac{\partial^2 S}{\partial q_j \partial Q_i} + \frac{\partial H}{\partial y} P_i \right),$$

y así,

$$\frac{dP_i}{dt} = \sum_{j=1}^n \frac{\partial^2 S}{\partial q_j \partial Q_i} \left( \frac{dq_j}{dt} - \frac{\partial H}{\partial p_j} \right) - \frac{\partial H}{\partial y} P_i.$$

Si imponemos las condiciones  $\det\left(\frac{\partial^2 S}{\partial \mathbf{q} \partial \mathbf{Q}}\right) \neq 0$  y  $\frac{dP_i}{dt} = -\frac{\partial H}{\partial y} P_i$  entonces obtenemos las condiciones

$$\frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}. \quad (4.24)$$

Ahora, para  $\mathbf{Q}$  fijo, un cálculo similar al anterior muestra que

$$\frac{dp_i}{dt} = \sum_{j=1}^n \frac{\partial^2 S}{\partial q_j \partial q_i} \left( \frac{dq_j}{dt} - \frac{\partial H}{\partial p_j} \right) - \left( \frac{\partial H}{\partial q_i} + \frac{\partial H}{\partial y} p_i \right) = - \left( \frac{\partial H}{\partial q_i} + \frac{\partial H}{\partial y} p_i \right). \quad (4.25)$$

Por último,

$$\frac{dS}{dt} = \sum_{j=1}^n \frac{\partial S}{\partial q_j} \frac{dq_j}{dt} + \frac{\partial S}{\partial t} = \sum_{j=1}^n p_j \frac{\partial H}{\partial p_j} - H. \quad (4.26)$$

Las ecuaciones (4.24), (4.25) y (4.26) son las mismas que las ecuaciones (3.3) y (3.4). Así, bajo las condiciones  $\det\left(\frac{\partial^2 S}{\partial \mathbf{q} \partial \mathbf{Q}}\right) \neq 0$  y  $\frac{dP_i}{dt} = -\frac{\partial H}{\partial y} P_i$  podemos aplicar el método de Hamilton-Jacobi al igual que en el caso de sistemas conservativos. Para ver un ejemplo de esto consultar [7].

## Capítulo 5

# Relatividad Especial

Al principio del capítulo anterior presentamos las leyes de Newton, estas leyes dependen fuertemente de estar utilizando un sistema de referencia inercial como sistema coordinado. Después vimos que bajo ciertas condiciones las ecuaciones de Newton son equivalentes a las ecuaciones de Euler-Lagrange, la ventaja de estas últimas es que no dependen del sistema coordinado que se está utilizando, en unas coordenadas  $q_1, \dots, q_n$  cualesquiera basta calcular el Lagrangiano  $L$  para obtener dichas ecuaciones. Una desventaja de este método es que nociones físicas muy claras como velocidad y momento se pierden y en su lugar hablamos de velocidades generalizadas  $\dot{\mathbf{q}}$  y momentos generalizados  $\frac{\partial L}{\partial \dot{\mathbf{q}}}$ . En este capítulo volveremos a usar sistemas de referencia inerciales para discutir algunas nociones físicas de mucho interés para la relatividad espacial.

A finales del siglo XIX se creía que había una sustancia llamada éter la cual se encontraba en el espacio. Estrellas, planetas y demás objetos celestes se desplazaban a través del éter, la luz también usaba el éter para propagarse a través del espacio. Así pues hubo un gran interés por medir la velocidad de esta sustancia. En 1887 Michelson y Morley diseñaron un experimento para determinar la dirección en la cual se movía el éter, de acuerdo a este experimento la luz debería moverse más rápido cuando la dirección del vector que va del Sol a la Tierra coincide con la dirección en la que el éter se desplaza. Como bien sabemos los resultados del experimento fueron desconcertantes, la velocidad de la luz parecía ser la misma sin importar la dirección en la que se le medía. Esto contrastaba fuertemente con un principio del cual casi todos hemos sido testigos cuando viajamos en un carro en una avenida de dos sentidos: Si un observador A se mueve a velocidad constante  $v$  respecto a otro observador B, entonces si un objeto se mueve con velocidad  $w$  respecto a B, el mismo objeto se mueve con velocidad  $v + w$  respecto al sistema A. Esto es una consecuencia del principio de relatividad galileano.

## 5.1. El principio de relatividad

Si nos encontramos viajando en un carro o en el metro más o menos a velocidad constante, prácticamente todos hemos sido testigos de que los objetos dentro del vehículo en cuestión se comportan casi como se comportarían si estuvieran en un parque o en un laboratorio los cuales, para los ojos humanos, parecen estar en un reposo a diferencia del carro o el metro. Si estamos observando algún fenómeno físico y giramos nuestro cuerpo en alguna dirección, las leyes de la física obviamente son las mismas una vez que hemos girado a todos los objetos en cuestión en la misma dirección en la que giramos nuestro cuerpo. A través del tiempo hay muchos fenómenos físicos que no cambian, la gravedad o un animal corriendo se comportan de la misma manera hoy que hace 1000 años. Lo mismo podemos decir de la fuerza y el perro aquí y en China, en ambos lugares las leyes que rigen el movimiento son las mismas.

Podemos sintetizar todo lo anterior de la manera siguiente: las ecuaciones de movimiento son invariantes bajo un cierto grupo de transformaciones. Para que lo anterior tenga sentido debemos precisar 1) quiénes son las ecuaciones de movimiento y 2) cuál es el grupo de transformaciones. En el caso de la mecánica clásica tenemos a la ecuación de Newton  $m\ddot{x} = F(x, \dot{x}, t)$ . La ecuación de Newton sólo es válida en sistemas de referencia inerciales y como mencionamos al principio de esta sección, otro sistema de referencia que se mueve a velocidad constante respecto al primero también parece ser un sistema de referencia inercial. Para cada tiempo  $t$  las coordenadas de ambos sistemas están relacionadas por una traslación  $x = x' + vt$ , donde  $x'$  son las coordenadas del segundo sistema y  $v$  es la velocidad con la que el primer sistema coordenado se mueve respecto al segundo. También mencionamos que las rotaciones, traslaciones espaciales y temporales parecen no alterar las leyes de la mecánica.

La fuerza es una función  $F: \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3$  y ahora postulamos que esta fuerza debe de ser invariante bajo movimientos a velocidad constante  $g(x, t) = (x + vt, t)$ , traslaciones  $g(x, t) = (x + a, t + s)$  y rotaciones  $g(x, t) = (Ax, t)$  con  $A$  una rotación. Al grupo formado por estas transformaciones se le conoce como *el grupo de Galileo* y al enunciado: las ecuaciones de Newton son invariantes bajo el grupo de Galileo se le conoce como *el principio de relatividad Galileano*. De este principio se sigue que la fuerza no depende del tiempo, de la misma manera, la fuerza que ejerce una partícula situada en el punto  $x_i$  en otra partícula situada en el punto  $x_j$  no depende de las posiciones de  $x_i$  y  $x_j$  pues las ecuaciones de movimiento son invariantes bajo traslaciones espaciales. Así esta fuerza sólo depende de las posiciones relativas  $x_i - x_j$  de ambas partículas. La invarianza bajo movimientos a velocidad constante nos dice que la fuerza entre estas dos partículas tampoco puede depender de las velocidades  $\dot{x}$  (la velocidad de una partícula depende del sistema inercial que estemos usando) si no sólo de su velocidad relativa  $\dot{x}_i - \dot{x}_j$ , más aún, la invarianza bajo rotaciones nos dice que la fuerza no depende de la dirección de los vectores  $x_i - x_j$  y  $\dot{x}_i - \dot{x}_j$  sino sólo de su norma.

Es claro que hay muchos fenómenos físicos que no satisfacen este principio de

relatividad, por ejemplo, no es lo mismo poner un satélite con cierta velocidad inicial alrededor de la Tierra que ponerlo con esa misma velocidad alrededor del Sol, sin embargo existen muchísimos fenómenos físicos que, con un alto grado de precisión, si lo satisfacen. Por último, si tenemos dos sistemas de referencia inerciales y hacemos una rotación para que sus ejes coordenados coincidan, tenemos que las coordenadas de ambos se relacionan por  $x = x' + vt$  y en consecuencia

$$w = \frac{dx}{dt} = \frac{dx'}{dt} + v = w' + v.$$

Es decir, si el observador primado  $x'$  ve que un objeto se desplaza con velocidad  $w$ , entonces el observador no primado  $x$  ve a este mismo objeto moverse con velocidad  $w' + v$ . Posiblemente este último hecho es bien conocido y aceptado por todos los humanos pero como mencionamos al principio del capítulo, a principios del siglo XIX la gente empezó a dudar de esto pues los experimentos mostraban que la luz tiene la misma velocidad en cualquier sistema de referencia inercial.

## 5.2. Fundamentos de la relatividad especial

Desde que la gente aprendió la gran utilidad del principio de relatividad Galileano y las leyes de Newton, hasta finales del siglo XIX, nadie dudaba de la utilidad y veracidad de estos conceptos. Esto se debió en gran medida a que los fenómenos que no satisfacen estos principios con un buen grado de precisión, se mueven, de acuerdo a un sistema de referencia inercial, con velocidades cercanas a la velocidad de la luz. En esta sección veremos algunos de los fundamentos de la relatividad especial, esta teoría funciona para describir la mecánica de partículas que se mueven con velocidades cercanas a la velocidad de la luz.

La primera diferencia de la relatividad especial con la mecánica clásica es en la forma en que interactúan las partículas. En la mecánica clásica la independencia de la fuerza  $F$  del tiempo y las velocidades implica que la interacción de dos partículas ocurre de manera inmediata, por ejemplo, la fuerza de gravedad entre dos planetas se presenta de manera inmediata sin importar que se encuentren en el sistema solar o en galaxias distintas.

Así pues, un primer postulado de la relatividad especial es el siguiente: la interacción entre dos partículas o una partícula y un campo de fuerzas tiene un tiempo de propagación y antes de ese tiempo las partículas no se ven afectadas por las otras partículas o el campo. Para todos los fenómenos físicos existe una velocidad máxima de propagación, esta velocidad determina el tiempo tras el cual las partículas comienzan a interactuar. El hecho de que haya una velocidad máxima de propagación implica que ninguna partícula se puede mover con una velocidad mayor a la velocidad máxima de propagación. Experimentalmente se ha comprobado que nada se puede mover más rápido que la velocidad de la luz y, como mencionamos al principio del capítulo, la velocidad de la luz es la misma en todos los sistemas de referencia inerciales. La velocidad de la luz tiene un valor de  $c = 2,988 \times 10^{10} \text{ cm/seg.}$

El grandísimo valor de la velocidad de la luz explica por qué la mecánica clásica, que no toma en cuenta la propagación de interacciones, funciona tan bien: en un sistema de referencia inercial, cuando las velocidades de nuestras partículas son muy pequeñas en comparación con la velocidad de la luz entonces la velocidad de la luz tiene un valor relativo muy grande y pareciera que todas las interacciones suceden de manera inmediata. Más adelante, en las fórmulas que obtendremos, veremos que si el valor relativo de  $c$  tiende a infinito entonces recuperamos la mecánica clásica.

En lo siguiente el tiempo va a jugar un papel muy importante, vamos a describir un evento mediante sus tres coordenadas espaciales  $(x, y, z)$  y el tiempo  $t$ . Así un evento quedará especificado por un punto en  $\mathbb{R}^4$ , donde la primera coordenada corresponde al tiempo. A este conjunto le llamaremos *el espacio tiempo* y a veces llamaremos a los puntos del espacio tiempo como *eventos*. A continuación vamos a expresar de manera matemática el hecho de que la luz tiene la misma velocidad en todos los marcos de referencia inerciales. Supongamos que en el sistema de referencia inercial  $(t, x, y, z)$  se emite una señal que se propaga a la velocidad de la luz desde el punto  $a = (t_1, x_1, y_1, z_1)$ . Si la señal llega al punto  $b = (t_2, x_2, y_2, z_2)$  entonces la señal recorrió una distancia  $c(t_2 - t_1)$  y esta distancia al cuadrado es igual a  $(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$  y por lo tanto

$$0 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2.$$

Si en otro sistema de referencia inercial los puntos  $a$  y  $b$  tienen coordenadas  $(t'_1, x'_1, y'_1, z'_1)$  y  $(t'_2, x'_2, y'_2, z'_2)$  entonces el mismo razonamiento nos lleva a concluir que

$$\begin{aligned} c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 = \\ c^2(t'_2 - t'_1)^2 - (x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2. \end{aligned}$$

A la cantidad  $s^2 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$  se le conoce como el intervalo entre los eventos  $(t_1, x_1, y_1, z_1)$  y  $(t_2, x_2, y_2, z_2)$ . De la invarianza de la velocidad de la luz se sigue que si en un sistema de referencia inercial el intervalo entre dos eventos es cero, entonces es cero en todos los demás sistemas de referencia inerciales. Si los eventos están infinitamente cerca podemos escribir

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2.$$

Hemos visto que si  $ds = 0$  en un sistema de referencia inercial entonces  $ds' = 0$  en cualquier otro sistema de referencia inercial, así para cada  $p \in \mathbb{R}^4$ ,  $ds^2$  y  $ds'^2$  determinan la misma cuádrlica en cada plano tangente  $T_p\mathbb{R}^4$ . Esta cuádrlica resulta ser un cono al cual se le conoce como *el cono de luz*. De esto concluimos que  $ds^2 = a ds'^2$  donde  $a$  es una función que sólo depende del valor absoluto de la velocidad relativa de ambos sistemas de referencia y no de las coordenadas espaciales ni del tiempo pues esto contradiría el hecho a que, en ausencia de fuerzas el espacio es homogéneo e isotrópico (el espacio se ve igual en todos los

puntos (homogeneidad) y en todas las direcciones (isotropía)). Consideremos tres sistemas de referencia inerciales  $K, K_1$  y  $K_2$  y sean  $v_1, v_2$  las velocidades de  $K_1$  y  $K_2$  relativas a  $K$  y  $v_{12}$  la velocidad de  $K_2$  relativa a  $K_1$ . Tenemos que

$$ds^2 = a(|v_1|)ds_1^2 \quad ds^2 = a(|v_2|)ds_2^2 \quad ds_1^2 = a(|v_{12}|)ds_2^2,$$

entonces,

$$ds^2 = a(|v_1|)a(|v_2|)ds_2^2 = \frac{a(|v_1|)}{a(|v_2|)}a(|v_{12}|)ds^2, \quad (5.1)$$

de aquí obtenemos

$$\frac{a(|v_2|)}{a(|v_1|)} = a(|v_{12}|).$$

Pero  $v_{12}$  depende de la magnitud de  $v_1$  y  $v_2$  y del ángulo entre ellos (por ejemplo si  $v_1$  y  $v_2$  son colineales obtenemos una cierta  $v_{12}$  y para  $v_1$  y  $-v_2$  claramente obtenemos una  $v_{12}$  distinta, ver figura 5.2) y entonces  $a$  tiene que ser una constante, la igualdad (5.1) nos dice que esta constante es igual a 1. Por lo tanto concluimos que  $ds = ds'$ , esto quiere decir que el intervalo entre dos eventos es el mismo en cualquier sistema de referencia inercial.

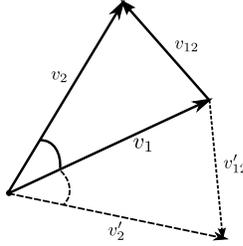


Figura 5.1: El vector  $v_{12}$  depende de  $v_1$  y  $v_2$

Veamos ahora que si para cualesquiera dos sistemas de referencia inerciales ocurre que  $ds = ds'$  entonces una partícula que se mueve a la velocidad de la luz en uno de los dos sistemas coordenados, también se mueve a la velocidad de la luz en cualquier otro sistema coordenado. En efecto, en el sistema coordenado en el cual la partícula se mueve a la velocidad de la luz tenemos que

$$c^2 = v^2 = \frac{dx^2 + dy^2 + dz^2}{dt^2},$$

y así  $ds = 0$ , como  $ds' = ds$  entonces en el otro sistema de referencia inercial  $(t', x', y', z')$  ocurre que  $c^2 dt'^2 = dx'^2 + dy'^2 + dz'^2$  y por lo tanto,

$$v' = \frac{dx'^2 + dy'^2 + dz'^2}{dt'^2} = c^2.$$

Resumiendo tenemos la siguiente proposición:

**Proposición 5.1.** *El hecho de que la velocidad de la luz sea igual en todos los sistemas de referencia inerciales es equivalente a que  $c^2 dt^2 - dx^2 - dy^2 - dz^2 = c^2 dt'^2 - dx'^2 - dy'^2 - dz'^2$  para cualesquiera dos sistemas de referencia inerciales.*

Supongamos ahora que tenemos dos eventos  $(t_1, x_1, y_1, z_1)$  y  $(t_2, x_2, y_2, z_2)$  tales que en otro marco de referencia ambos eventos suceden en el mismo punto del espacio físico, es decir,  $x'_2 = x'_1, y'_2 = y'_1$  y  $z'_2 = z'_1$ . Así, tenemos que  $s^2 = s'^2 = c^2(t'_2 - t'_1)^2 > 0$ . En la siguiente sección veremos que la condición  $s^2 > 0$  es suficiente para la existencia de un sistema de referencia inercial en el cual ambos eventos ocurren en el mismo punto del espacio físico.

Si para los dos eventos existe un sistema de referencia inercial en el cual  $t'_2 = t'_1$ , los eventos son simultáneos y entonces  $s^2 = s'^2 = -((x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2) < 0$ . También veremos que esta última condición,  $s^2 < 0$ , basta para que en algún sistema de referencia inercial ambos eventos sean simultáneos.

### 5.3. La transformación de Lorentz

En la sección anterior vimos que la cantidad  $ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$  no depende del sistema de referencia inercial que estemos utilizando para dar coordenadas a los eventos del espacio tiempo. Así pues la transformación que relaciona las coordenadas de dos sistemas de referencia inerciales debe de preservar esta forma cuadrática, al grupo de transformaciones que preservan esta cantidad y la orientación del eje temporal se le conoce como *el grupo de Lorentz*. Notemos que todas las rotaciones del espacio físico  $(x, y, z)$  preservan a la forma cuadrática  $ds^2$  pues ellas preservan la forma  $dx^2 + dy^2 + dz^2$  y la coordenada del tiempo no se ve afectada por esta rotación. Supongamos que tenemos dos sistemas de referencia inerciales  $(t, x, y, z)$  y  $(t', x', y', z')$  y que el sistema primado se mueve con velocidad  $V$  respecto al primero, vamos a calcular la transformación de Lorentz que relaciona a ambos sistemas coordenados. Por lo anterior podemos suponer que los ejes  $(x, y, z)$  y  $(x', y', z')$  coinciden y que  $V$  está en la dirección del eje  $x$ . Así,  $y = y', z = z'$  y por lo tanto la transformación de Lorentz que relaciona las coordenadas de ambos sistemas de referencia debe de preservar a la forma  $c^2 dt^2 - dx^2$ . Esta forma cuadrática define una familia de hipérbolas ( $c^2 dt^2 - dx^2 \neq 0$ ) y un cono ( $c^2 dt^2 - dx^2 = 0$ ) en el plano  $(t, x)$ . Por lo tanto la transformación de Lorentz que buscamos debe de preservar esta familia de curvas. Al igual que las transformaciones que fijan a circunferencias concéntricas en el origen son rotaciones de la forma

$$x = x' \cos(\theta) - y' \sin(\theta) \qquad y = x' \sin(\theta) + y' \cos(\theta),$$

las transformaciones que preservan nuestra familia de hipérbolas son de la forma

$$x = x' \cosh(\psi) + ct' \sinh(\psi) \qquad t = \frac{x'}{c} \sinh(\psi) + t' \cosh(\psi).$$

En efecto de las expresiones anteriores obtenemos

$$\begin{aligned} dt^2 &= \frac{\sinh^2(\psi)}{c^2} dx'^2 + \frac{2}{c} \sinh(\psi) \cosh(\psi) dx' dt' + \cosh^2(\psi) dt'^2 \\ dx^2 &= \cosh^2(\psi) dx'^2 + 2c \cosh(\psi) \sinh(\psi) dx' dt' + c^2 \sinh^2(\psi) dt'^2, \end{aligned}$$

y por lo tanto,

$$\begin{aligned} c^2 dt^2 - dx^2 &= (\sinh^2(\psi) - \cosh^2(\psi)) dx'^2 + c^2(\cosh^2(\psi) + \sinh^2(\psi)) dt'^2 \\ &= c^2 dt'^2 - dx'^2. \end{aligned}$$

Para escribir esta transformación en términos de la velocidad  $V$  con la cual el sistema primado se mueve respecto al no primado, notemos que si en las ecuaciones anteriores ponemos  $x' = 0$  obtenemos  $x = ct' \sinh(\psi)$  y  $t = t' \cosh(\psi)$  y  $x$  es la posición en el espacio que el sistema no primado asigna al sistema de referencia primado. Por ende

$$V = \frac{dx}{dt} = \frac{c dt' \sinh(\psi)}{dt' \cosh(\psi)} = c \tanh(\psi),$$

de la identidad  $\cosh^2(\psi) - \sinh^2(\psi) = 1$  obtenemos  $\cosh^2(\psi) = \frac{1}{1 - \tanh^2(\psi)}$  y sustituyendo lo anterior obtenemos

$$\cosh(\psi) = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}} \quad \sinh(\psi) = \frac{\frac{V}{c}}{\sqrt{1 - \frac{V^2}{c^2}}}.$$

Si sustituimos esto en la transformación de Lorentz obtenemos que las coordenadas de ambos sistemas de referencia inerciales están relacionadas por:

$$x = \frac{x' + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}} \quad t = \frac{t' + \frac{V}{c^2}x'}{\sqrt{1 - \frac{V^2}{c^2}}}. \quad (5.2)$$

Notemos que cuando  $c \rightarrow \infty$  (físicamente esto quiere decir que el valor de la velocidad  $V$  con la que el sistema primado se mueve respecto al sistema no primado es muy pequeña en comparación con la velocidad de la luz) recuperamos la transformación del grupo de Galileo

$$x = x' + Vt' \quad t = t'.$$

Para obtener la transformación inversa de la transformación de Lorentz que acabamos de obtener notemos que, de acuerdo al sistema primado el sistema no primado se mueve con velocidad  $-V$  y por lo tanto

$$x' = \gamma(x - Vt) \quad t' = \gamma\left(t - \frac{V}{c^2}x\right), \quad (5.3)$$

donde  $\gamma = \left(\sqrt{1 - \frac{V^2}{c^2}}\right)^{-1}$ .

**Proposición 5.2.** Consideremos un sistema de referencia inercial, dos eventos  $(t_1, x_1, y_1, z_1), (t_2, x_2, y_2, z_2)$  del espacio tiempo y el intervalo

$$s^2 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2.$$

Entonces:

1. Existe un sistema de referencia inercial en el cual los eventos ocurren en el mismo punto en el espacio ( $x'_1 = x'_2, y'_1 = y'_2, z'_1 = z'_2$ ) si y sólo si  $s^2 > 0$ . Si  $t_2 > t_1$  entonces en cualquier sistema de referencia inercial ocurre que  $t'_2 > t'_1$ .
2. Existe un sistema de referencia inercial en el cual los eventos son simultáneos ( $t'_1 = t'_2$ ) si y sólo si  $s^2 < 0$ . Si  $t_2 > t_1$  también podemos encontrar un sistema de referencia inercial en el cual  $t'_2 > t'_1$ .

*Demostración.* Ya hemos visto que si los eventos ocurren en el mismo punto en el espacio o son simultáneos entonces  $s^2 > 0$  o  $s^2 < 0$ . Supongamos que  $s^2 > 0$ , primero vamos a encontrar un sistema de referencia inercial en el cual  $x'_1 = x'_2$ . De la transformación 5.3 obtenemos que

$$0 = x'_2 - x'_1 = \gamma(x_2 - x_1 - V(t_2 - t_1))$$

ya así  $V = \frac{x_2 - x_1}{t_2 - t_1}$ . Esta velocidad es menor a la velocidad de la luz pues como  $s^2 > 0$  tenemos que

$$c^2(t_2 - t_1)^2 > (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 > (x_2 - x_1)^2,$$

ya por lo tanto  $c^2 > \frac{(x_2 - x_1)^2}{(t_2 - t_1)^2} = V^2$ . Si  $t_2 > t_1$ , supongamos que en algún sistema de referencia inercial ocurre que  $t'_2 < t'_1$ , en este caso

$$0 < t'_1 - t'_2 = -\gamma(t_2 - t_1 - \frac{V}{c^2}(x_2 - x_1)),$$

ya entonces  $t_2 - t_1 < \frac{V}{c^2}(x_2 - x_1)$ , como  $\frac{V}{c} < 1$  tenemos que  $c(t_2 - t_1) < (x_2 - x_1)$  ya al elevar al cuadrado obtenemos una contradicción.

Si ahora  $s^2 < 0$  podemos hacer una rotación en el espacio tridimensional  $(x, y, z)$  de tal forma que la recta que une a los puntos  $(x_1, y_1, z_1)$  con  $(x_2, y_2, z_2)$  sea paralela al eje  $x$ . Después de este cambio los eventos tienen coordenadas  $(t_1, x_1, 0, 0)$  y  $(t_2, x_2, 0, 0)$  ya así  $c^2(t_2 - t_1)^2 < (x_2 - x_1)^2$ , usando nuevamente la transformación 5.3,

$$0 = t'_2 - t'_1 = \gamma(t_2 - t_1 - \frac{V_0}{c^2}(x_2 - x_1))$$

ya entonces  $V_0 = c^2 \frac{t_2 - t_1}{x_2 - x_1}$ . Como  $V^2 = c^4 \frac{(t_2 - t_1)^2}{(x_2 - x_1)^2} < c^2$ , la velocidad  $V$  si está permitida. Si  $t_2 > t_1$  ya queremos que

$$\gamma(t_2 - t_1 - \frac{V}{c^2}(x_2 - x_1)) = t'_2 - t'_1 < 0,$$

entonces debe ocurrir que  $V_0 = c^2 \frac{t_2 - t_1}{x_2 - x_1} < V < c$ , como  $V_0 < c$  podemos encontrar una velocidad que cumpla lo anterior.  $\square$

La proposición anterior nos dice que, para cada evento  $(t_1, x_1, y_1, z_1)$  la forma cuadrática divide al espacio tiempo en tres regiones distintas. Para todos los eventos  $(t, x, y, z)$  para los cuales  $s^2 > 0$ , el tiempo  $t$  tiene dos posibilidades  $t > t_1$  o  $t < t_1$ , a la primera región se le conoce como *el futuro* del evento  $(t_1, x_1, y_1, z_1)$  y a la segunda como *el pasado*. Tal como afirma la proposición estar en el futuro o el pasado de un cierto evento no depende del sistema de referencia inercial que estemos utilizando para dar coordenadas a los eventos del espacio tiempo. Por otro lado, para los eventos tales que  $s^2 < 0$ , la proposición nos dice que las nociones de futuro y pasado no tienen sentido. Esto va de acuerdo a que, como hemos postulado, nada puede viajar a una velocidad superior a la velocidad de la luz, precisamente para los eventos para los cuales  $s^2 < 0$  esta última condición se viola. Así, estos eventos jamás se verán afectados por una señal o campo de fuerzas que tenga origen en el evento  $(t_1, x_1, y_1, z_1)$ .

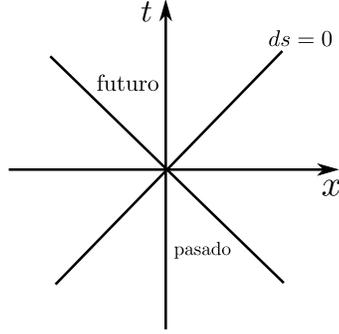


Figura 5.2: El futuro y pasado de un evento en el espacio tiempo.

Usando la transformación de Lorentz (5.2) podemos obtener la regla de correspondencia que relaciona las velocidades que dos sistemas de referencia inerciales asignan a una misma partícula, tenemos que:

$$\begin{aligned} dt &= \gamma(dt' + \frac{V}{c^2} dx') & dx &= \gamma(dx' + V dt') \\ dy &= dy' & dz &= dz', \end{aligned}$$

por lo tanto, si  $v_x = \frac{dx}{dt}$  y  $v'_x = \frac{dx'}{dt'}$

$$\begin{aligned} v_x &= \frac{dx' + V dt'}{dt' + \frac{V}{c^2} dx'} = \frac{dx' + V dt'}{dt' (1 + \frac{V}{c^2} v'_x)} \\ &= \frac{v'_x + V}{1 + \frac{V}{c^2} v'_x}. \end{aligned}$$

De manera análoga obtenemos:

$$v_y = \gamma^{-1} \left( \frac{v'_y}{1 + \frac{V}{c^2} v_x} \right)$$

$$v_z = \gamma^{-1} \left( \frac{v'_z}{1 + \frac{V}{c^2} v_x} \right).$$

Puesto que en el límite cuando  $c \rightarrow \infty$ ,  $\gamma \rightarrow 1$ , al tomar este límite recuperamos la regla para transformar las velocidades entre dos sistemas de referencia inerciales de la mecánica clásica.

$$v_x = V + v'_x \qquad v_y = v'_y \qquad v_z = v'_z.$$

## 5.4. Las ecuaciones de la Relatividad Especial

De acuerdo al principio de mínima acción, en la mecánica clásica una partícula libre sigue una trayectoria para la cual la integral  $\int \frac{m}{2} (dx^2 + dy^2 + dz^2)$  alcanza un mínimo. En efecto, una partícula libre se desplaza en una línea recta a velocidad constante y sabemos que de todas las trayectorias recorridas a velocidad constante que unen a dos puntos, aquella que minimiza la distancia euclidiana  $\sqrt{dx^2 + dy^2 + dz^2}$  es una línea recta. En analogía con esto, postulamos que en la relatividad especial una partícula libre sigue una trayectoria para la cual

$$\int \alpha ds,$$

con  $\alpha \in \mathbb{R}$  adquiere un mínimo. Tenemos que,

$$ds = \sqrt{c^2 dt^2 - (dx^2 + dy^2 + dz^2)} = c \sqrt{1 - \frac{v^2}{c^2}},$$

donde  $\mathbf{v}$  es la velocidad de la partícula libre en cuestión y  $v^2 = \mathbf{v} \cdot \mathbf{v} = v_x^2 + v_y^2 + v_z^2$  es su rapidez. Así, la función

$$L = \alpha c \sqrt{1 - \frac{v^2}{c^2}}$$

debe de ser el Lagrangiano de la partícula libre. Para determinar el valor de  $\alpha$ , en el límite cuando  $c \rightarrow \infty$ , debe ocurrir que este Lagrangiano se convierte en aquel de la partícula libre en la mecánica clásica. Cuando  $c \rightarrow \infty$  tenemos que  $\frac{v}{c} \rightarrow 0$  y expandiendo  $L$  en series de potencias de la variable  $\frac{v}{c}$  obtenemos,

$$L = \alpha c \sqrt{1 - \frac{v^2}{c^2}} \approx \alpha c - \frac{\alpha v^2}{2c}.$$

El término  $\alpha c$  no afecta a las ecuaciones de movimiento y así debe ocurrir que  $-\frac{\alpha v^2}{2c} = m \frac{v^2}{2}$  y entonces,  $\alpha = -mc$ . Por lo tanto el Lagrangiano de la partícula

libre en la relatividad especial es:

$$L = -mc^2 \sqrt{1 - \frac{v^2}{c^2}}.$$

De esta expresión obtenemos el momento relativista:

$$\frac{\partial L}{\partial v} = \mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

En el caso de la partícula libre tenemos que la energía  $H$  es igual a la energía cinética  $T$  y en este caso tenemos que

$$\begin{aligned} T = \mathbf{p} \cdot \mathbf{v} - L &= \frac{mv^2}{\sqrt{1 - \frac{v^2}{c^2}}} + mc^2 \sqrt{1 - \frac{v^2}{c^2}} \\ &= \left( \sqrt{1 - \frac{v^2}{c^2}} \right)^{-1} \left( mv^2 + mc^2 \left( 1 - \frac{v^2}{c^2} \right) \right) \\ &= \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}}. \end{aligned}$$

Para escribir la energía cinética en términos del momento  $\mathbf{p}$  calculemos

$$\begin{aligned} \frac{T^2}{c^2} - \mathbf{p} \cdot \mathbf{p} &= \frac{m^2 c^2 - m^2 v^2}{1 - \frac{v^2}{c^2}} \\ &= m^2 c^2, \end{aligned}$$

y por lo tanto  $\frac{T}{c^2} = \mathbf{p}^2 + m^2 c^2$  ( $\mathbf{p}^2 = \mathbf{p} \cdot \mathbf{p}$ ). Así

$$T = c\sqrt{\mathbf{p}^2 + m^2 c^2}. \quad (5.4)$$

La interacción de varias partículas puede describirse usando el concepto de campo de fuerzas. Cuando decimos que una partícula actúa en otra podemos pensar que la partícula crea un campo de fuerzas alrededor de ella y las demás partículas se ven afectadas por este campo de fuerzas. En la mecánica clásica un campo de fuerzas creado por una partícula afecta de manera inmediata a todas las demás partículas. En la relatividad, debido a la velocidad de propagación de las interacciones, un campo de fuerzas creado por una partícula afecta a las demás sólo después de que ha transcurrido un cierto tiempo y el campo de fuerzas alcanza a las demás partículas. Así, en la relatividad el campo de fuerzas adquiere un significado físico, una partícula interactúa con el campo de fuerzas y después de un cierto tiempo la segunda partícula interactúa con el campo de fuerzas.

Dos campos de fuerza muy conocidos son el electromagnético y el gravitatorio, este último es el objeto de estudio de la relatividad general. En este trabajo únicamente hablaremos del campo electromagnético. Experimentalmente se ha

comprobado que la interacción de una partícula con el campo electromagnético depende de un parámetro  $e$  llamada *la carga eléctrica*, este parámetro puede tomar cualquier valor real. La fuerza que una partícula con carga  $e$  siente debido a un campo electromagnético puede escribirse como

$$\mathbf{F} = e\mathbf{E} + \frac{e}{c}\mathbf{v} \times \mathbf{H}.$$

Al campo vectorial  $\mathbf{E}$  se le conoce como *el campo eléctrico* y a  $\mathbf{H}$  como *el campo magnético*. Cabe mencionar que la partícula a su vez afecta al campo electromagnético pero, si el valor absoluto de  $e$  es muy pequeño podemos despreciar estos cambios. La fuerza debida al campo electromagnético puede obtenerse mediante el Lagrangiano

$$L = -mc^2\sqrt{1 - \frac{v^2}{c^2}} + \frac{e}{c}\mathbf{A} \cdot \mathbf{v} - e\phi,$$

donde  $\mathbf{A}: \mathbb{R}^4 \rightarrow \mathbb{R}^3$  y  $\phi: \mathbb{R}^4 \rightarrow \mathbb{R}$  son funciones en el espacio tiempo conocidas como *el potencial vectorial* y *el potencial escalar* respectivamente. Con este Lagrangiano el momento generalizado es

$$\frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}} + \frac{e}{c}\mathbf{A} = \mathbf{p} + \frac{e}{c}\mathbf{A}.$$

Con esto la ecuación de Euler-Lagrange puede escribirse como

$$\frac{d}{dt} \left( \mathbf{p} + \frac{e}{c}\mathbf{A} \right) = \frac{e}{c}\nabla\mathbf{A} \cdot \mathbf{v} - e \text{grad} \phi + \frac{e}{c}\mathbf{v} \times \text{rot}\mathbf{A},$$

donde  $\nabla\mathbf{A} \cdot \mathbf{v}$  es la matriz derivada de  $\mathbf{A}$  respecto a  $x, y, z$  aplicada al vector  $\mathbf{v}$ . Usando que

$$\frac{d\mathbf{A}}{dt} = \frac{\partial\mathbf{A}}{\partial t} + \nabla\mathbf{A} \cdot \mathbf{v},$$

podemos escribir lo anterior como:

$$\begin{aligned} \frac{d\mathbf{p}}{dt} &= -e \left( \frac{1}{c} \frac{\partial\mathbf{A}}{\partial t} + \text{grad} \phi \right) + \frac{e}{c}\mathbf{v} \times \text{rot}\mathbf{A} \\ &= e\mathbf{E} + \frac{e}{c}\mathbf{v} \times \mathbf{H} \end{aligned} \quad (5.5)$$

Estas son las ecuaciones de movimiento de la relatividad especial. Tenemos que:

$$\begin{aligned} \mathbf{E} &= - \left( \frac{1}{c} \frac{\partial\mathbf{A}}{\partial t} + \text{grad} \phi \right) \\ \mathbf{H} &= \text{rot}\mathbf{A}. \end{aligned} \quad (5.6)$$

Usando que  $\frac{d}{dt} \left( (1 - \frac{v^2}{c^2})^{-\frac{1}{2}} \right) = \frac{1}{c^2} (1 - \frac{v^2}{c^2})^{-\frac{3}{2}} \mathbf{v} \cdot \dot{\mathbf{v}}$  obtenemos que

$$\frac{dT}{dt} = \frac{d}{dt} \left( mc^2 \left( 1 - \frac{v^2}{c^2} \right)^{-\frac{1}{2}} \right) = m \left( 1 - \frac{v^2}{c^2} \right)^{-\frac{3}{2}} \mathbf{v} \cdot \dot{\mathbf{v}},$$

por otro lado,

$$\begin{aligned}\frac{d\mathbf{p}}{dt} &= m\dot{\mathbf{v}} \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} + m\frac{\mathbf{v}}{c^2} \left(1 - \frac{v^2}{c^2}\right)^{-\frac{3}{2}} \mathbf{v} \cdot \dot{\mathbf{v}} \\ &= m\dot{\mathbf{v}} \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} + \frac{\mathbf{v}}{c^2} \frac{dT}{dt},\end{aligned}$$

y así,

$$\begin{aligned}\dot{\mathbf{p}} \cdot \mathbf{v} &= m \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} \mathbf{v} \cdot \dot{\mathbf{v}} + \frac{v^2}{c^2} \frac{dT}{dt} \\ &= \left(1 - \frac{v^2}{c^2}\right) \frac{dT}{dt} + \frac{v^2}{c^2} \frac{dT}{dt} = \frac{dT}{dt}.\end{aligned}$$

Sustituyendo la expresión (5.5) para  $\dot{\mathbf{p}}$  obtenemos

$$\frac{dT}{dt} = e\mathbf{E} \cdot \mathbf{v}. \quad (5.7)$$

**Ejemplo 20.** Consideremos un campo electromagnético con  $\mathbf{E} = 0$  y  $\mathbf{H}$  constante en el espacio y en el tiempo. A un campo como este se le conoce como *campo magnético constante y uniforme*. Si escogemos la dirección de dicho campo en dirección del eje  $z$  entonces  $\mathbf{H} = (0, 0, H)$  y las ecuaciones de movimiento (5.5) son:

$$\dot{p}_x = \frac{e}{c} H v_y \quad \dot{p}_y = -\frac{e}{c} H v_x \quad \dot{p}_z = 0.$$

Notemos que,

$$\frac{T}{c^2} \mathbf{v} = \left( \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \right) \frac{\mathbf{v}}{c^2} = \frac{m\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}} = \mathbf{p},$$

y de la igualdad (5.7) obtenemos que en este caso  $\frac{dT}{dt} = 0$  y en consecuencia  $\dot{\mathbf{v}} = \frac{c^2}{T} \dot{\mathbf{p}}$ . Así, tenemos las ecuaciones:

$$\dot{v}_x = \omega v_y \quad \dot{v}_y = -\omega v_x \quad \dot{v}_z = 0,$$

donde  $\omega = \frac{ce}{T} H$ . Por lo tanto en el plano  $(v_x, v_y)$  las soluciones a las ecuaciones anteriores son circunferencias y están dadas por

$$v_x = v_0 \cos(\omega t + \alpha) \quad v_y = -v_0 \sin(\omega t + \alpha) \quad v_z = v_{0z}.$$

Las condiciones iniciales son el radio de la circunferencia  $v_0$ , el ángulo inicial  $\alpha$  y la velocidad inicial en el eje  $z$ ,  $v_{0z}$ . Integrando obtenemos que la trayectoria que sigue la partícula está dada por:

$$x = x_0 + \frac{v_0}{\omega} \sin(\omega t + \alpha) \quad y = y_0 + \frac{v_0}{\omega} \cos(\omega t + \alpha) \quad z = z_0 + v_{0z} t.$$

Así, bajo el campo de fuerzas de un campo magnético constante y uniforme, la partícula describe un hélice con eje en dirección del campo magnético. Si la velocidad inicial en dirección del eje  $z$  es cero entonces la partícula se mueve en un círculo en el plano  $z = z_0$ .

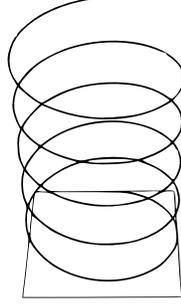


Figura 5.3: Trayectoria de una partícula bajo un campo magnético constante.

## 5.5. Ondas electromagnéticas

Como mencionamos anteriormente, en la teoría de la relatividad especial, el campo electromagnético adquiere una importancia fundamental para describir las interacciones entre partículas cargadas: un campo electromagnético se propaga en el espacio y sólo después de un determinado tiempo las partículas interactúan con el campo. Hasta ahora hemos supuesto que las cargas de las partículas no afectan al campo electromagnético y encontramos que

$$\mathbf{E} = -\frac{1}{c} \left( \frac{\partial \mathbf{A}}{\partial t} + \text{grad } \phi \right) \quad \mathbf{H} = \text{rot } \mathbf{A}.$$

De estas ecuaciones es fácil obtener ecuaciones que no involucren a los potenciales  $\mathbf{A}$  y  $\phi$ . En efecto, recordando que  $\text{rot grad} = 0$  y  $\text{div rot} = 0$  obtenemos

$$\begin{aligned} \text{rot } \mathbf{E} &= -\frac{1}{c} \frac{\partial}{\partial t} (\text{rot } \mathbf{A}) - \text{rot grad } \phi \\ &= -\frac{1}{c} \frac{\partial}{\partial t} \mathbf{H}, \end{aligned}$$

y

$$\text{div } \mathbf{H} = \text{div rot } \mathbf{A} = 0.$$

Por lo tanto,

$$\begin{aligned} \text{rot } \mathbf{E} &= -\frac{\partial}{\partial t} \mathbf{H} \\ \text{div } \mathbf{H} &= 0. \end{aligned} \tag{5.8}$$

Estas son las primeras dos *ecuaciones de Maxwell* y ellas no determinan completamente al campo electromagnético. Las otras dos ecuaciones de Maxwell son

$$\begin{aligned} \operatorname{rot} \mathbf{H} &= \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c} \mathbf{j} \\ \operatorname{div} \mathbf{E} &= 4\pi\rho, \end{aligned}$$

donde  $\rho$  es la *densidad de carga* y  $\mathbf{j}$  es el *vector de densidad de corriente*. Para ver como obtener estas dos ecuaciones consultar [5]. La integral de la densidad de carga  $\rho$  en algún volumen  $V$  nos dice cuál es la cantidad de carga contenida en ese volumen y el vector de densidad de corriente  $\mathbf{j}$  es una medida de cuanto carga fluye a través de un elemento de área. En este trabajo estudiaremos el caso en el que  $\rho = 0$  y  $\mathbf{j} = \bar{0}$ . Bajo estas condiciones las ecuaciones de Maxwell se vuelven

$$\begin{aligned} \operatorname{rot} \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} & \operatorname{div} \mathbf{H} &= 0 \\ \operatorname{rot} \mathbf{H} &= \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} & \operatorname{div} \mathbf{E} &= 0. \end{aligned} \quad (5.9)$$

A los campos electromagnéticos que satisfacen estas ecuaciones se les conoce como *ondas electromagnéticas*. A continuación veremos que las ondas electromagnéticas satisfacen la ecuación de onda. Para esto recordemos que el campo eléctrico y magnético están definidos por

$$\mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \operatorname{grad} \phi \quad \mathbf{H} = \operatorname{rot} \mathbf{A}.$$

donde  $\mathbf{A}$  y  $\phi$  son los potenciales vectorial y escalar. Un cálculo sencillo muestra que si reemplazamos a estos potenciales por

$$\mathbf{A}' = \mathbf{A} + \operatorname{grad} f \quad \phi' = \phi - \frac{1}{c} \frac{\partial f}{\partial t},$$

entonces estos nuevos potenciales determinan al mismo campo electromagnético que  $\mathbf{A}$  y  $\phi$ . En particular, resolviendo  $\phi = \frac{1}{c} \frac{\partial f}{\partial t}$  siempre podemos tomar el potencial escalar igual a cero. Así,  $\mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}$  y sustituyendo esto en la igualdad  $\operatorname{rot} \mathbf{H} = \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}$  de (5.9) obtenemos,

$$\operatorname{rot} \operatorname{rot} \mathbf{A} = -\Delta \mathbf{A} + \operatorname{grad} \operatorname{div} \mathbf{A} = -\frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2}.$$

En esta igualdad  $\Delta$  es el operador Laplaciano y significa que se lo aplicamos a cada coordenada del vector  $\mathbf{A}$ . Ahora veremos que podemos tomar al potencial vectorial  $\mathbf{A}$  de tal forma que  $\operatorname{div} \mathbf{A} = 0$ . De la igualdad  $\operatorname{div} \mathbf{E} = 0$  de (5.9) obtenemos

$$\frac{\partial}{\partial t} \operatorname{div} \mathbf{A} = \operatorname{div} \frac{\partial \mathbf{A}}{\partial t} = \operatorname{div} c\mathbf{E} = 0.$$

Así, la función  $\operatorname{div} \mathbf{A}$  no depende del tiempo. Una consecuencia del teorema de Helmholtz es que dada una función  $g: \mathbb{R}^3 \rightarrow \mathbb{R}$  y un campo vectorial  $\mathbf{v}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  con  $\operatorname{div} \mathbf{v} = 0$  que decaen al infinito más rápido que  $\frac{1}{r^2}$  (físicamente esta

última hipótesis siempre es razonable) entonces es posible encontrar un campo vectorial  $\mathbf{B}$  tal que  $\text{div } \mathbf{B} = g$  y  $\text{rot } \mathbf{B} = \mathbf{v}$ . Para darnos una idea de lo anterior, supongamos que tenemos un campo vectorial  $\mathbf{B}$  que satisface las dos condiciones anteriores y consideremos la ecuación de Poisson

$$-\Delta \mathbf{F} = \mathbf{B}.$$

Si las componentes de  $\mathbf{B}$  son acotadas y decaen al infinito más rápido que  $\frac{1}{r^2}$  entonces la ecuación anterior admite una solución. Si definimos  $U := \text{div } \mathbf{F}$  y  $\mathbf{W} := \text{rot } \mathbf{F}$  entonces

$$\text{rot } \mathbf{W} = \text{rot } \text{rot } \mathbf{F} = -\Delta \mathbf{F} + \text{grad } \text{div } \mathbf{F} = \mathbf{B} + \text{grad } U,$$

y por lo tanto,

$$\mathbf{B} = -\text{grad } U + \text{rot } \mathbf{W}.$$

Veamos las ecuaciones que satisfacen la función  $U$  y el campo vectorial  $\mathbf{W}$ . Tomando la divergencia,

$$g = \text{div } \mathbf{B} = -\text{div } \text{grad } U + \text{div } \text{rot } \mathbf{W} = -\Delta U.$$

Si ahora tomamos el rotacional,

$$\begin{aligned} \mathbf{v} = \text{rot } \mathbf{B} &= -\text{rot } \text{grad } U + \text{rot } \text{rot } \mathbf{W} \\ &= -\Delta \mathbf{W} + \text{grad } \text{div } \text{rot } \mathbf{F} \\ &= -\Delta \mathbf{W}. \end{aligned}$$

Así pues, dados una función  $g: \mathbb{R}^3 \rightarrow \mathbb{R}$  y un campo vectorial  $\mathbf{v}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , si resolvemos las ecuaciones  $-\Delta U = g$  y  $-\Delta \mathbf{W} = \mathbf{v}$ , entonces el campo vectorial  $\mathbf{B} = -\text{grad } U + \text{rot } \mathbf{W}$  tiene como divergencia a  $g$  y como rotacional a  $\mathbf{v}$ . Aplicando esto a  $g = \text{div } \mathbf{A}$  y  $\mathbf{v} = 0$  obtenemos un campo vectorial  $\mathbf{B}$  y como  $\mathbb{R}^3$  es simplemente conexo  $\mathbf{B} = \text{grad } f$  para alguna función  $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ . Así,  $\mathbf{A}' = \mathbf{A} - \text{grad } f$  tiene divergencia cero. Notemos que este cambio de potencial vectorial no afecta nuestra previa elección del potencial escalar  $\phi = 0$  pues  $f$  no depende del tiempo. Así pues tenemos que el potencial vectorial para las ondas electromagnéticas satisface

$$\Delta \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = 0$$

Como el orden en que tomamos las derivadas no importa, al tomar la parcial respecto al tiempo de esta igualdad obtenemos que el campo eléctrico  $\mathbf{E}$  también satisface esta ecuación. Tomando el rotacional vemos que el campo magnético también la satisface. Así, si  $f: \mathbb{R}^4 \rightarrow \mathbb{R}$  es una entrada del campo eléctrico o magnético entonces  $f$  satisface:

$$\Delta f - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = 0. \quad (5.10)$$

A esta ecuación se le conoce como *la ecuación de onda*. Si la onda en cuestión sólo depende de una dirección en el espacio entonces la onda se llama *onda plana* si tomamos el eje  $x$  como esta dirección, entonces la ecuación de onda se vuelve

$$\frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} - \frac{\partial^2 f}{\partial x^2} = 0.$$

No es difícil convencerse que las soluciones a esta ecuación son de la forma  $f(x, t) = f_1(t - \frac{x}{c}) + f_2(t + \frac{x}{c})$ . La función  $f_1$  es una onda propagándose con velocidad  $c$  en el sentido positivo del eje  $x$ . En efecto,  $f_1$  toma los mismos valores en rectas de la forma  $x = ct + a$ . Así, si a  $t = 0$  tenemos que  $f_1(0 - \frac{x}{c}) = \tilde{f}(x)$ , a tiempo  $t > 0$  la onda se ve igual que  $\tilde{f}(x)$  pero trasladada una distancia  $ct$  en el sentido positivo del eje  $x$ . Lo mismo ocurre con la función  $f_2(t + \frac{x}{c})$  pero ahora esta onda viaja en el sentido negativo del eje  $x$ .

Para simplificar nuestro estudio de las ondas electromagnéticas vamos a consi-

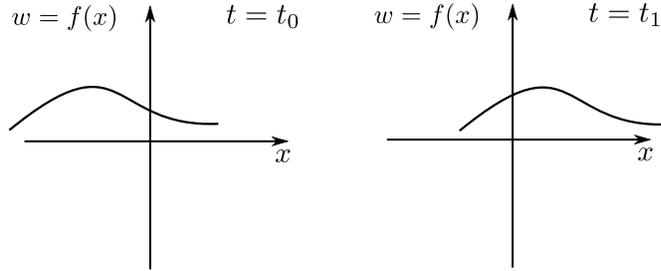


Figura 5.4: Una onda propagándose en la dirección positiva del eje  $x$ .

derar *ondas monocromáticas*, estas ondas están caracterizadas por ser periódicas en el tiempo. Así, una onda monocromática se puede escribir de la forma  $f(x, y, z, t) = \tilde{f}(x, y, z) \cos(\omega t + \alpha)$  y la ecuación de onda se vuelve

$$\Delta f + \frac{\omega}{c^2} f = 0.$$

Si además la onda es plana tenemos que,

$$\frac{\partial^2 f}{\partial x^2} + \frac{\omega}{c^2} f = 0.$$

Las soluciones a esta ecuación son de la forma  $a \cos(\omega(t + \frac{x}{c}) + \alpha)$ . En general podemos escribir a las componentes de una onda plana monocromática que se propaga en la dirección del vector unitario  $\mathbf{n} \in \mathbb{R}^3$  como  $a \cos(\omega(t - \frac{\mathbf{n} \cdot \mathbf{x}}{c}) + \alpha)$ , donde  $\mathbf{n} \cdot \mathbf{x}$  es el producto escalar usual de  $\mathbb{R}^3$ . Si definimos al vector  $\mathbf{k} = (\frac{\omega}{c^2}, \frac{\omega}{c} \mathbf{n}) \in \mathbb{R}^4$  y denotamos por  $\langle , \rangle$  al producto escalar definido por  $\langle (t, x, y, z), (t', x', y', z') \rangle$  podemos escribir a una onda plana monocromática como

$$a \cos(\langle \mathbf{k}, (t, x, y, z) \rangle + \alpha).$$

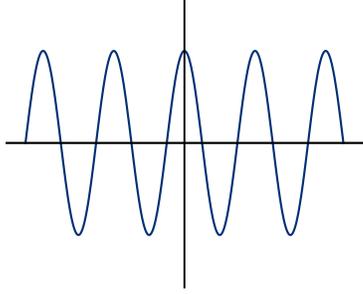


Figura 5.5: Onda monocromática.

Al argumento del coseno se le conoce como *la fase de la onda*. Notemos que el vector  $\mathbf{k}$  satisface  $\langle \mathbf{k}, \mathbf{k} \rangle = 0$ .

En general las ondas electromagnéticas no son planas pero, si la amplitud y la dirección de propagación permanecen prácticamente constantes en distancias del orden de la longitud de onda  $\lambda = \frac{2\pi c}{\omega}$  entonces, en una vecindad  $U$  de un punto  $\mathbf{x}_0 \in \mathbb{R}^3$  podemos introducir el concepto de *superficie de onda*. Esta superficie está caracterizada por todos los puntos que a un tiempo fijo  $t_0$  tienen la misma fase que el punto  $\mathbf{x}_0$  a tiempo  $t_0$ . Para una onda plana monocromática las superficies de onda son planos perpendiculares a la dirección de propagación dados por la ecuación  $\mathbf{n} \cdot \mathbf{x} = b + \omega t_0 + \alpha$  donde  $b$  es la fase de la onda a tiempo  $t_0$  en el punto  $\mathbf{x}_0$ . En la vecindad  $U$  podemos hablar de la dirección de propagación de la onda que es la dirección normal a la superficie de onda. A las curvas que son tangentes a la dirección de propagación de la onda se les conoce como *rayos*. Al estudio de la propagación de ondas mediante el uso de rayos se le conoce como *óptica geométrica*. Así, la óptica geométrica estudia la propagación de ondas cuya longitud de onda  $\lambda$  es muy pequeña.

Para una onda monocromática podemos escribir sus componentes como  $f =$

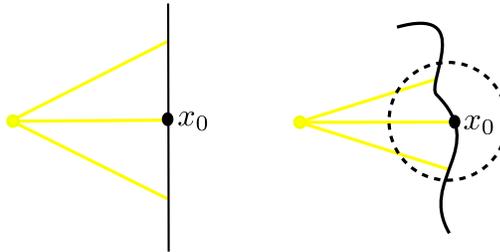


Figura 5.6: Superficies de onda de una onda monocromática plana y de otra no plana. Localmente podemos sustituir a la onda no plana por una que sí lo es.

$a \cos(\Psi)$  con  $\Psi: \mathbb{R}^4 \rightarrow \mathbb{R}$ . En una vecindad  $U$  del punto  $(t_0, \mathbf{x}_0)$  podemos hacer

la expansión  $\Psi = \Psi_0 + \frac{\partial \Psi}{\partial t} t + \text{grad } \Psi \cdot \mathbf{x}$ . En esta vecindad podemos sustituir a la superficie de onda por su plano tangente y suponer que la onda es plana. Comparando la expansión de  $\Psi$  con el vector  $\mathbf{k} = (\frac{\omega}{c^2}, \frac{\omega}{c} \mathbf{n})$  que da la fase de una onda monocromática plana, obtenemos:

$$\frac{\partial \Psi}{\partial t} = \omega \qquad -\text{grad } \Psi = \frac{\omega}{c} \mathbf{n},$$

Recordando que  $\langle \mathbf{k}, \mathbf{k} \rangle = 0$  obtenemos que  $\Psi$  satisface la ecuación:

$$\left(\frac{\partial \Psi}{\partial t}\right)^2 - \left(\frac{\partial \Psi}{\partial x}\right)^2 - \left(\frac{\partial \Psi}{\partial y}\right)^2 - \left(\frac{\partial \Psi}{\partial z}\right)^2 = 0. \quad (5.11)$$

Esta ecuación se le conoce como *la ecuación Eikonal* y sus curvas de nivel son las superficies de onda.

Al final del capítulo 3 estudiamos un modelo para la propagación de señales en una variedad Riemanniana  $(B, g)$ . Como vimos en esta sección, este modelo se corresponde con la propagación de ondas electromagnéticas cuya longitud de onda  $\lambda$  es muy pequeña.



## Capítulo 6

# Conclusiones

A lo largo del capítulo 4 estudiamos distintas maneras para obtener las ecuaciones de movimiento de una partícula en  $\mathbb{R}^3$ . En particular estudiamos sistemas de una partícula, restringida a una superficie o curva diferenciable de  $\mathbb{R}^3$ , en la cual actúan fuerzas que son derivables de un potencial o a lo más, las fuerzas no potenciales son proporcionales a la velocidad de la partícula. A lo largo de los años estos dos casos han sido ampliamente estudiados pero, el segundo de ellos, presentado como aquí, en términos de Hamiltonianos de contacto ha sido estudiado recientemente en [7]. Hay algunas preguntas que son de interés en este tema. Si a un sistema conservativo le añadimos fuerzas de fricción lineales, para obtener las ecuaciones (3.3) y (3.4) necesitamos fijar un sistema coordenado. Una vez hecho esto calculamos la 1-forma  $\alpha = y - \sum p dq$  y con ello, el campo vectorial  $X_H$  queda totalmente determinado por las ecuaciones (3.1). En coordenadas rectangulares vimos que las ecuaciones que definen al campo  $X_H$  coinciden con las ecuaciones de movimiento de una partícula bajo la influencia de una fuerza potencial y a fuerzas de fricción lineales. En los ejemplos 15 y 16 vimos que para el caso de una partícula moviéndose en un círculo o en  $\mathbb{R}^3$ , usar coordenadas angulares o coordenadas esféricas obtenemos las ecuaciones de movimiento correctas. Esto da la esperanza a que haya algún resultado que relacione las ecuaciones en distintos sistemas coordenados o más aún, que este modelo aparentemente dependiente de las coordenadas en realidad no dependa de ellas. En el caso de fuerzas conservativas,  $\frac{\partial H}{\partial y} = 0$  y la última ecuación (3.4) del campo vectorial  $X_H$  resulta ser  $\dot{y} = \sum p_i \frac{\partial H}{\partial p_i} - H = L$ . Así la última coordenada del flujo  $X_H$  es la acción  $S$  calculada a lo largo de las soluciones a las ecuaciones de movimiento. La acción  $S$  también es una solución a la ecuación de Hamilton-Jacobi. En el caso de contacto es interesante saber quién es la función  $S$  y si se puede interpretar como una especie de acción a través de la cual también podemos obtener las ecuaciones de movimiento. También sería interesante saber si se puede hacer depender al Hamiltoniano de un sistema conservativo de un parámetro adicional  $y$  para obtener las ecuaciones de otro tipo de fuerzas no potenciales.

Hay otro tipo de sistemas que se pueden estudiar mediante sistemas Hamilto-

nianos. Por ejemplo, las ecuaciones de movimiento de Relatividad Especial se pueden obtener mediante un Lagrangiano. Haciendo la transformada de Legendre de este Lagrangiano podemos escribir las ecuaciones de movimiento mediante un Hamiltoniano.

Las variedades de contacto también son útiles para obtener soluciones locales a ecuaciones diferenciales parciales de primer orden. Las ecuaciones que definen ambos tipos de sistemas Hamiltonianos (simpléctico o de contacto) son parte de las ecuaciones de las características de la ecuación de Hamilton-Jacobi. En el caso simpléctico el método de Hamilton-Jacobi nos permite obtener las soluciones a las ecuaciones de movimiento haciendo uso de las soluciones a la ecuación de Hamilton-Jacobi. En el caso de contacto, con una condición extra a la solución a la ecuación de Hamilton-Jacobi, el método de Hamilton-Jacobi también funciona.

En años recientes se ha descubierto que muchos otros tipos de sistemas físicos se pueden modelar mediante el uso de variedades de contacto [8]. En [10] se da una descomposición del campo electromagnético como suma de tres campos vectoriales dos de los cuales resultan ser los llamados campos de Beltrami. Estos campos vectoriales siempre inducen estructuras de contacto en el espacio en el que están definidos. La importancia de la descomposición anterior es que desacopla a las ecuaciones de Maxwell. Para entender la física de este procedimiento hace falta conocer el concepto matemático de la helicidad y el concepto físico de la polaridad. Así pues no siempre es del todo claro como cierto tipo de fenómenos físicos se pueden modelar mediante estructuras de contacto o simplécticas pero sin duda hay un gran número de ellos .

Es por eso que para entender a fondo cierto tipo de sistemas Hamiltonianos es bueno tener una buena idea de la física que hay detrás de ellos. Por otro lado, en los últimos años la geometría de contacto se ha desarrollado bastante (al grado de que ya hay topología de contacto) y esto da pie a que cada vez se pueda hacer un estudio más detallado de aquellos fenómenos que se pueden modelar mediante sistemas Hamiltonianos.

# Bibliografía

- [1] V.I. Arnol'd. *Mathematical Methods of Classical Mechanics*, Springer 1989.
- [2] V.I. Arnol'd. *Lectures on Partial Differential Equations*, Springer 2004.
- [3] H. Geiges *An Introduction to Contact Topology*, Cambridge University Press 2008.
- [4] L. D. Landau, E.M. Lifshitz *Mechanics*, Butterworth Heinemann 1976.
- [5] L. D. Landau, E.M. Lifshitz *The Classical Theory of Fields*, Butterworth Heinemann 1975.
- [6] J. M. Lee. *Introduction to Smooth Manifolds*, Springer 2003.
- [7] A. Bravetti, C. Hans, D. Tapias. *Contact Hamiltonian Dynamics*, Annals of Physics 376(2017) pp. 17-39.
- [8] A. Bravetti. *Contact Hamiltonian Dyanmics: The Concept and Its Use*, Entropy vol.19 2017-10 pp 535.
- [9] C. Boyer. *The history of calculus and its conceptual development*, Dover Publications 1959.
- [10] M Dahl. *Contact Geometry in Electromagnetism*, Progress In Electromagnetics Research, PIER 46, 2004, 46, 77-104.