

---

---

**Dep. Matemática Pura. FCUP**

---

---

**Geometria do Cálculo de Variações**

**Geometria Simpléctica<sup>1</sup>**

**Resumo das aulas teóricas**

**Mestrado em Matemática - Fundamentos e Aplicações**

**Ano lectivo de 2002/03**

**João Nuno Tavares**

---

<sup>1</sup>*simpléctico* do grego *symplektikós*, “que serve para ligar”.

---

---

---

---

# ÍNDICE:

<b>1</b>	<b>Elementos de cálculo de variações</b>	<b>1</b>
1.1	O problema clássico do cálculo de variações . . . . .	1
1.2	Transformada de Legendre. Equações canónicas . . . . .	7
1.3	Sistemas mecânicos conservativos . . . . .	11
1.4	A forma de Poincaré-Cartan . . . . .	16
1.5	Problema com extremidades móveis . . . . .	19
1.6	A função de acção $S$ . Equação de Hamilton-Jacobi $S_t + H(t, \mathbf{x}, S_{\mathbf{x}}) = 0$ . . . . .	23
1.7	Um princípio variacional para sistemas Hamiltonianos. Princípio de Maupertuis . . . . .	31
1.8	Os princípios variacionais de Jacobi e de Fermat. Analogia óptico-mecânica . . . . .	36
1.9	Feixes de extremais. A iconal . . . . .	38
1.10	O integral invariante de Hilbert . . . . .	45
1.11	Transformações canónicas. Método de Hamilton . . . . .	47
1.12	Método de Jacobi para integrar as equações canónicas de Hamilton. Teorema de Jacobi . . . . .	52
1.13	Invariantes integrais . . . . .	58
1.13.1	Preliminares de álgebra linear . . . . .	58
1.13.2	Subvariedades integrais. Teorema de Darboux . . . . .	59
1.13.3	Invariantes integrais . . . . .	63
<b>2</b>	<b>Problemas variacionais paramétricos</b>	<b>67</b>
2.1	Lagrangeanos paramétricos homogéneos . . . . .	67
2.2	Formalismo canónico . . . . .	71
2.3	Campos de Meyer . . . . .	79
2.4	Indicatriz. Função excesso. Fórmula de Weierstrass . . . . .	87
2.5	Discussão geométrica da equação de Hamilton-Jacobi reduzida . . . . .	90

---

2.6	Aplicações à óptica geométrica . . . . .	98
2.6.1	Construção das frentes de onda a partir dos raios . . . . .	98
2.6.2	Propagação das frentes de onda através de uma descontinuidade do meio. Lei de Snell . . . . .	101
2.6.3	Princípio de Huygens . . . . .	102
2.7	Apêndice. Equações de Maxwell e óptica geométrica . . . . .	104
2.7.1	Propagação da luz num meio isotrópico não homogêneo . . . . .	104
2.7.2	Representação integral das equações de Maxwell . . . . .	106
2.7.3	Propagação das descontinuidades. Frentes de onda . . . . .	108
2.7.4	A equação iconal da óptica geométrica . . . . .	112
<b>3</b>	<b>Geometria Simpléctica e Mecânica</b>	<b>115</b>
3.1	Variedades simplécticas . . . . .	116
3.2	Sistemas mecânicos com simetria. Aplicação momento. Redução . . . . .	124
3.2.1	O Problema de Kepler . . . . .	126
3.2.2	Movimento livre de um sólido com um ponto fixo . . . . .	127

# Capítulo 1

## Elementos de cálculo de variações

### 1.1 O problema clássico do cálculo de variações

Começemos por discutir o seguinte problema clássico do cálculo de variações:

- ♣ **Problema 1.1** ... *Entre as curvas  $\mathbf{x}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ , que satisfazem as condições de fronteira:*

$$\mathbf{x}(t_0) = \mathbf{x}_0, \quad \mathbf{x}(t_1) = \mathbf{x}_1 \quad (1.1.1)$$

onde  $\mathbf{x}_0, \mathbf{x}_1$  são dois pontos fixos em  $\mathbb{R}^n$ , calcular a curva para a qual o valor do funcional:

$$I[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} L(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \quad (1.1.2)$$

é mínimo <sup>1</sup>.

♣.

Figure 1.1:

---

<sup>1</sup>Mais geralmente,  $\mathbb{R}^n$  pode ser substituído por uma variedade suave  $M$  de dimensão  $n$ .

Em (1.1.2) a função  $L : \mathbb{R} \times T\mathbb{R}^n \cong \mathbb{R}^{2n+1} \rightarrow \mathbb{R}$ , chamada **Lagrangeano**, supõe-se de classe  $C^2$ , e o mínimo é entendido como **mínimo fraco**, no sentido seguinte - o funcional  $I$  atinge um mínimo fraco numa curva  $\widehat{\mathbf{x}}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$  se existir  $\delta > 0$  tal que, para toda a curva  $\mathbf{x}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$  que satisfaz as condições de fronteira (2.1.4) e a condição  $\|\mathbf{x}(\cdot) - \widehat{\mathbf{x}}(\cdot)\|_{C^1} < \delta$ , se tem  $I[\mathbf{x}(\cdot)] \geq I[\widehat{\mathbf{x}}(\cdot)]$ .

Outra possibilidade consiste em alargar a classe das funções admissíveis à classe  $SC^1([t_0, t_1], \mathbb{R}^n)$  das funções contínuas, de classe  $C^1$  por pedaços, munida da norma  $C^o$ . Neste caso, o mínimo é entendido como **mínimo forte**, no sentido seguinte - o funcional  $I$  atinge um mínimo forte numa curva  $\widehat{\mathbf{x}}(\cdot) \in SC^1([t_0, t_1], \mathbb{R}^n)$  se existir  $\delta > 0$  tal que, para toda a curva  $\mathbf{x}(\cdot) \in SC^1([t_0, t_1], \mathbb{R}^n)$  que satisfaz as condições de fronteira (2.1.4) e a condição  $\|\mathbf{x}(\cdot) - \widehat{\mathbf{x}}(\cdot)\|_{C^o} < \delta$ , se tem  $I[\mathbf{x}(\cdot)] \geq I[\widehat{\mathbf{x}}(\cdot)]$ .

É claro que um mínimo forte é necessariamente um mínimo fraco.

Vamos agora supôr que  $\widehat{\mathbf{x}}(\cdot)$  é uma solução do Problema 1.1, e vejamos uma condição necessária para que a curva  $\widehat{\mathbf{x}}(\cdot)$  seja mínimo fraco. Esta será também uma condição necessária para que essa mesma curva seja mínimo forte. Para isso, consideremos uma família a 1-parâmetro  $\lambda$  de curvas em  $C^1([t_0, t_1], \mathbb{R}^n)$ :

$$\lambda \longmapsto \mathbf{x}(\cdot; \lambda), \quad \lambda \in \mathbb{R} \quad (1.1.3)$$

que dependa diferenciavelmente do parâmetro  $\lambda$  e tal que:

$$\begin{aligned} \mathbf{x}(\cdot; \lambda = 0) &= \widehat{\mathbf{x}}(\cdot) \\ \mathbf{x}(t_0; \lambda) &= \mathbf{x}_0 \quad \text{e} \quad \mathbf{x}(t_1; \lambda) = \mathbf{x}_1, \quad \forall \lambda \\ \delta \widehat{\mathbf{x}}(\cdot) = \boldsymbol{\eta}(\cdot) &\stackrel{\text{def}}{=} \left. \frac{d}{d\lambda} \right|_{\lambda=0} \mathbf{x}(\cdot; \lambda) \end{aligned} \quad (1.1.4)$$

onde  $\boldsymbol{\eta}(\cdot) = \delta \widehat{\mathbf{x}}(\cdot)$  é uma variação com extremidades fixas, isto é, uma função em  $C^1([t_0, t_1], \mathbb{R}^n)$ , tal que  $\boldsymbol{\eta}(t_0) = 0 = \boldsymbol{\eta}(t_1)$ . Podemos por exemplo tomar a família a 1-parâmetro  $\lambda \mapsto \widehat{\mathbf{x}}(\cdot) + \lambda \boldsymbol{\eta}(\cdot)$ .

Então a função real de variável real:

$$\begin{aligned} \phi(\lambda) &= I[\mathbf{x}(\cdot; \lambda)] \\ &= \int_{t_0}^{t_1} L(t, \mathbf{x}(\cdot; \lambda), \dot{\mathbf{x}}(t; \lambda)) dt \end{aligned} \quad (1.1.5)$$

é uma função de classe  $C^2$ , que atinge um mínimo local em  $\lambda = 0$ . Portanto  $\phi'(0) = 0$  e  $\phi''(0) \geq 0$ .

Sejam  $x^i$  coordenadas locais em  $\mathbb{R}^n$  e  $(x^i, \dot{x}^i)$  as correspondentes coordenadas para  $T\mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n$ . Suponhamos que  $\mathbf{x}(\cdot) = x^i(\cdot)$  e  $\boldsymbol{\eta}(\cdot) = \eta^i(\cdot)$ , e calculemos  $\phi'(0)$  usando a regra da cadeia e a derivação sob o sinal integral. Usando sistematicamente a convenção de Einstein, vem que:

$$\phi'(0) = \int_{t_0}^{t_1} \left[ L_{x^i} \left( t, \widehat{\mathbf{x}}(t), \dot{\widehat{\mathbf{x}}}(t) \right) \eta^i(t) + L_{\dot{x}^i} \left( t, \widehat{\mathbf{x}}(t), \dot{\widehat{\mathbf{x}}}(t) \right) \dot{\eta}^i(t) \right] dt \quad (1.1.6)$$

onde pusemos  $L_{x^i} = \frac{\partial L}{\partial x^i}$  e  $L_{\dot{x}^i} = \frac{\partial L}{\partial \dot{x}^i}$ . Podemos ainda escrever o integral (1.1.6) na seguinte forma vectorial simplificada:

$$\int_{t_0}^{t_1} [L_{\mathbf{x}}(t) \boldsymbol{\eta}(t) + L_{\dot{\mathbf{x}}}(t) \dot{\boldsymbol{\eta}}(t)] dt \quad (1.1.7)$$

A este integral vamos aplicar o Lema seguinte:

- **♣ Lema 1.1 (Du Bois-Reymond)** ... *Sejam  $\mathbf{f}, \mathbf{g} \in C^0([t_0, t_1], \mathbb{R}^n)$  duas funções contínuas tais que:*

$$\int_{t_0}^{t_1} [\mathbf{f}(t) \boldsymbol{\eta}(t) + \mathbf{g}(t) \dot{\boldsymbol{\eta}}(t)] dt = 0 \quad (1.1.8)$$

para toda a função  $\boldsymbol{\eta} \in C^1([t_0, t_1], \mathbb{R}^n)$  que satisfaz  $\boldsymbol{\eta}(t_0) = \boldsymbol{\eta}(t_1) = 0$ . Então a função  $\mathbf{g}$  é de classe  $C^1$  e:

$$-\frac{d}{dt} \mathbf{g}(t) + \mathbf{f}(t) = 0 \quad (1.1.9)$$

**Dem.:** Seja  $\mathbf{F}(t) = \int_{t_0}^t \mathbf{f}(\tau) d\tau + \mathbf{k}$  uma primitiva da função  $\mathbf{f}$ , e calculemos o integral (1.1.8) por partes. Vem que:

$$\int_{t_0}^{t_1} [-\mathbf{F}(t) + \mathbf{g}(t)] \dot{\boldsymbol{\eta}}(t) dt = 0 \quad (1.1.10)$$

Consideremos agora a função:

$$\boldsymbol{\eta}(t) = \int_{t_0}^t [-\mathbf{F}(\tau) + \mathbf{g}(\tau)] d\tau$$

Temos então que  $\boldsymbol{\eta}(t_0) = 0$ , e, escolhendo convenientemente a constante  $\mathbf{k}$ , garantimos também a condição  $\boldsymbol{\eta}(t_1) = 0$ . Substituindo esta função  $\boldsymbol{\eta}$  no integral (1.1.8), obtemos:

$$\int_{t_0}^{t_1} [-\mathbf{F}(t) + \mathbf{g}(t)]^2 dt = 0$$

donde se deduz que  $\mathbf{F}(t) = \mathbf{g}(t)$  e portanto  $\mathbf{g}$  é de classe  $C^1$  e é válida a equação (1.1.9).

♣

Aplicando este lema a (1.1.6), com  $\mathbf{f}(t) = L_{\mathbf{x}}(t, \hat{\mathbf{x}}(t), \dot{\hat{\mathbf{x}}}(t))$  e  $\mathbf{g}(t) = L_{\dot{\mathbf{x}}} (t, \hat{\mathbf{x}}(t), \dot{\hat{\mathbf{x}}}(t))$ , concluímos que  $\hat{\mathbf{x}}(\cdot)$  deve satisfazer a chamada **equação de Euler-Lagrange**:

$$-\frac{d}{dt} L_{\dot{\mathbf{x}}} (t, \hat{\mathbf{x}}(t), \dot{\hat{\mathbf{x}}}(t)) + L_{\mathbf{x}} (t, \hat{\mathbf{x}}(t), \dot{\hat{\mathbf{x}}}(t)) = 0 \quad (1.1.11)$$

que é um sistema de  $n$  ODE's de segunda ordem, que escrevemos na forma simplificada seguinte:

$$-\frac{d}{dt} L_{\dot{x}^i} + L_{x^i} = 0, \quad i = 1, \dots, n \quad (1.1.12)$$

ou ainda, em forma vectorial:

$$\boxed{-\frac{d}{dt}L_{\dot{\mathbf{x}}} + L_{\mathbf{x}} = \mathbf{0}} \quad (1.1.13)$$

A solução geral deste sistema depende pois de  $2n$  parâmetros que devem ser escolhidos para que as condições de fronteira (2.1.4) sejam verificadas pela solução (note que (2.1.4) são  $2n$  condições de fronteira). No entanto, nada se afirma acerca da existência de solução que, de facto, pode não existir.

Qualquer solução das equações de Euler-Lagrange diz-se uma **extremal** do problema variacional 1.1.

### Notas ...

1. Quando o Lagrangeano  $L$  não depende explicitamente de  $t$ , isto é,  $L = L(\mathbf{x}, \dot{\mathbf{x}})$ , a chamada **energia** de  $L$ :

$$E_L(\mathbf{x}, \dot{\mathbf{x}}) \stackrel{\text{def}}{=} L_{\dot{\mathbf{x}}} \dot{\mathbf{x}} - L(\mathbf{x}, \dot{\mathbf{x}}) \quad (1.1.14)$$

ou mais detalhadamente,  $E_L(x^i, \dot{x}^i) = L_{\dot{x}^i} \dot{x}^i - L(x^i, \dot{x}^i)$ , é um integral primeiro da equação de Euler-Lagrange. De facto, se  $\mathbf{x}(t)$  é solução da equação (1.1.11), então:

$$\begin{aligned} \frac{d}{dt}E_L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) &= \frac{d}{dt}(L_{\dot{\mathbf{x}}} \dot{\mathbf{x}}) - \frac{d}{dt}L(\mathbf{x}, \dot{\mathbf{x}}) \\ &= \left(\frac{d}{dt}L_{\dot{\mathbf{x}}}\right) \dot{\mathbf{x}} + L_{\dot{\mathbf{x}}} \ddot{\mathbf{x}} - L_{\mathbf{x}} \dot{\mathbf{x}} - L_{\dot{\mathbf{x}}} \ddot{\mathbf{x}} \\ &= L_{\mathbf{x}} \dot{\mathbf{x}} + L_{\dot{\mathbf{x}}} \ddot{\mathbf{x}} - L_{\mathbf{x}} \dot{\mathbf{x}} - L_{\dot{\mathbf{x}}} \ddot{\mathbf{x}} \\ &= 0 \end{aligned} \quad (1.1.15)$$

2. Quando o Lagrangeano  $L$  não depende explicitamente da variável  $x^i$ , para um certo  $i \in \{1, \dots, n\}$ , o chamado **momento conjugado a  $x^i$** :

$$p_i(t, \dot{\mathbf{x}}) \stackrel{\text{def}}{=} L_{x^i}(t, \dot{\mathbf{x}}) \quad (1.1.16)$$

é um integral primeiro da equação de Euler-Lagrange. De facto,  $L_{x^i} = 0$  e a  $i$ -ésima equação (1.1.11) fica apenas  $-\frac{d}{dt}L_{x^i} = 0$ , isto é  $p_i(t, \dot{\mathbf{x}}(t)) \equiv c$ . Diz-se neste caso que a variável  $x^i$  é **cíclica**.



♣ **Exemplo 1.1 (A braquistócrona)** (Johann Bernoulli, 1696) ... É a curva que une dois pontos  $P_1$  e  $P_2$  num plano vertical, de tal modo que um ponto material de massa  $m$ , deslizando sem atrito sobre essa curva, sujeito apenas à gravidade, a percorre num tempo mínimo (do grego *brakhystós*: “o mais curto” + *khronos*: “tempo”).

Suponhamos que o plano vertical é o plano  $xy$ ,  $P_1 = (0, 0)$ ,  $P_2 = (a, b)$ , com  $a > 0$  e  $b > 0$  e que  $y = \varphi(x)$  é a equação da curva braquistócrona (o eixo dos  $y$ 's orienta-se para baixo).

Figure 1.2: Braquistócrona

A velocidade do ponto material, deslizando sem atrito sobre essa curva, é:

$$v = \frac{ds}{dt} = \sqrt{1 + [y'(x)]^2} \frac{dx}{dt}$$

e como, por outro lado,  $v$  é também determinada pela equação de conservação de energia:

$$\frac{1}{2}mv^2 = mgy \quad \Rightarrow \quad v = \sqrt{2gy}$$

deduzimos que o tempo  $T$  de descida de  $P_1 = (0, 0)$  até  $P_2 = (a, b)$  é dado por:

$$\begin{aligned} T[y(x)] &= \int_0^a \frac{ds}{v} \\ &= \int_0^a \frac{\sqrt{1 + [y'(x)]^2} dx}{\sqrt{2gy}} \\ &= \frac{1}{(2g)^{1/2}} \int_0^a \left[ \frac{1 + y'^2}{y} \right]^{1/2} dx \end{aligned} \quad (1.1.17)$$

com condições de fronteira  $y(0) = 0$  e  $y(a) = b$ .

Como o Lagrangeano  $L$  não depende do parâmetro  $x$ , há conservação da energia  $E_L = L_{y'}y' - L$ :

$$E_L = \frac{y'^2}{\sqrt{y(1 + y'^2)}} - \frac{\sqrt{1 + y'^2}}{\sqrt{y}} \equiv c$$

Simplificando, vem que:

$$\frac{1}{\sqrt{y(1 + y'^2)}} = C \quad \Rightarrow \quad y(1 + y'^2) = C_1$$

Introduzindo um parâmetro  $t$  e pondo  $y' = \cotg t$ , vem que:

$$y = \frac{C_1}{1 + \cotg^2 t} = C_1 \sin^2 t = \frac{C_1}{2}(1 - \cos 2t)$$

Por outro lado:

$$\begin{aligned} dx &= \frac{dy}{y'} = \frac{2C_1 \sin t \cos t dt}{\cotg t} = 2C_1 \sin^2 t dt = C_1(1 - \cos 2t)dt \\ \Rightarrow x &= C_1 \left( t - \frac{\sin 2t}{2} \right) + C_2 = \frac{C_1}{2} (2t - \sin 2t) + C_2 \end{aligned} \quad (1.1.18)$$

e a forma paramétrica da solução é:

$$\begin{cases} x - C_2 &= \frac{C_1}{2} (2t - \sin 2t) \\ y &= \frac{C_1}{2} (1 - \cos 2t) \end{cases}$$

Pondo  $\tau = 2t$  e como  $C_2 = 0$ , já que  $y(0) = 0$ , obtem-se a família de ciclóides:

$$\begin{cases} x(\tau) &= \frac{C}{2} (\tau - \sin \tau) \\ y(\tau) &= \frac{C}{2} (1 - \cos \tau) \end{cases} \quad (1.1.19)$$

onde  $C$  se calcula impondo a condição  $y(a) = b$ .

♣ **Exemplo 1.2** ... Calcular as extremais de:

$$I[x(t)] = \int_1^2 (\dot{x}^2 - 2tx) dt, \quad x(1) = 0, \quad x(2) = -1$$

Como  $L(t, x, \dot{x}) = \dot{x}^2 - 2tx$ , a equação de Euler-Lagrange é:

$$-\frac{d}{dt}L_{\dot{x}} + L_x = -2\ddot{x} - 2t = 0$$

cuja solução geral é:

$$x(t) = -\frac{t^3}{6} + at + b$$

Utilizando as condições de fronteira, calculamos  $a = 1/6$  e  $b = 0$ . A extremal é pois  $x(t) = \frac{t}{6}(1 - t^2)$ .

♣ **Exemplo 1.3** ... Calcular as extremais de:

$$I[y(x)] = \int_1^2 (y'(1 + x^2y')) dx, \quad y(1) = 3, \quad y(2) = 5$$

Como  $L(x, y, y') = y'(1 + x^2y')$ , não depende de  $y$ , o momento  $p = L_{y'}(x, y') = 1 + 2x^2y'$  é constante. A equação de Euler-Lagrange é:

$$\frac{d}{dx}L_{y'} = 1 + 2x^2y' = 0$$

cuja solução geral é:

$$1 + 2x^2y' \equiv c$$

Portanto  $y' = \frac{c-1}{2x^2}$ , i.e.,  $y = \frac{a}{x} + b$ , onde  $a = \frac{1-c}{2}$ . As extremais são pois uma família de hipérbolas. Utilizando as condições de fronteira, calculamos  $a = -4$  e  $b = 7$ . A extremal é pois  $y(x) = 7 - \frac{4}{x}$ .

♣ **Exemplo 1.4** ... Calcular as extremais de:

$$I[y(x)] = \int_a^b \frac{\sqrt{1+(y')^2}}{y} dx, \quad y(a) = A, \quad y(b) = B$$

onde os pontos  $(a, A), (b, B)$  pertencem ao semiplano superior  $\mathbf{H}^+ = \{(x, y) : y > 0\}$ .

Como  $L(x, y, y') = \frac{\sqrt{1+(y')^2}}{y}$ , não depende do parâmetro  $x$ , a energia total:

$$E_L(y, y') = y' L_{y'}(y, y') - L(y, y') = y' \frac{y'}{y\sqrt{1+(y')^2}} - \frac{\sqrt{1+(y')^2}}{y}$$

é constante. Depois de simplificar, obtemos:

$$y\sqrt{1+(y')^2} \equiv r > 0$$

Pondo  $y' = \operatorname{tg} t$ , vem que:

$$y^2 = \frac{r^2}{1+y'^2} = \frac{r^2}{1+\operatorname{tg}^2 t} = r^2 \cos^2 t$$

ou  $y = r \cos t$ . Por outro lado:

$$\frac{dy}{dx} = y' \Rightarrow dx = \frac{dy}{y'} = \frac{-r \sin t dt}{\operatorname{tg} t} = -r \cos t dt$$

ou  $x = -r \sin t + C$ . A solução é pois, em forma paramétrica:

$$\begin{cases} x - C &= -r \sin t \\ y &= r \cos t \end{cases}$$

e, eliminando  $t$ , obtemos uma família de circunferências centradas no eixo dos  $x'$ s:  $(x - C)^2 + y^2 = r^2$ . A extremal pedida será a que passa pelos dois pontos dados, e é única.

♣.

## 1.2 Transformada de Legendre. Equações canônicas

A transformada de Legendre de uma função  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  é, grosso modo, a equação da família de hiperplanos tangentes ao gráfico de  $F$ . Por exemplo, para  $n = 2$ ,  $F$  é uma função de 2 variáveis e o seu gráfico é a superfície de  $\mathbb{R}^3$ :

$$\operatorname{gr} F = \{(x^1, x^2, z) : z = F(x^1, x^2)\}$$

A superfície  $\operatorname{gr} F$ , em  $\mathbb{R}_{x^1, x^2, z}^3$ , pode ser descrita por dois processos duais - ou como o conjunto de pontos determinado pela equação  $z = F(x^1, x^2)$ , ou como a envolvente dos seus planos tangentes. Vejamos qual a equação a que deve satisfazer um plano afim em

$\mathbb{R}^3$  para que seja tangente a  $\text{gr } F$ . A equação de um plano afim não vertical em  $\mathbb{R}^3$ , pode ser sempre escrita na forma:

$$Z - p_1 X^1 - p_2 X^2 + u = 0$$

onde  $(X^1, X^2, Z)$  são as coordenadas correntes de um ponto desse plano. Neste caso, chamamos a  $(p_1, p_2, u)$  as coordenadas desse plano, que é pois o plano perpendicular ao vector  $(p_1, p_2, -1)$  e que intersecta o eixo dos  $zz$  no ponto  $(0, 0, -u)$ .

Como o plano tangente a  $\text{gr } F$ , no ponto  $(x^1, x^2, z = F(x^1, x^2)) \in \text{gr } F$ , é o plano de equação  $[(X^1, X^2, Z) - (x^1, x^2, F(\mathbf{x}))] \cdot (\frac{\partial F}{\partial x^1}(\mathbf{x}), \frac{\partial F}{\partial x^2}(\mathbf{x}), -1) = 0$ , onde  $\mathbf{x} = (x^1, x^2)$ , isto é:

$$Z - F(\mathbf{x}) - \frac{\partial F}{\partial x^1}(\mathbf{x})(X^1 - x^1) - \frac{\partial F}{\partial x^2}(\mathbf{x})(X^2 - x^2) = 0, \quad \mathbf{x} = (x^1, x^2)$$

as coordenadas desse plano são portanto:

$$\begin{aligned} p_1 &= \frac{\partial F}{\partial x^1}(x^1, x^2) \\ p_2 &= \frac{\partial F}{\partial x^2}(x^1, x^2) \\ u &= x^1 \frac{\partial F}{\partial x^1}(x^1, x^2) + x^2 \frac{\partial F}{\partial x^2}(x^1, x^2) - F(x^1, x^2) \end{aligned} \quad (1.2.1)$$

que se dizem as **coordenadas tangenciais** da superfície  $\text{gr } F$ . A superfície fica também determinada se conhecermos  $u$  como função de  $p_1$  e  $p_2$ , isto é, se conhecermos a família a dois parâmetros de planos tangentes ao  $\text{gr } F$ . Esta relação  $u = \Phi(p_1, p_2)$ , que se diz a **equação tangencial** do  $\text{gr } F$ , pode ser deduzida a partir de  $z = F(x^1, x^2)$ , calculando (se possível) os valores de  $x^1$  e  $x^2$ , como função de  $p_1$  e  $p_2$ , a partir das equações:

$$p_1 = \frac{\partial F}{\partial x^1}(x^1, x^2), \quad p_2 = \frac{\partial F}{\partial x^2}(x^1, x^2)$$

e substituindo esses valores  $x^1(p_1, p_2)$  e  $x^2(p_1, p_2)$  em  $u$ , dado por (1.2.1):

$$\begin{aligned} u &= \Phi(p_1, p_2) \\ &= x^1 \frac{\partial F}{\partial x^1}(x^1, x^2) + x^2 \frac{\partial F}{\partial x^2}(x^1, x^2) - F(x^1, x^2) \\ &= p_1 x^1(p_1, p_2) + p_2 x^2(p_1, p_2) - F(x^1(p_1, p_2), x^2(p_1, p_2)) \end{aligned} \quad (1.2.2)$$

A esta função  $\Phi(p_1, p_2)$  chamamos a **transformada de Legendre** da função  $F(x^1, x^2)$ .

Reciprocamente, para determinar as coordenadas pontuais a partir das coordenadas tangenciais, calculamos as derivadas parciais de  $\Phi(p_1, p_2)$ , dada por (1.2.2). Como  $p_1 = \frac{\partial F}{\partial x^1}(x^1, x^2)$  e  $p_2 = \frac{\partial F}{\partial x^2}(x^1, x^2)$ , obtemos:

$$\frac{\partial \Phi}{\partial p_1} = x^1 + p_1 \frac{\partial x^1}{\partial p_1} + p_2 \frac{\partial x^2}{\partial p_1} - \frac{\partial F}{\partial x^1} \frac{\partial x^1}{\partial p_1} - \frac{\partial F}{\partial x^2} \frac{\partial x^2}{\partial p_1} = x^1$$

e análogamente:

$$\frac{\partial \Phi}{\partial p_2} = x^2$$

Concluindo - obtemos o seguinte conjunto de fórmulas:

$$\begin{aligned} \Phi(p_1, p_2) + F(x^1, x^2) &= p_1 x^1 + p_2 x^2 \\ p_1 &= \frac{\partial F}{\partial x^1} & x^1 &= \frac{\partial \Phi}{\partial p_1} \\ p_2 &= \frac{\partial F}{\partial x^2} & x^2 &= \frac{\partial \Phi}{\partial p_2} \end{aligned} \quad (1.2.3)$$

que ilustra o carácter dual da passagem entre coordenadas pontuais e coordenadas tangenciais.

A transformada de Legendre de uma função  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  pode ser sempre calculada se as duas equações  $p_1 = \frac{\partial F}{\partial x^1}, p_2 = \frac{\partial F}{\partial x^2}$  puderem ser resolvidas em ordem a  $x^1, x^2$ , o que é possível se:

$$\frac{\partial^2 F}{\partial (x^1)^2} \frac{\partial^2 F}{\partial (x^2)^2} - \left( \frac{\partial^2 F}{\partial x^1 \partial x^2} \right)^2 \neq 0$$

A generalização para funções  $F : \mathbb{R}_x^n \rightarrow \mathbb{R}$  é óbvia - a transformada de Legendre de  $F$  é a função  $\Phi : \mathbb{R}_p^n \rightarrow \mathbb{R}$ , definida da seguinte forma. Primeiro definimos os  $\mathbf{p} = (p_i)$  através de:

$$\mathbf{p} = \mathbf{p}(\mathbf{x}) = \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}), \quad \text{isto é} \quad p_i = \frac{\partial F}{\partial x^i}, \quad i = 1, \dots, n \quad (1.2.4)$$

Supondo que:

$$\det \left[ \frac{\partial^2 F}{\partial x^i \partial x^j} \right] \neq 0 \quad (1.2.5)$$

podemos inverter a relação  $\mathbf{p} = \mathbf{p}(\mathbf{x})$ , para calcular os  $x^{i'}$ s como função dos  $p_i$ 's:  $\mathbf{x} = \mathbf{x}(\mathbf{p})$ . Definimos então a transformada de Legendre  $\Phi : \mathbb{R}_p^n \rightarrow \mathbb{R}$ , de  $F$ , através de:

$$\Phi(\mathbf{p}) = \mathbf{p}\mathbf{x}(\mathbf{p}) - F(\mathbf{x}(\mathbf{p})) \quad (1.2.6)$$

Suponhamos agora que temos um Lagrangeano  $L(t, \mathbf{x}, \dot{\mathbf{x}})$ , definido em  $\mathbb{R}_t \times \mathbb{R}_x^n \times \mathbb{R}_{\dot{x}}^n$ . Para cada  $(t, \mathbf{x})$  fixo, consideremos a função parcial  $F(\dot{\mathbf{x}}) = L(t, \mathbf{x}, \dot{\mathbf{x}})$  e calculemos a transformada de Legendre de  $F$ . Essa transformada é uma função  $H(t, \mathbf{x}, \mathbf{p})$ , a que se chama o **Hamiltoniano** correspondente ao Lagrangeano  $L$ , e que é definida em  $\mathbb{R}_t \times \mathbb{R}_x^n \times \mathbb{R}_p^n$ , através de:

$$\boxed{H(t, \mathbf{x}, \mathbf{p}) = \mathbf{p}\dot{\mathbf{x}} - L(t, \mathbf{x}, \dot{\mathbf{x}})|_{\dot{\mathbf{x}}=\dot{\mathbf{x}}(t, \mathbf{x}, \mathbf{p})}} \quad (1.2.7)$$

Nesta fórmula,  $\dot{\mathbf{x}} = \dot{\mathbf{x}}(t, \mathbf{x}, \mathbf{p})$  obtem-se invertendo a relação (com  $(t, \mathbf{x})$  fixo):

$$\begin{aligned} \mathbf{p} &= \mathbf{p}(t, \mathbf{x}, \dot{\mathbf{x}}) \\ &= \frac{\partial L}{\partial \dot{\mathbf{x}}}(t, \mathbf{x}, \dot{\mathbf{x}}) = L_{\dot{\mathbf{x}}}(t, \mathbf{x}, \dot{\mathbf{x}}) \end{aligned} \quad (1.2.8)$$

o que é possível se suposermos  $L$  **hiperregular**, isto é, se:

$$\det \left[ \frac{\partial L}{\partial \dot{x}^i \partial \dot{x}^j} \right] \neq 0 \quad (1.2.9)$$

Notemos que  $H$  não é mais do que a energia do Lagrangeano  $L$ , expressa nas coordenadas  $(t, \mathbf{x}, \mathbf{p})$ .

Calculemos agora a diferencial do Hamiltoniano:

$$\begin{aligned} H &= H(t, \mathbf{x}, \mathbf{p}) \\ &= \mathbf{p}\dot{\mathbf{x}}(t, \mathbf{x}, \mathbf{p}) - L(t, \mathbf{x}, \dot{\mathbf{x}}(t, \mathbf{x}, \mathbf{p})) \end{aligned} \quad (1.2.10)$$

Vem sucessivamente que:

$$\begin{aligned} dH &= H_t dt + H_x dx + H_p dp \\ &= (\mathbf{p}\dot{\mathbf{x}}_t - L_t - L_{\dot{\mathbf{x}}}\dot{\mathbf{x}}_t) dt + (\mathbf{p}\dot{\mathbf{x}}_x - L_x - L_{\dot{\mathbf{x}}}\dot{\mathbf{x}}_x) dx + (\dot{\mathbf{x}} + \mathbf{p}\dot{\mathbf{x}}_p - L_{\dot{\mathbf{x}}}\dot{\mathbf{x}}_p) dp \\ &= (\mathbf{p}\dot{\mathbf{x}}_t - L_t - \mathbf{p}\dot{\mathbf{x}}_t) dt + (\mathbf{p}\dot{\mathbf{x}}_x - L_x - \mathbf{p}\dot{\mathbf{x}}_x) dx + (\dot{\mathbf{x}} + \mathbf{p}\dot{\mathbf{x}}_p - \mathbf{p}\dot{\mathbf{x}}_p) dp \\ &= -L_t dt - L_x dx + \dot{\mathbf{x}} dp \end{aligned} \quad (1.2.11)$$

donde se deduz que:

$$H_t = -L_t, \quad H_p = \dot{\mathbf{x}} \quad H_x = -L_x \quad (1.2.12)$$

O sistema de Euler-Lagrange  $-\frac{d}{dt}L_{\dot{\mathbf{x}}} + L_x = 0$  escreve-se portanto na seguinte **forma canônica**:

$$\boxed{\begin{cases} \dot{\mathbf{x}} = H_p(t, \mathbf{x}, \mathbf{p}) \\ \dot{\mathbf{p}} = -H_x(t, \mathbf{x}, \mathbf{p}) \end{cases}} \quad (1.2.13)$$

já que  $\dot{\mathbf{p}} = \frac{d}{dt}L_{\dot{\mathbf{x}}} = L_x = -H_x$ . Estas equações dizem-se as **equações canônicas de Hamilton** associadas às equações de Euler-Lagrange.

Mais detalhadamente - uma curva  $t \mapsto \mathbf{x}(t)$  é solução das equações de Euler-Lagrange se e só se a curva:

$$t \mapsto \left( \mathbf{x}(t), \mathbf{p}(t) = L_{\dot{\mathbf{x}}}(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) \right)$$

é solução das equações canônicas.

♣ **Exemplo 1.5** ... Calcular as equações canônicas para o funcional:

$$I[\mathbf{x}(t)] = \int_0^\pi (2xy - 2x^2 + \dot{x}^2 - \dot{y}^2) dt, \quad \text{onde } \mathbf{x} = (x, y) \in \mathbb{R}^2$$

O Lagrangeano  $L(\mathbf{x}, \dot{\mathbf{x}}) = 2xy - 2x^2 + \dot{x}^2 - \dot{y}^2$  é hiperregular, já que:

$$\det \begin{bmatrix} L_{\dot{x}\dot{x}} & L_{\dot{x}\dot{y}} \\ L_{\dot{y}\dot{x}} & L_{\dot{y}\dot{y}} \end{bmatrix} = \det \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} = -4 \neq 0$$

Portanto os momentos  $\mathbf{p} = (p, q)$  são dados por:

$$\begin{aligned} p &= L_{\dot{x}} = 2\dot{x} &\Rightarrow &\dot{x} = \frac{p}{2} \\ q &= L_{\dot{y}} = -2\dot{y} &\Rightarrow &\dot{y} = -\frac{q}{2} \end{aligned}$$

O Hamiltoniano é:

$$\begin{aligned} H(x, y, p, q) &= p\dot{x} + q\dot{y} - L(x, y, \dot{x}, \dot{y})|_{\dot{x}=\frac{p}{2}, \dot{y}=-\frac{q}{2}} \\ &= 2x^2 - 2xy + \frac{p^2}{4} - \frac{q^2}{4} \end{aligned}$$

e as equações canônicas são:

$$\begin{cases} \dot{x} = \frac{p}{2} \\ \dot{y} = -\frac{q}{2} \\ \dot{p} = -4x + 2y \\ \dot{q} = 2x \end{cases}$$



### 1.3 Sistemas mecânicos conservativos

Para **sistemas mecânicos conservativos**, o Lagrangeano é dado, em notação matricial, por:

$$\boxed{L(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}\dot{\mathbf{x}}^T \mathbf{g}(\mathbf{x}) \dot{\mathbf{x}} - V(\mathbf{x})} \quad (1.3.1)$$

onde  $\mathbf{g}(\mathbf{x})$  é uma matriz simétrica definida positiva, que representa uma métrica Riemanniana em  $\mathbb{R}_x^n$ . As parcelas da soma (1.3.1) chamam-se respectivamente:

$$\begin{array}{ll} \frac{1}{2}\dot{\mathbf{x}}^T \mathbf{g}(\mathbf{x}) \dot{\mathbf{x}} & \text{energia cinética} \\ V(\mathbf{x}) & \text{energia potencial} \end{array} \quad (1.3.2)$$

O Lagrangeano (1.3.1) é hiperregular, uma vez que a matriz  $\mathbf{g}(\mathbf{x})$  é inversível  $\forall \mathbf{x}$ . Portanto a matriz inversa  $\mathbf{G}(\mathbf{x}) = \mathbf{g}(\mathbf{x})^{-1}$  define uma métrica contravariante em  $\mathbb{R}^n$ . Como:

$$\mathbf{p} = L_{\dot{\mathbf{x}}} = \dot{\mathbf{x}}^T \mathbf{g}(\mathbf{x}) \quad \Rightarrow \quad \dot{\mathbf{x}} = \mathbf{G}(\mathbf{x})\mathbf{p}^T, \quad \text{onde } \mathbf{G}(\mathbf{x}) = \mathbf{g}(\mathbf{x})^{-1}$$

o Hamiltoniano é dado por:

$$\begin{aligned} H(\mathbf{x}, \mathbf{p}) &= \mathbf{p}\dot{\mathbf{x}} - L(\mathbf{x}, \dot{\mathbf{x}})|_{\dot{\mathbf{x}}=\mathbf{G}(\mathbf{x})\mathbf{p}^T} \\ &= \mathbf{p}\mathbf{G}(\mathbf{x})\mathbf{p}^T - \frac{1}{2}\mathbf{p}\mathbf{G}(\mathbf{x})\mathbf{g}(\mathbf{x})\mathbf{G}(\mathbf{x})\mathbf{p}^T + V(\mathbf{x}) \\ &= \frac{1}{2}\mathbf{p}\mathbf{G}(\mathbf{x})\mathbf{p}^T + V(\mathbf{x}) \\ &= \frac{1}{2}G^{ij}(\mathbf{x})p_i p_j + V(\mathbf{x}) \end{aligned} \quad (1.3.3)$$

Como  $H$  não depende explicitamente de  $t$ ,  $H$  é um integral primeiro das equações de Hamilton. De facto, se  $(\mathbf{x}(t), \mathbf{p}(t))$  é uma solução dessas equações:

$$\begin{cases} \dot{\mathbf{x}}(t) &= H_{\mathbf{p}}(\mathbf{x}(t), \mathbf{p}(t)) \\ \dot{\mathbf{p}}(t) &= -H_{\mathbf{x}}(\mathbf{x}(t), \mathbf{p}(t)) \end{cases}$$

então:

$$\begin{aligned} \frac{d}{dt}H(\mathbf{x}(t), \mathbf{p}(t)) &= H_{\mathbf{x}}\dot{\mathbf{x}} + H_{\mathbf{p}}\dot{\mathbf{p}} \\ &= H_{\mathbf{x}}H_{\mathbf{p}} - H_{\mathbf{p}}H_{\mathbf{x}} \\ &= 0 \end{aligned} \tag{1.3.4}$$

de tal forma que:

$$H(\mathbf{x}(t), \mathbf{p}(t)) \equiv h \tag{1.3.5}$$

para uma certa constante  $h$ , dita o **nível de energia** da solução  $(\mathbf{x}(t), \mathbf{p}(t))$ .

Por exemplo, para uma partícula de massa  $m$ , movendo-se em  $\mathbb{R}^3$  sob a acção de um campo de forças  $\mathbf{F}(\mathbf{x}) = -\nabla V(\mathbf{x})$ , o Lagrangeano é:

$$L(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}m\dot{\mathbf{x}}^2 - V(\mathbf{x})$$

onde  $\dot{\mathbf{x}}^2 = \dot{\mathbf{x}} \cdot \dot{\mathbf{x}} = \|\dot{\mathbf{x}}\|^2 = \dot{\mathbf{x}}^T \dot{\mathbf{x}}$ , e o Hamiltoniano é:

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2m}\mathbf{p}^2 + V(\mathbf{x})$$

As equações de Hamilton são:

$$\begin{cases} \dot{\mathbf{x}} &= \mathbf{p}^T/m \\ \dot{\mathbf{p}} &= -\nabla V(\mathbf{x})^T \end{cases}$$

donde se deduz que:

$$\ddot{\mathbf{x}} = \frac{\dot{\mathbf{p}}^T}{m} = -\frac{\nabla V(\mathbf{x})}{m}$$

e como  $\mathbf{F}(\mathbf{x}) = -\nabla V(\mathbf{x})$ :

$$m\ddot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) \tag{1.3.6}$$

que é a famosa **equação de Newton**.

Mais geralmente, dado um sistema de  $N$  partículas de massas  $m_i$  e coordenadas  $\mathbf{x}_i = (x_i, y_i, z_i)$ , para  $i = 1, \dots, N$ , que se movem sob a acção de forças  $\mathbf{F}_i$  que derivam de um potencial  $V = V(\mathbf{x}_1, \dots, \mathbf{x}_N)$ , que depende apenas das posições das partículas:

$$\mathbf{F}_i = -\nabla_{\mathbf{x}_i} V, \quad i = 1, \dots, N$$

a energia cinética é:

$$\frac{1}{2} \sum_{i=1}^N m_i \dot{\mathbf{x}}_i^2$$

e a energia potencial é  $V$ . As equações do movimento são:

$$m_i \ddot{\mathbf{x}}_i = \mathbf{F}_i, \quad i = 1, \dots, N \tag{1.3.7}$$

♣ **Exemplo 1.6 (Oscilador harmónico)** ... O oscilador harmónico (de dimensão 1) é descrito pela seguinte ODE de segunda ordem em  $\mathbb{R}_x$ :

$$\ddot{x} + \omega^2 x = 0, \quad \omega \neq 0 \quad (1.3.8)$$

cuja solução geral é:

$$x(t) = A \cos(\omega t + b), \quad A, b \text{ constantes}$$

A equação (1.3.8) é a equação de Euler-Lagrange correspondente ao Lagrangeano (que não depende de  $t$ ):

$$L(x, \dot{x}) = \frac{\dot{x}^2}{2\omega} - \frac{\omega x^2}{2} \quad (1.3.9)$$

De facto:

$$-\frac{d}{dt}L_{\dot{x}} + L_x = -\frac{\ddot{x}}{\omega} - \omega x$$

Aplicando a transformada de Legendre a  $L$ , vem que  $p = L_{\dot{x}} = \frac{\dot{x}}{\omega}$ , donde  $\dot{x} = \omega p$ , e portanto o Hamiltoniano é:

$$H(x, p) = p\dot{x} - L(x, \dot{x})|_{\dot{x}=\omega p} = \frac{\omega}{2}(x^2 + p^2) \quad (1.3.10)$$

As equações canónicas são pois:

$$\begin{cases} \dot{x} = H_p = \omega p \\ \dot{p} = -H_x = -\omega x \end{cases} \quad (1.3.11)$$

cuja solução geral é:

$$\begin{cases} x(t) = A \cos(\omega t + b) \\ p(t) = -A \sin(\omega t + b) \end{cases}, \quad A = \sqrt{2a}, b \text{ constantes} \quad (1.3.12)$$

Note que de facto  $p(t) = L_{\dot{x}}(x(t), \dot{x}(t))$ .

♣ **Exemplo 1.7 (Movimento num campo central)** ... Consideremos um ponto material de massa  $m$ , que se move em  $\mathbb{R}^3 - \{\mathbf{0}\}$ , sob a influência de um **campo de forças central**:

$$F(\mathbf{x}) = \frac{\varphi(r)}{r} \mathbf{x}, \quad \text{onde } r = \|\mathbf{x}\| > 0 \quad (1.3.13)$$

Em (1.3.13),  $\varphi : ]0, \infty[ \rightarrow \mathbb{R}$  representa uma função contínua. Podemos então escrever:

$$F(\mathbf{x}) = -\nabla V(\mathbf{x}) \quad (1.3.14)$$

onde:

$$V(\mathbf{x}) = -\Phi(r), \quad \text{com } \Phi(r) = \int_{r_0}^r \varphi(\rho) d\rho \quad (1.3.15)$$

O Lagrangeano é:

$$L(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}m \dot{\mathbf{x}}^2 - V(\mathbf{x})$$

que, como não depende de  $t$ , implica a conservação de energia:

$$E_L = \frac{m \dot{\mathbf{x}}^2}{2} + V(\mathbf{x}) \equiv E \quad (1.3.16)$$

para alguma constante  $E$  (o nível de energia).

Definamos agora o **momento**  $\mathbf{p}(t)$  e o **momento angular**  $\boldsymbol{\ell}(t)$ , do movimento  $\mathbf{x}(t)$ , através de:

$$\mathbf{p}(t) = L_{\dot{\mathbf{x}}} = m\dot{\mathbf{x}}(t), \quad \boldsymbol{\ell}(t) = \mathbf{x}(t) \times \mathbf{p}(t) \quad (1.3.17)$$

Deduzimos então que:

$$\begin{aligned} \dot{\boldsymbol{\ell}}(t) &= \dot{\mathbf{x}}(t) \times \mathbf{p}(t) + \mathbf{x}(t) \times \dot{\mathbf{p}}(t) \\ &= \dot{\mathbf{x}} \times m\dot{\mathbf{x}} + \mathbf{x} \times m \frac{\varphi(r)}{r} \mathbf{x} \\ &= \mathbf{0} \end{aligned} \quad (1.3.18)$$

isto é, o momento angular  $\boldsymbol{\ell}(t)$  é conservado:

$$\boldsymbol{\ell}(t) \equiv \mathbf{a} \quad (1.3.19)$$

para algum vector constante  $\mathbf{a}$ . Os 4 integrais primeiros independentes (1.3.16) e (1.3.19), são suficientes para integrar as equações do movimento (??). De facto, podemos escolher um sistema de eixos tal que  $\mathbf{a}$  aponte na direcção positiva do eixo dos  $x'$ s:

$$\mathbf{a} = (0, 0, a), \quad a \geq 0$$

De  $\boldsymbol{\ell}(t) = \mathbf{x}(t) \times \mathbf{p}(t) = m(\mathbf{x}(t) \times \dot{\mathbf{x}}(t)) \equiv \mathbf{a}$ , obtemos que  $\mathbf{a} \cdot \mathbf{x}(t) = 0, \forall t$ . Portanto, se  $a > 0$ ,  $\mathbf{x}(t)$  está sempre no plano  $xy$  e o movimento processa-se neste plano:

$$\mathbf{x}(t) = (x(t), y(t), 0)$$

A conservação do momento angular (1.3.19), pode então ser escrita na forma:

$$x\dot{y} - y\dot{x} = \frac{a}{m} \quad (1.3.20)$$

que é a chamada lei das áreas de Kepler: “*as áreas varridas pelo vector de posição  $\mathbf{x}(t)$ , em tempos iguais, são iguais*”. Em particular, o movimento ou é linear ( $a = 0$ ) ou  $\mathbf{x}(t)$  e  $\dot{\mathbf{x}}(t)$  nunca são colineares.

A lei de conservação de energia (1.3.16) toma agora a forma seguinte:

$$\frac{m}{2}(\dot{x}^2 + \dot{y}^2) = E + \Phi(r) \quad (1.3.21)$$

onde  $\Phi$  é dada por (1.3.15), e  $r = \sqrt{x^2 + y^2}$ . Introduzindo coordenadas polares de pólo na origem:

$$x = r \cos \theta, \quad y = r \sin \theta$$

podemos escrever (1.3.20) e (1.3.21), em termos de  $r(t)$  e  $\theta(t)$ , na forma seguinte:

$$r^2 \dot{\theta} = \frac{a}{m} \quad (1.3.22)$$

$$\frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) = E + \Phi(r) \quad (1.3.23)$$

Analisemos mais detalhadamente o **problema de Kepler** em que:

$$F(\mathbf{x}) = \frac{-\gamma m M}{r^3} \mathbf{x}, \quad r = \|\mathbf{x}\| \quad (1.3.24)$$

Esta é a força gravitacional que um ponto material de massa  $M$ , fixo no centro  $\mathbf{0}$ , exerce sobre um ponto material de massa  $m$ , situado em  $\mathbf{x} \neq \mathbf{0}$ , de acordo com a **lei da atração universal de Newton**.  $\gamma$  é a constante universal de gravitação. Neste caso, temos que  $F(\mathbf{x}) = -\nabla V(\mathbf{x})$ , onde:

$$V(\mathbf{x}) = \Phi(r) = \frac{\gamma m M}{r}$$

Vamos supôr que o movimento não é linear ( $a > 0$ , em (1.3.22)). Então (1.3.22) e (1.3.23) ficam com o aspecto:

$$\dot{\theta} = \frac{C}{r^2}, \quad \text{onde} \quad C = a/m \quad (1.3.25)$$

$$\frac{1}{2}(\dot{r}^2 + r^2\dot{\theta}^2) = \frac{\gamma M}{r} + W, \quad \text{onde} \quad W = E/m \quad (1.3.26)$$

Destas duas equações deduzimos que:

$$\frac{1}{2}C^2 \left[ r^{-4} \left( \frac{dr}{d\theta} \right)^2 + r^{-2} \right] = \frac{\gamma M}{r} + W$$

e portanto a função  $s(\theta) = 1/r(\theta)$  satisfaz:

$$\frac{1}{2}C^2 \left[ \left( \frac{ds}{d\theta} \right)^2 + s^2 \right] - \gamma M s = W \quad (1.3.27)$$

Derivando esta equação em ordem a  $\theta$  obtemos:

$$\frac{ds}{d\theta} \left( C^2 \left[ \frac{d^2s}{d\theta^2} + s \right] - \gamma M \right) = 0$$

Como  $\frac{ds}{d\theta} \neq 0$ , excepto em pontos isolados, vemos que:

$$\frac{d^2s}{d\theta^2} + s = \frac{\gamma M}{C^2} \quad (1.3.28)$$

e portanto:

$$s(\theta) = \frac{\gamma M}{C^2} + \frac{\alpha}{C} \cos(\theta + \theta_0) \quad (1.3.29)$$

onde  $\alpha$  e  $\theta_0$  são constantes arbitrárias,  $\alpha > 0$ . Pondo:

$$k = \frac{C^2}{\gamma M}, \quad \epsilon = \frac{\alpha C}{\gamma M}$$

e recordando que  $r(\theta) = 1/s(\theta)$ , obtemos:

$$r(\theta) = \frac{k}{1 + \epsilon \cos(\theta + \theta_0)} \quad (1.3.30)$$

que é a equação polar de uma cónica com excentricidade  $\epsilon$ . Esta equação descreve uma elipse, uma parábola ou uma hipérbole, conforme  $0 < \epsilon < 1$ ,  $\epsilon = 1$  ou  $\epsilon > 1$ , respectivamente. Inserindo:

$$s(\theta) = \frac{1}{k} [1 + \epsilon \cos(\theta + \theta_0)], \quad s'(\theta) = -\frac{\epsilon}{k} \sin(\theta + \theta_0)$$

em (1.3.27), obtemos:

$$\epsilon^2 = 1 + \frac{2}{m} \left( \frac{C}{\gamma M} \right)^2 E$$

portanto  $E < 0$  corresponde a  $0 < \epsilon < 1$ , i.e., a uma elipse,  $E = 0$  corresponde a  $\epsilon = 1$ , i.e., a uma parábola, e, finalmente,  $E > 1$  corresponde a  $\epsilon > 1$ , i.e., a uma hipérbole.

O **problema geral dos dois corpos** reduz-se facilmente ao problema anterior. De facto, consideremos dois pontos materiais  $\mathbf{x}_1 = (x_1, y_1, z_1)$  de massa  $M > 0$  e  $\mathbf{x}_2 = (x_2, y_2, z_2)$  de massa  $m > 0$ , em  $\mathbb{R}^3$ . A equações de Newton são:

$$M\ddot{\mathbf{x}}_1 = -\frac{\gamma m M}{\|\mathbf{x}_1 - \mathbf{x}_2\|^3}(\mathbf{x}_1 - \mathbf{x}_2), \quad m\ddot{\mathbf{x}}_2 = -\frac{\gamma m M}{\|\mathbf{x}_1 - \mathbf{x}_2\|^3}(\mathbf{x}_2 - \mathbf{x}_1) \quad (1.3.31)$$

Introduzindo o baricentro  $\mathbf{x}_b$  através de:

$$(m + M) \mathbf{x}_b = M \mathbf{x}_1 + m \mathbf{x}_2 \quad (1.3.32)$$

vem que  $\ddot{\mathbf{x}}_b(t) = 0$  e portanto:

$$\mathbf{x}_b = \mathbf{a}t + \mathbf{c}$$

onde  $\mathbf{a}, \mathbf{c} \in \mathbb{R}^3$  são constantes. Podemos pois escolher o baricentro como origem de um sistema de coordenadas onde as equações de Newton permanecem inalteradas (sistema inercial). Temos então que:

$$\mathbf{x}(t) \equiv \mathbf{0}$$

Introduzindo coordenadas relativas  $\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1$  deduzimos que:

$$m\ddot{\mathbf{x}} = -\frac{kmM^*}{r^3} \mathbf{x}, \quad r = \|\mathbf{x}\|, \quad M^* = m + M$$

que é o problema de Kepler com um centro de massa  $M^*$  no baricentro  $\mathbf{x}_b = \mathbf{0}$ .

♣.

## 1.4 A forma de Poincaré-Cartan

Vamos agora discutir o problema seguinte:

- ♣ **Problema 1.2** ... Consideremos uma família a um parâmetro  $\alpha \in \mathbb{R}$  de curvas  $\mathbf{x}(\cdot; \alpha) \in C^1([t_0(\alpha), t_1(\alpha)], \mathbb{R}^n)$ , cujas extremidades variam com o parâmetro  $\alpha$ , e o funcional:

$$J(\alpha) = \int_{t_0(\alpha)}^{t_1(\alpha)} L(t, \mathbf{x}(t; \alpha), \dot{\mathbf{x}}(t; \alpha)) dt \quad (1.4.1)$$

O problema é mais uma vez calcular a curva da família para a qual  $J(\alpha)$  tem um mínimo local.



Figure 1.3:

No problema anterior, supomos que:

$$\alpha \longmapsto \left( t_0(\alpha), \mathbf{x}_0(\alpha) \stackrel{\text{def}}{=} \mathbf{x}(t_0(\alpha); \alpha) \right)$$

é uma curva suave em  $\mathbb{R}^{n+1}$ , ao longo da qual a extremidade esquerda das diversas curvas da família  $\mathbf{x}(\cdot; \alpha)$ , varia quando  $\alpha$  varia. Anàlogamente, supomos que:

$$\alpha \longmapsto \left( t_1(\alpha), \mathbf{x}_1(\alpha) \stackrel{\text{def}}{=} \mathbf{x}(t_1(\alpha); \alpha) \right)$$

é uma curva suave em  $\mathbb{R}^{n+1}$ , ao longo da qual a extremidade direita das diversas curvas da família  $\mathbf{x}(\cdot; \alpha)$ , varia quando  $\alpha$  varia.

Suponhamos que  $J(\alpha)$  tem um mínimo local para  $\alpha = 0$ , e representemos por  $\widehat{\mathbf{x}}(\cdot) = \mathbf{x}(\cdot; \alpha = 0)$  a curva onde esse mínimo é atingido. Adoptamos ainda as seguintes notações:

$$\widehat{t}_0 = t_0(0), \quad \widehat{t}_1 = t_1(0), \quad \dot{\widehat{\mathbf{x}}}(t) = \dot{\mathbf{x}}(\cdot; 0), \quad \frac{\partial \mathbf{x}}{\partial \alpha}(t; 0) = \boldsymbol{\eta}(t), \quad \frac{\partial \dot{\mathbf{x}}}{\partial \alpha}(t; 0) = \dot{\boldsymbol{\eta}}(t)$$

Vamos agora calcular  $\left. \frac{dJ(\alpha)}{d\alpha} \right|_{\alpha=0}$ . Pelo teorema fundamental do cálculo vem que:

$$\begin{aligned} \frac{dJ(\alpha)}{d\alpha} &= L\left(t_1(\alpha), \mathbf{x}(t_1(\alpha); \alpha), \dot{\mathbf{x}}(t_1(\alpha); \alpha)\right) \frac{dt_1}{d\alpha} \\ &\quad - L\left(t_0(\alpha), \mathbf{x}(t_0(\alpha); \alpha), \dot{\mathbf{x}}(t_0(\alpha); \alpha)\right) \frac{dt_0}{d\alpha} \\ &\quad + \int_{t_0(\alpha)}^{t_1(\alpha)} \left( L_{\mathbf{x}} \frac{\partial \mathbf{x}}{\partial \alpha} + L_{\dot{\mathbf{x}}} \frac{\partial \dot{\mathbf{x}}}{\partial \alpha} \right) dt \end{aligned}$$

Para  $\alpha = 0$ , podemos escrever isto na forma abreviada:

$$\left. \frac{dJ(\alpha)}{d\alpha} \right|_{\alpha=0} = \widehat{L}_1 \delta t_1 - \widehat{L}_0 \delta t_0 + \int_{\widehat{t}_0}^{\widehat{t}_1} \left( \widehat{L}_{\mathbf{x}}(t) \boldsymbol{\eta}(t) + \widehat{L}_{\dot{\mathbf{x}}}(t) \dot{\boldsymbol{\eta}}(t) \right) dt \quad (1.4.2)$$

onde adoptamos as seguintes notações:

$$\begin{aligned}\delta t_0 &= \frac{dt_0}{d\alpha}(0), & \delta t_1 &= \frac{dt_1}{d\alpha}(0) \\ \widehat{L}_1 &= L(t_1(0), \mathbf{x}(t_1(0); 0), \dot{\mathbf{x}}(t_1(0); 0)) = L(\widehat{t}_1, \widehat{\mathbf{x}}(t_1), \dot{\widehat{\mathbf{x}}}(t_1)) \\ \widehat{L}_0 &= L(t_0(0), \mathbf{x}(t_0(0); 0), \dot{\mathbf{x}}(t_0(0); 0)) = L(\widehat{t}_0, \widehat{\mathbf{x}}(t_0), \dot{\widehat{\mathbf{x}}}(t_0)) \\ \widehat{L}_{\mathbf{x}}(t) &= L_{\mathbf{x}}(t, \mathbf{x}(t; 0), \dot{\mathbf{x}}(t; 0)) = L_{\mathbf{x}}(t, \widehat{\mathbf{x}}(t), \dot{\widehat{\mathbf{x}}}(t)) \\ \widehat{L}_{\dot{\mathbf{x}}}(t) &= L_{\dot{\mathbf{x}}}(t, \mathbf{x}(t; 0), \dot{\mathbf{x}}(t; 0)) = L_{\dot{\mathbf{x}}}(t, \widehat{\mathbf{x}}(t), \dot{\widehat{\mathbf{x}}}(t))\end{aligned}$$

Agora integramos por partes a última parcela em (1.4.2), e obtemos:

$$\left. \frac{dJ(\alpha)}{d\alpha} \right|_{\alpha=0} = \widehat{L}_1 \delta t_1 - \widehat{L}_0 \delta t_0 + \widehat{L}_{\dot{\mathbf{x}}}\boldsymbol{\eta} \Big|_{\widehat{t}_0}^{\widehat{t}_1} + \int_{\widehat{t}_0}^{\widehat{t}_1} \left( \widehat{L}_{\mathbf{x}}(t) - \frac{d}{dt} \widehat{L}_{\dot{\mathbf{x}}}(t) \right) \boldsymbol{\eta}(t) dt \quad (1.4.3)$$

Vamos finalmente modificar as parcelas fora do integral. Para isso, começamos por derivar as identidades  $\mathbf{x}_0(\alpha) = \mathbf{x}(t_0(\alpha); \alpha)$  e  $\mathbf{x}_1(\alpha) = \mathbf{x}(t_1(\alpha); \alpha)$  em ordem a  $\alpha$ , para  $\alpha = 0$ , para obter:

$$\begin{aligned}\delta \mathbf{x}_0 &\stackrel{\text{def}}{=} \frac{d\mathbf{x}_0}{d\alpha}(0) \\ &= \dot{\mathbf{x}}(t_0(0); 0) \frac{dt_0}{d\alpha}(0) + \frac{\partial \mathbf{x}}{\partial \alpha}(t_0(0); 0) \\ &= \dot{\widehat{\mathbf{x}}}(\widehat{t}_0) \delta t_0 + \boldsymbol{\eta}(\widehat{t}_0)\end{aligned}$$

o que implica que:

$$\boldsymbol{\eta}(\widehat{t}_0) = \delta \mathbf{x}_0 - \dot{\widehat{\mathbf{x}}}(\widehat{t}_0) \delta t_0$$

Anàlogamente se obtém:

$$\boldsymbol{\eta}(\widehat{t}_1) = \delta \mathbf{x}_1 - \dot{\widehat{\mathbf{x}}}(\widehat{t}_1) \delta t_1$$

Agora substituímos estes valores de  $\boldsymbol{\eta}(\widehat{t}_0)$  e  $\boldsymbol{\eta}(\widehat{t}_1)$  na terceira parcela de (1.4.3). Após reordenar os termos, vem que:

$$\begin{aligned}\left. \frac{dJ(\alpha)}{d\alpha} \right|_{\alpha=0} &= \left[ \widehat{L}_1 - (\widehat{L}_{\dot{\mathbf{x}}})_1 \dot{\widehat{\mathbf{x}}}(\widehat{t}_1) \right] \delta t_1 - \left[ \widehat{L}_0 - (\widehat{L}_{\dot{\mathbf{x}}})_0 \dot{\widehat{\mathbf{x}}}(\widehat{t}_0) \right] \delta t_0 \\ &+ (\widehat{L}_{\dot{\mathbf{x}}})_1 \delta \mathbf{x}_1 - (\widehat{L}_{\dot{\mathbf{x}}})_0 \delta \mathbf{x}_0 + \int_{\widehat{t}_0}^{\widehat{t}_1} \left( \widehat{L}_{\mathbf{x}}(t) - \frac{d}{dt} \widehat{L}_{\dot{\mathbf{x}}}(t) \right) \boldsymbol{\eta}(t) dt \quad (1.4.4)\end{aligned}$$

Mas recordemos que:

$$E_L(t, \mathbf{x}, \dot{\mathbf{x}}) = L_{\dot{\mathbf{x}}}\dot{\mathbf{x}} - L(t, \mathbf{x}, \dot{\mathbf{x}})$$

o que permite escrever (1.4.4) na forma:

$$\boxed{\left. \frac{dJ(\alpha)}{d\alpha} \right|_{\alpha=0} = (L_{\dot{\mathbf{x}}}d\mathbf{x} - E_L dt) \Big|_{\widehat{t}_0}^{\widehat{t}_1} + \int_{\widehat{t}_0}^{\widehat{t}_1} \left( \widehat{L}_{\mathbf{x}}(t) - \frac{d}{dt} \widehat{L}_{\dot{\mathbf{x}}}(t) \right) \boldsymbol{\eta}(t) dt} \quad (1.4.5)$$

onde:

$$\begin{aligned}
(L_{\dot{\mathbf{x}}}d\mathbf{x} - E_L dt)(\widehat{t}_0) &\stackrel{\text{def}}{=} (L_{\dot{\mathbf{x}}}d\mathbf{x} - E_L dt)_{(\widehat{t}_0, \widehat{\mathbf{x}}(\widehat{t}_0))}(\delta t_0, \delta \mathbf{x}_0) \\
&= (\widehat{L}_{\dot{\mathbf{x}}})_0 \delta \mathbf{x}_0 - \left[ (\widehat{L}_{\dot{\mathbf{x}}})_0 \dot{\widehat{\mathbf{x}}}(\widehat{t}_0) - \widehat{L}_0 \right] \delta t_0 \\
(L_{\dot{\mathbf{x}}}d\mathbf{x} - E_L dt)(\widehat{t}_1) &\stackrel{\text{def}}{=} (L_{\dot{\mathbf{x}}}d\mathbf{x} - E_L dt)_{(\widehat{t}_1, \widehat{\mathbf{x}}(\widehat{t}_1))}(\delta t_1, \delta \mathbf{x}_1) \\
&= (\widehat{L}_{\dot{\mathbf{x}}})_1 \delta \mathbf{x}_1 - \left[ (\widehat{L}_{\dot{\mathbf{x}}})_1 \dot{\widehat{\mathbf{x}}}(\widehat{t}_1) - \widehat{L}_1 \right] \delta t_1
\end{aligned}$$

com:

$$(\delta t_0, \delta \mathbf{x}_0) \in T_{(\widehat{t}_0, \widehat{\mathbf{x}}(\widehat{t}_0))} \mathbb{R}^{n+1}, \quad \text{e} \quad (\delta t_1, \delta \mathbf{x}_1) \in T_{(\widehat{t}_1, \widehat{\mathbf{x}}(\widehat{t}_1))} \mathbb{R}^{n+1}$$

A fórmula (1.4.5) é fundamental para o que se segue. Nela surge a 1-forma em  $\mathbb{R} \times T\mathbb{R}^n$ :

$$\boxed{\boldsymbol{\theta}_L = L_{\dot{\mathbf{x}}}d\mathbf{x} - E_L dt} \quad (1.4.6)$$

onde  $E_L(t, \mathbf{x}, \dot{\mathbf{x}}) = L_{\dot{\mathbf{x}}}\dot{\mathbf{x}} - L(t, \mathbf{x}, \dot{\mathbf{x}})$ , chamada a **forma de Poincaré-Cartan** e que será muito importante em breve. Para Lagrangeanos hiperregulares, podemos transportar esta forma para  $\mathbb{R} \times T^*\mathbb{R}^n$ , via transformada de Legendre, para obter a forma:

$$\boxed{\boldsymbol{\theta}_H = \mathbf{p}d\mathbf{x} - H dt} \quad (1.4.7)$$

a que também chamamos **forma de Poincaré-Cartan**.

## 1.5 Problema com extremidades móveis

Como aplicação da teoria exposta na secção anterior, vamos discutir o problema seguinte:

- **♣ Problema 1.3** ... *Entre as curvas  $\mathbf{x}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ , que satisfazem as condições de fronteira:*

$$\Phi_0(t_0, \mathbf{x}(t_0)) = \mathbf{0}, \quad \Phi_1(t_1, \mathbf{x}(t_1)) = \mathbf{0} \quad (1.5.1)$$

onde  $\Phi_0 : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^k$  e  $\Phi_1 : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^\ell$ , calcular a curva para a qual o valor do funcional:

$$I[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} L(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \quad (1.5.2)$$

é mínimo.

♣.

Estamos a supôr mais uma vez que o intervalo  $[t_0, t_1]$  depende da curva  $\mathbf{x}(\cdot)$ . Supômos ainda que  $\Phi_0$  e  $\Phi_1$  são submersões de tal forma que  $\Sigma_0 = \Phi_0^{-1}(\mathbf{0})$  e  $\Sigma_1 = \Phi_1^{-1}(\mathbf{0})$  são subvariedades em  $\mathbb{R}^{n+1}$  de codimensão  $k$  e  $\ell$ , respectivamente.

Figure 1.4:

Se  $\widehat{\mathbf{x}}(\cdot) \in C^1([\widehat{t}_0, \widehat{t}_1], \mathbb{R}^n)$  é uma solução do problema 1.3, então também será solução do problema 1.1, com extremidades fixas  $\widehat{\mathbf{x}}(\widehat{t}_0)$  e  $\widehat{\mathbf{x}}(\widehat{t}_1)$ . Portanto  $\widehat{\mathbf{x}}(\cdot)$  satisfaz a equação de Euler-Lagrange (1.1.11):

$$-\frac{d}{dt}\widehat{L}_{\dot{\mathbf{x}}} + \widehat{L}_{\mathbf{x}} = \mathbf{0} \quad (1.5.3)$$

No entanto, para calcular  $\widehat{t}_0, \widehat{t}_1$  e ainda os  $2n$  parâmetros que caracterizam a solução pretendida (portanto, ao todo  $2n + 2$  parâmetros), apenas dispomos, para já, das  $k + \ell$  condições de fronteira (1.5.1),  $\Phi_0 = \mathbf{0}$  e  $\Phi_1 = \mathbf{0}$ . No entanto, como vamos ver, essa solução tem de verificar outras condições de fronteira adicionais, que fornecem as condições que faltam para determinar unívocamente a solução optimal (se esta existir!).

De facto, seja  $(\delta t_1, \delta \mathbf{x}_1) \in T_{(\widehat{t}_1, \widehat{\mathbf{x}}_1)}\Sigma_1$  um vector tangente arbitrário a  $\Sigma_1 = \Phi_1^{-1}(\mathbf{0})$  no ponto  $(\widehat{t}_1, \widehat{\mathbf{x}}_1)$ , e  $\gamma : \alpha \mapsto (t_1(\alpha), \mathbf{x}_1(\alpha))$  uma curva suave em  $\Sigma_1$ , tal que  $\gamma(0) = (t_1(0), \mathbf{x}_1(0)) = (\widehat{t}_1, \widehat{\mathbf{x}}_1)$  e  $\frac{d\gamma}{d\alpha}(0) = (\delta t_1, \delta \mathbf{x}_1)$ . Seja  $\mathbf{x}(t; \alpha)$  uma família a um parâmetro  $\alpha$  de curvas tais que  $\mathbf{x}(t; 0) = \widehat{\mathbf{x}}(t)$ ,  $\mathbf{x}(t_1(\alpha); \alpha) = \mathbf{x}_1(\alpha)$  e ainda  $\mathbf{x}(\widehat{t}_0; \alpha) \equiv \widehat{\mathbf{x}}(\widehat{t}_0)$ , isto é, a extremidade esquerda está fixa (figura 1.4).

Aplicando à família  $\mathbf{x}(t; \alpha)$  a teoria exposta na resolução do problema 1.2, nomeadamente a fórmula (1.4.5), obtemos:

$$\left. \frac{dJ(\alpha)}{d\alpha} \right|_{\alpha=0} = (L_{\dot{\mathbf{x}}}d\mathbf{x} - E_L dt) \Big|_{\widehat{t}_0}^{\widehat{t}_1} + \int_{\widehat{t}_0}^{\widehat{t}_1} \left( \widehat{L}_{\mathbf{x}}(t) - \frac{d}{dt}\widehat{L}_{\dot{\mathbf{x}}}(t) \right) \boldsymbol{\eta}(t) dt = 0 \quad (1.5.4)$$

Mas, atendendo a que  $\widehat{\mathbf{x}}(\cdot)$  satisfaz a equação de Euler-Lagrange (1.5.3), e ainda ao facto de que todas as curvas  $\mathbf{x}(t; \alpha)$  passam pelo ponto fixo  $(\widehat{t}_0, \widehat{\mathbf{x}}_0(\widehat{t}_0))$ , e portanto  $(\delta t_0, \delta \mathbf{x}_0) = (0, \mathbf{0})$ , concluímos que:

$$\begin{aligned} (L_{\dot{\mathbf{x}}}d\mathbf{x} - E_L dt)_{(\widehat{t}_1, \widehat{\mathbf{x}}(\widehat{t}_1))}(\delta t_1, \delta \mathbf{x}_1) &= (\widehat{L}_{\dot{\mathbf{x}}})_1 \delta \mathbf{x}_1 - (E_L)_1 \delta t_1 \\ &= (\widehat{L}_{\dot{\mathbf{x}}})_1 \delta \mathbf{x}_1 - \left[ (\widehat{L}_{\dot{\mathbf{x}}})_1 \dot{\widehat{\mathbf{x}}}(\widehat{t}_1) - \widehat{L}_1 \right] \delta t_1 \\ &= 0 \end{aligned} \quad (1.5.5)$$

para todo o vector tangente  $(\delta t_1, \delta \mathbf{x}_1)$  a  $\Sigma_1$  no ponto  $(\widehat{t}_1, \widehat{\mathbf{x}}(\widehat{t}_1))$ . De forma completamente análoga se deduz que:

$$\begin{aligned} (L_{\dot{\mathbf{x}}} d\mathbf{x} - E_L dt)_{(\widehat{t}_0, \widehat{\mathbf{x}}(\widehat{t}_0))} (\delta t_0, \delta \mathbf{x}_0) &= (\widehat{L}_{\dot{\mathbf{x}}})_0 \delta \mathbf{x}_0 - (E_L)_0 \delta t_0 \\ &= (\widehat{L}_{\dot{\mathbf{x}}})_0 \delta \mathbf{x}_0 - \left[ (\widehat{L}_{\dot{\mathbf{x}}})_0 \dot{\widehat{\mathbf{x}}}(\widehat{t}_0) - \widehat{L}_0 \right] \delta t_0 \\ &= 0 \end{aligned} \quad (1.5.6)$$

para todo o vector tangente  $(\delta t_0, \delta \mathbf{x}_0)$  a  $\Sigma_0$  no ponto  $(\widehat{t}_0, \widehat{\mathbf{x}}(\widehat{t}_0))$ . Estas relações (1.5.5) e (1.5.6) chamam-se **condições de transversalidade** para o problema 1.3.

Por exemplo, quando a extremidade direita se pode mover apenas no hiperplano  $t \equiv t_1$ , a condição (1.5.5) reduz-se a:

$$L_{\dot{\mathbf{x}}}(\widehat{t}_1, \widehat{\mathbf{x}}(\widehat{t}_1), \dot{\widehat{\mathbf{x}}}(\widehat{t}_1)) = \mathbf{0}$$

Como a dimensão de  $\Sigma_1 = \Phi_1^{-1}(\mathbf{0})$  é  $n + 1 - \ell$ , a relação de transversalidade (1.5.5), fornece  $n + 1 - \ell$  equações independentes. Anàlogamente, como a dimensão de  $\Sigma_0 = \Phi_0^{-1}(\mathbf{0})$  é  $n + 1 - k$ , a relação de transversalidade (1.5.6), fornece  $n + 1 - k$  equações independentes. Adicionando as  $k + \ell$  condições de fronteira  $\Phi_0 = \mathbf{0}$  e  $\Phi_1 = \mathbf{0}$ , que já tínhamos, obtemos finalmente as  $2n + 2$  relações que precisamos para determinar univocamente a solução optimal (se esta existir!).

♣ **Exemplo 1.8** ... Discutir as condições de transversalidade para o problema:

$$I[y(x)] = \int_{x_0}^{x_1} n(x, y) \sqrt{1 + (y')^2} dx, \quad y(x_0) = y_0, \quad y = \psi(x)$$

isto é, a extremidade esquerda está fixa, enquanto a direita se move na curva  $y = \psi(x)$ .

Como  $L(x, y, y') = n(x, y) \sqrt{1 + (y')^2}$ , vem que  $L_{y'} = \frac{y' n(x, y)}{\sqrt{1 + (y')^2}}$ . Qualquer vector tangente à curva  $y = \psi(x)$ , no ponto  $(x, \psi(x))$ , é da forma:

$$(\delta x, \delta y) = \lambda (1, \psi'(x)), \quad \lambda \in \mathbb{R}$$

Portanto a condição de transversalidade (1.5.5) tem a forma (com as correspondentes adaptações de notação  $t \mapsto x, \mathbf{x} \mapsto y$ ):

$$\begin{aligned} 0 &= (L_{y'} dy - E_L dx)_{(x, y = \psi(x))} (\delta x, \delta y) \\ &= L_{y'} \delta y - E_L \delta x \\ &= \lambda (L_{y'} \psi'(x) - E_L) \end{aligned} \quad (1.5.7)$$

onde  $E_L = L_{y'} y' - L$ . Portanto:

$$\begin{aligned} L_{y'} \psi'(x) - E_L &= L_{y'} \psi'(x) - L_{y'} y' + L \\ &= L_{y'} (\psi'(x) - y') + L \\ &= \frac{y' n(x, y)}{\sqrt{1 + (y')^2}} (\psi'(x) - y') + n(x, y) \sqrt{1 + (y')^2} \\ &= 0 \end{aligned} \quad (1.5.8)$$

ou:

$$\frac{n(x, y)(1 + \psi' y')}{\sqrt{1 + (y')^2}} = 0$$

Supondo que  $n(x, y) \neq 0$ , na extremidade direita da extremal, obtemos  $1 + \psi' y' = 0$  ou  $y' = -\frac{1}{\psi'}$ , isto é, a condição de transversalidade reduz-se neste caso a uma condição de ortogonalidade usual - a extremal deve intersectar perpendicularmente a curva  $y = \psi(x)$ , na sua extremidade direita.

♣ **Exemplo 1.9** ... Calcular a distância entre a parábola  $y = x^2$  e a recta  $x - y = 5$ .

O problema consiste em calcular o valor extremo de:

$$I[y(x)] = \int_{x_0}^{x_1} \sqrt{1 + (y')^2} dx, \quad y(x_0) = x_0^2, \quad y(x_1) = x_1 - 5$$

isto é, a extremidade esquerda move-se na parábola  $y = x^2$ , enquanto a direita se move na recta  $x - y = 5$ .

A solução geral da equação de Euler-Lagrange é  $y(x) = ax + b$ . Pretende-se pois calcular a extremal  $y(x) = ax + b$ ,  $x \in [x_0, x_1]$ , onde  $a, b, x_0$  e  $x_1$  são constantes a determinar. Como  $L = \sqrt{1 + (y')^2}$  e  $L_{y'} = \frac{y'}{\sqrt{1 + (y')^2}}$ , as condições de transversalidade (1.5.5) e (1.5.6) têm, neste caso, a forma:

$$\begin{aligned} (L_{y'} \delta y - E_L \delta x)|_{x=x_1} &= \lambda (L_{y'} - L_{y'} y' + L)|_{x=x_1} \\ &= \lambda \left( \frac{y'}{\sqrt{1 + (y')^2}} - y' \frac{y'}{\sqrt{1 + (y')^2}} + \sqrt{1 + (y')^2} \right) \Big|_{x=x_1} \\ &= 0 \end{aligned}$$

e:

$$\begin{aligned} (L_{y'} \delta y - E_L \delta x)|_{x=x_0} &= \lambda (L_{y'} 2x - L_{y'} y' + L)|_{x=x_0} \\ &= \lambda \left( (2x - y') \frac{y'}{\sqrt{1 + (y')^2}} + \sqrt{1 + (y')^2} \right) \Big|_{x=x_0} \\ &= 0 \end{aligned}$$

onde  $y' = a$ . Por outro lado, as condições de fronteira são  $y(x_0) = x_0^2$  e  $y(x_1) = x_1 - 5$ , isto é:

$$\begin{aligned} ax_0 + b &= x_0^2 \\ ax_1 + b &= x_1 - 5 \end{aligned}$$

e portanto, temos um sistema de 4 equações a 4 incógnitas  $x_0, x_1, a$  e  $b$ :

$$\left\{ \begin{array}{l} (1 - a) \frac{a}{\sqrt{1 + a^2}} + \sqrt{1 + a^2} = 0 \\ (2x_0 - a) \frac{a}{\sqrt{1 + a^2}} + \sqrt{1 + a^2} = 0 \\ ax_0 + b = x_0^2 \\ ax_1 + b = x_1 - 5 \end{array} \right.$$

cuja solução é:

$$a = -1, \quad b = 3/4, \quad x_0 = 1/2, \quad x_1 = 23/8$$

A equação da extremal é pois  $y(x) = -x + 3/4$ , e a distância pedida é:

$$\ell = \int_{1/2}^{23/8} \sqrt{1 + (-1)^2} dx = 19\sqrt{2}/8$$

♣.

## 1.6 A função de acção $S$ . Equação de Hamilton-Jacobi

$$S_t + H(t, \mathbf{x}, S_{\mathbf{x}}) = 0$$

Consideremos de novo o problema de minimizar o funcional  $I[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} L(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt$  com condições de fronteira  $\mathbf{x}(t_0) = \mathbf{x}_0$  e  $\mathbf{x}(t) = \mathbf{x}$ . Agora estamos a fixar a extremidade esquerda  $P_0 = (t_0, \mathbf{x}_0)$ , e variamos a extremidade direita  $P = (t, \mathbf{x}) \in \mathbb{R}^{n+1}$ . Como na secção anterior, supomos que o intervalo  $[t_0, t]$  depende da curva  $\mathbf{x}(\cdot)$ .

Vamos supôr ainda que existe um aberto  $\mathcal{U} \subseteq \mathbb{R}_{t\mathbf{x}}^{n+1}$  tal que, para todo o ponto  $P = (t, \mathbf{x}) \in \mathcal{U}$  existe uma única extremal que une  $P_0$  a  $P$ . Diz-se neste caso que, em  $\mathcal{U}$ , está definido um **feixe central de extremais**, de pólo  $P_0$ .

Neste caso, em cada ponto  $P = (t, \mathbf{x}) \in \mathcal{U}$  fica seleccionado um único vector tangente  $(1, \widehat{\mathbf{x}}(t)) \in T_P \mathbb{R}^{n+1}$ , que não é mais do que o vector velocidade, em  $t$ , do gráfico da única extremal  $\widehat{\mathbf{x}} : [t_0, t] \rightarrow \mathbb{R}^{n+1}$ , que une  $P_0 = (t_0, \mathbf{x}_0)$  a  $P = (t, \mathbf{x})$ . Consideremos a função  $\mathfrak{J} : \mathcal{U} \rightarrow \mathbb{R}^n$  definida por:

$$\mathfrak{J}(t, \mathbf{x}) = \widehat{\mathbf{x}}(t) \tag{1.6.1}$$

a que chamamos o **campo de inclinações** do feixe de extremais em  $\mathcal{U}$ . O respectivo gráfico será uma subvariedade de dimensão  $n + 1$  em  $\mathbb{R} \times T\mathbb{R}^n$ , parametrizado por:

$$\varphi(t, \mathbf{x}) = (t, \mathbf{x}, \mathfrak{J}(t, \mathbf{x})), \quad (t, \mathbf{x}) \in \mathcal{U} \tag{1.6.2}$$

Definimos agora uma função  $S$ , no aberto  $\mathcal{U}$ , chamada a **função acção**, cujo valor num ponto  $P = (t, \mathbf{x}) \in \mathcal{U}$ , é igual ao valor do funcional  $I$  na única extremal que une  $P_0$  a  $P$ . Portanto:

$$S(t, \mathbf{x}) = S(t_0, \mathbf{x}_0; t, \mathbf{x}) = \int_{t_0}^t L(\tau, \widehat{\mathbf{x}}(\tau), \dot{\widehat{\mathbf{x}}}(\tau)) dt \tag{1.6.3}$$

onde  $\widehat{\mathbf{x}} : [t_0, t] \rightarrow \mathbb{R}^{n+1}$  é a única extremal que une  $P_0 = (t_0, \mathbf{x}_0)$  a  $P = (t, \mathbf{x})$ . O nosso objectivo é calcular a diferencial de  $S$ .

Figure 1.5: Feixe central de extremais.

- **♣ Proposição 1.1** ... A diferencial da acção dada por:

$$dS = \varphi^* \boldsymbol{\theta}_L \quad (1.6.4)$$

onde  $\boldsymbol{\theta}_L = L_{\dot{\mathbf{x}}} d\mathbf{x} - E_L dt$  é a forma de Poincaré-Cartan em  $\mathbb{R} \times T\mathbb{R}^n$ .

**Dem.:** Seja  $(\delta t, \delta \mathbf{x}) \in T_{(t, \mathbf{x})} \mathbb{R}^{n+1}$  um vector tangente arbitrário a  $\mathbb{R}^{n+1}$  no ponto  $(t, \mathbf{x}) \in \mathcal{U}$ , e  $\gamma : \alpha \mapsto (t(\alpha), \mathbf{x}(\alpha))$  uma curva suave, em  $\mathcal{U}$ , tal que  $\gamma(0) = (t(0), \mathbf{x}(0)) = (t, \mathbf{x})$  e  $\frac{d\gamma}{d\alpha}(0) = (\delta t, \delta \mathbf{x})$ . Seja  $\mathbf{x}(t; \alpha)$  a família a um parâmetro  $\alpha$  de extremais, tal que  $\mathbf{x}(t; 0) = \widehat{\mathbf{x}}(t)$ ,  $\mathbf{x}(t(\alpha); \alpha) = \mathbf{x}(\alpha)$  e ainda  $\mathbf{x}(t_0; \alpha) \equiv \mathbf{x}_0$ . Por definição da acção:

$$S(t(\alpha), \mathbf{x}(\alpha)) = J(\alpha) = \int_{t_0}^{t(\alpha)} L(t, \mathbf{x}(t; \alpha), \dot{\mathbf{x}}(t; \alpha)) dt$$

Aplicando à família  $\mathbf{x}(t; \alpha)$  a teoria exposta na resolução do problema 1.2, nomeadamente a fórmula (1.4.5), obtemos:

$$\begin{aligned} dS_{(t, \mathbf{x})}(\delta t, \delta \mathbf{x}) &= \left. \frac{dS(t(\alpha), \mathbf{x}(\alpha))}{d\alpha} \right|_{\alpha=0} = \left. \frac{dJ(\alpha)}{d\alpha} \right|_{\alpha=0} \\ &= (L_{\dot{\mathbf{x}}} d\mathbf{x} - E_L dt) \Big|_{t_0}^t + \int_{t_0}^{t_1} \left( \widehat{L}_{\mathbf{x}}(t) - \frac{d}{dt} \widehat{L}_{\dot{\mathbf{x}}}(t) \right) \boldsymbol{\eta}(t) dt \\ &= (L_{\dot{\mathbf{x}}} d\mathbf{x} - E_L dt)_{(t, \mathbf{x})}(\delta t, \delta \mathbf{x}) \end{aligned} \quad (1.6.5)$$

atendendo a que  $\widehat{\mathbf{x}}(\cdot)$  satisfaz a equação de Euler-Lagrange (1.5.3), e ainda ao facto de que todas as curvas  $\mathbf{x}(t; \alpha)$  passam pelo ponto fixo  $P_0 = (t_0, \mathbf{x}_0)$ . Portanto  $dS = \varphi^*(L_{\dot{\mathbf{x}}} d\mathbf{x} - E_L dt)$ , como se pretendia.

♣.

Supondo agora que o Lagrangeano é hiperregular, podemos passar ao formalismo canónico, via transformada de Legendre. Definimos então o chamado **campo de momentos** do feixe central de extremais em  $\mathcal{U}$ , através de:

$$\mathbf{p}(t, \mathbf{x}) = L_{\dot{\mathbf{x}}}(t, \mathbf{x}, \mathfrak{J}(t, \mathbf{x})), \quad (t, \mathbf{x}) \in \mathcal{U} \quad (1.6.6)$$

O respectivo gráfico é agora uma subvariedade de dimensão  $n+1$  em  $\mathbb{R} \times T^*\mathbb{R}^n$ , parametrizado por:

$$\psi(t, \mathbf{x}) = (t, \mathbf{x}, \mathbf{p}(t, \mathbf{x})), \quad (t, \mathbf{x}) \in \mathcal{U} \quad (1.6.7)$$

Como o Hamiltoniano  $H = H(t, \mathbf{x}, \mathbf{p})$  é a energia  $E_L$ , expressa nas coordenadas canónicas  $(t, \mathbf{x}, \mathbf{p})$ , vemos que:

$$dS = \psi^* \boldsymbol{\theta}_H \quad (1.6.8)$$

onde  $\boldsymbol{\theta}_H = \mathbf{p}d\mathbf{x} - H dt$  é a forma de Poincaré-Cartan em  $\mathbb{R} \times T^*\mathbb{R}^n$ . Mais detalhadamente:

$$dS(t, \mathbf{x}) = \mathbf{p}(t, \mathbf{x})d\mathbf{x} - H(t, \mathbf{x}, \mathbf{p}(t, \mathbf{x}))dt \quad (1.6.9)$$

Como  $-H dt + \mathbf{p} d\mathbf{x} = dS = \frac{\partial S}{\partial t} dt + \frac{\partial S}{\partial \mathbf{x}} d\mathbf{x}$ , concluímos que:

$$\mathbf{p} = \frac{\partial S}{\partial \mathbf{x}} \quad \text{e} \quad H = -\frac{\partial S}{\partial t}$$

e portanto  $S$  satisfaz a PDE de primeira ordem:

$$\frac{\partial S}{\partial t}(t, \mathbf{x}) + H\left(t, \mathbf{x}, \frac{\partial S}{\partial \mathbf{x}}(t, \mathbf{x})\right) = 0 \quad (1.6.10)$$

ou simplesmente:

$$\boxed{S_t + H(t, \mathbf{x}, S_{\mathbf{x}}) = 0} \quad (1.6.11)$$

que se chama a **equação de Hamilton-Jacobi** para a função  $S = S(t, \mathbf{x})$ .

♣ **Exemplo 1.10** ... Consideremos o funcional:

$$I[y(x)] = \frac{1}{2} \int_0^a (y'^2 - y^2) dx$$

A equação de Euler-Lagrange é:

$$0 = \frac{d}{dx} L_{y'} - L_y = y'' + y$$

cuja a solução geral é:

$$y(x) = a \cos x + b \sin x$$

Considerando as extremais que passam na origem  $\mathbf{0} = (0, 0)$ , obtemos a família:

$$y(x) = b \sin x$$

(já que  $0 = y(0) = a$ ) que constitui um feixe central de extremais de pólo  $\mathbf{0}$ , no aberto  $\mathcal{U} = ]0, \pi[ \times \mathbb{R} \subset \mathbb{R}_{xy}^2$ . A extremal que une o ponto  $(0, 0)$  ao ponto  $(x, y) \in \mathcal{U}$  é  $y(\tau) =$

$\frac{y}{\sin x} \sin \tau$ ,  $0 \leq \tau \leq x$  e, portanto a acção é:

$$\begin{aligned} S(x, y) &= \frac{1}{2} \int_0^x \left( y'(\tau)^2 - y(\tau)^2 \right) d\tau \\ &= \frac{1}{2} \int_0^x \left( \left( \frac{y}{\sin x} \right)^2 \cos^2 \tau - \left( \frac{y}{\sin x} \right)^2 \sin^2 \tau \right) d\tau \\ &= \frac{1}{2} \left( \frac{y}{\sin x} \right)^2 \int_0^x \cos 2\tau d\tau \\ &= \frac{1}{2} \left( \frac{y}{\sin x} \right)^2 \frac{\sin 2x}{2} \\ &= \frac{1}{2} y^2 \cotg x \end{aligned}$$

A sua diferencial é:

$$dS = -\frac{1}{2} y^2 \operatorname{cosec}^2 x dx + y \cotg x dy$$

Vamos verificar que  $dS = p dy - H dx$  (com uma óbvia simplificação de notação). O campo de inclinações do feixe é:

$$\mathfrak{J}(x, y) = \left. \frac{d}{d\tau} \right|_{\tau=x} \frac{y}{\sin x} \sin \tau = y \cotg x$$

e portanto o respectivo campo de momentos é:

$$p(x, y) = L_{y'}(x, y, \mathfrak{J}(x, y)) = \mathfrak{J}(x, y) = y \cotg x$$

já que  $L = \frac{1}{2}(y'^2 - y^2)$  e  $p = L_{y'} = y'$ .

Por outro lado:  $H(x, y, p) = py' - L|_{y'=p} = y'^2 - \frac{1}{2}(y'^2 - y^2)|_{y'=p} = \frac{1}{2}(p^2 + y^2)$ , e portanto:

$$\begin{aligned} p(x, y)dy - H(x, y, p(x, y))dx &= y \cotg x dy - \frac{1}{2} y^2 (\cotg^2 x + 1) dx \\ &= -\frac{1}{2} y^2 \operatorname{cosec}^2 x dx + y \cotg x dy \\ &= dS \end{aligned}$$

como se pretendia. A equação de Hamilton-Jacobi é pois:

$$S_x + \frac{1}{2}(S_y^2 + y^2) = 0$$

♣ **Exemplo 1.11 (Equação H-J para sistemas mecânicos conservativos)** ... A equação de Hamilton-Jacobi para a acção  $S = S(t, \mathbf{x})$  tem, neste caso, a forma:

$$S_t + H(\mathbf{x}, S_x) = 0 \tag{1.6.12}$$

onde o Hamiltoniano  $H$  é dado por (1.3.3):

$$\begin{aligned} H(\mathbf{x}, \mathbf{p}) &= \frac{1}{2} \mathbf{p} \mathbf{G}(\mathbf{x}) \mathbf{p}^T + V(\mathbf{x}) \\ &= \frac{1}{2} G^{ij}(\mathbf{x}) p_i p_j + V(\mathbf{x}) \end{aligned} \tag{1.6.13}$$

Portanto, a equação de Hamilton-Jacobi é:

$$S_t + \frac{1}{2} S_{\mathbf{x}}^T G(\mathbf{x}) S_{\mathbf{x}} + V(\mathbf{x}) = 0 \quad (1.6.14)$$

Usando o método de separação de variáveis, vamos procurar soluções da forma:

$$S(t, \mathbf{x}) = f(t) + W(\mathbf{x})$$

Vem então que  $S_t(t, \mathbf{x}) = f'(t)$  e  $S_{\mathbf{x}}(t, \mathbf{x}) = \nabla W(\mathbf{x})$ . Substituindo na equação (1.6.12), obtemos:

$$f'(t) = -H(\mathbf{x}, \nabla W(\mathbf{x}))$$

Como o primeiro membro depende apenas de  $t$  e o segundo apenas de  $\mathbf{x}$ , isto implica que:

$$f(t) \equiv -ht + a, \quad \text{e} \quad H(\mathbf{x}, \nabla W(\mathbf{x})) \equiv h$$

onde  $h$  e  $a$  são constantes. A solução geral da equação (1.6.12) será pois:

$$S(t, \mathbf{x}) = -ht + W(\mathbf{x}) + a$$

onde  $W$  é solução da chamada **equação reduzida de Hamilton-Jacobi**:

$$H(\mathbf{x}, W_{\mathbf{x}}) \equiv h \quad (1.6.15)$$

Por exemplo, para uma partícula de massa  $m$ , movendo-se em  $\mathbb{R}^3$  sob a acção de um campo de forças  $\mathbf{F}(\mathbf{x}) = -\nabla V(\mathbf{x})$ , o Hamiltoniano é:

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2m} \mathbf{p}^2 + V(\mathbf{x})$$

e a equação reduzida de Hamilton-Jacobi, para  $W = W(\mathbf{x})$ , tem a forma (2.5.12), isto é:

$$\frac{1}{2m} (\nabla W)^2 + V(\mathbf{x}) \equiv h \quad (1.6.16)$$

onde  $(\nabla W)^2 = \nabla W \cdot \nabla W = \|\nabla W\|^2$ .

♣ **Exemplo 1.12 (A acção para uma partícula livre)** ... Para uma partícula de massa  $m$ , movendo-se livremente em  $\mathbb{R}^3$  (sob a acção de um campo de forças nulo) o Hamiltoniano é:

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2m} \mathbf{p}^2$$

e a equação de Hamilton-Jacobi, para  $S = S(t, \mathbf{x})$ , tem a forma:

$$S_t - \frac{1}{2m} (S_{\mathbf{x}})^2 = 0 \quad (1.6.17)$$

Em coordenadas cartesianas  $\mathbf{x} = (x, y, z)$ , a equação (1.6.20) escreve-se na forma:

$$S_t - \frac{1}{2m} (S_x^2 + S_y^2 + S_z^2) = 0 \quad (1.6.18)$$

A equação de Newton é:

$$m\ddot{\mathbf{x}} = \mathbf{0}$$

cuja solução geral é:

$$\mathbf{x}(\tau) = \mathbf{a}\tau + \mathbf{b}$$

a que corresponde um movimento rectilíneo e uniforme (velocidade constante). Dado um ponto arbitrário  $\mathbf{x}_0 \in \mathbb{R}^3$ , a família de trajectórias que, no instante  $\tau = t_0$ , começam em  $\mathbf{x}_0$ , é dada por:

$$\mathbf{x}(\tau) = \mathbf{a}(\tau - t_0) + \mathbf{x}_0$$

Essa família constitui um feixe central de pólo  $P_0 = (t_0, \mathbf{x}_0)$ , definido em  $\mathcal{U} = \{(t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^3 : t \neq t_0\}$ . A única trajectória, desse feixe, que une  $P_0 = (t_0, \mathbf{x}_0)$  a um outro ponto  $P = (t, \mathbf{x}) \in \mathcal{U}$  é:

$$\mathbf{x}(\tau) = \frac{\mathbf{x} - \mathbf{x}_0}{t - t_0} (\tau - t_0) + \mathbf{x}_0, \quad t_0 \leq \tau \leq t$$

A acção é dada por:

$$\begin{aligned} S(t, \mathbf{x}) = S(t_0, \mathbf{x}_0; t, \mathbf{x}) &= \frac{1}{2}m \int_{t_0}^t \frac{(\mathbf{x} - \mathbf{x}_0)^2}{(t - t_0)^2} d\tau \\ &= \frac{1}{2}m \frac{(\mathbf{x} - \mathbf{x}_0)^2}{t - t_0}, \quad (t, \mathbf{x}) \in \mathcal{U} \end{aligned} \quad (1.6.19)$$

atendendo a que o Lagrangeano é  $L(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}m\dot{\mathbf{x}}^2$  e a que  $\dot{\mathbf{x}}(\tau) = \frac{\mathbf{x} - \mathbf{x}_0}{t - t_0}$ .

O campo de inclinações do feixe é:

$$\mathfrak{J}(t, \mathbf{x}) = \left. \frac{d}{d\tau} \right|_{\tau=t} \frac{\mathbf{x} - \mathbf{x}_0}{t - t_0} (\tau - t_0) + \mathbf{x}_0 = \frac{\mathbf{x} - \mathbf{x}_0}{t - t_0}$$

e o campo de momentos:

$$\mathbf{p}(t, \mathbf{x}) = L_{\dot{\mathbf{x}}}(t, \mathbf{x}, \mathfrak{J}(t, \mathbf{x})) = m \frac{\mathbf{x} - \mathbf{x}_0}{t - t_0}$$

O pull-back da forma de Poincaré-Cartan,  $\psi^*\theta_H$  é:

$$\begin{aligned} \psi^*\theta_H &= \mathbf{p}(t, \mathbf{x})d\mathbf{x} - H(t, \mathbf{x}, \mathbf{p}(t, \mathbf{x}))dt \\ &= m \frac{\mathbf{x} - \mathbf{x}_0}{t - t_0} d\mathbf{x} - \frac{1}{4m^2} m^2 \frac{(\mathbf{x} - \mathbf{x}_0)^2}{(t - t_0)^2} dt \\ &= m \frac{\mathbf{x} - \mathbf{x}_0}{t - t_0} d\mathbf{x} - \frac{1}{4} \frac{(\mathbf{x} - \mathbf{x}_0)^2}{(t - t_0)^2} dt \end{aligned}$$

e é óbvio que  $dS = \psi^*\theta_H$ . É fácil ver que  $S$ , dada por (1.6.19), é solução da equação de Hamilton-Jacobi  $S_t - \frac{1}{2m}(S_{\mathbf{x}})^2 = 0$ .

Consideremos a hipersuperfície  $\Sigma_m$ , em  $\mathcal{U}$ , definida por:

$$\frac{1}{2}m \frac{(\mathbf{x} - \mathbf{x}_0)^2}{t - t_0} \equiv \frac{1}{2}m$$

Para cada  $t > t_0$  fixo, a intersecção da hipersuperfície  $\Sigma_m$ , com o hiperplano  $\{t\} \times \mathbb{R}^3$ , é dada por:

$$(\mathbf{x} - \mathbf{x}_0)^2 = t - t_0$$

e projectando estas superfícies no espaço de configuração  $\mathbb{R}_{\mathbf{x}}^3$ , obtemos uma família de esferas, centradas em  $\mathbf{x}_0$ , de raio  $(t - t_0)^{1/2}$ , a que chamamos as superfícies de onda da partícula livre.

♣ **Exemplo 1.13 (A acção para uma partícula num campo constante)** ... Para uma partícula de massa  $m$ , movendo-se em  $\mathbb{R}^3$  sob a acção de um campo de forças constante  $\mathbf{F}(\mathbf{x}) \equiv \mathbf{F}$ , o Hamiltoniano é:

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2m} \mathbf{p}^2 + \mathbf{F}\mathbf{x}$$

e a equação de Hamilton-Jacobi, para  $S = S(t, \mathbf{x})$ , tem a forma:

$$S_t - \frac{1}{2m}(S_{\mathbf{x}})^2 + \mathbf{F}\mathbf{x} = 0 \quad (1.6.20)$$

Em coordenadas cartesianas  $\mathbf{x} = (x, y, z)$ , a equação (1.6.20) escreve-se na forma:

$$S_t - \frac{1}{2m}(S_x^2 + S_y^2 + S_z^2) + Ax + By + Cz = 0 \quad (1.6.21)$$

onde  $\mathbf{F} = (A, B, C)$ . A equação de Newton é:

$$m\ddot{\mathbf{x}} = \mathbf{F}$$

cuja solução geral é:

$$\mathbf{x}(\tau) = \frac{\mathbf{F}}{m} \frac{\tau^2}{2} + \mathbf{a}\tau + \mathbf{b}$$

a que corresponde um movimento uniformemente acelerado (aceleração constante). Dado um ponto arbitrário  $\mathbf{x}_0 \in \mathbb{R}^3$ , a família de trajectórias que, no instante  $\tau = t_0$ , começam em  $\mathbf{x}_0$ , é dada por:

$$\mathbf{x}(\tau) = \frac{\mathbf{F}}{2m} (\tau^2 - t_0^2) + \mathbf{a}(\tau - t_0) + \mathbf{x}_0$$

Essa família constitui um campo central de pólo  $P_0 = (t_0, \mathbf{x}_0)$ . A única trajectória, desse campo, que une  $(t_0, \mathbf{x}_0)$  a um outro ponto  $(t, \mathbf{x})$  é:

$$\mathbf{x}(\tau) = \frac{\mathbf{F}}{2m} (\tau^2 - t_0^2) - \left[ \frac{\mathbf{F}}{2m} (t + t_0) + \frac{\mathbf{x} - \mathbf{x}_0}{t - t_0} \right] (\tau - t_0) + \mathbf{x}_0, \quad t_0 \leq \tau \leq t$$

Efectuando os cálculos para a acção, vemos que ela é dada por:

$$S(t, \mathbf{x}) = S(t_0, \mathbf{x}_0; t, \mathbf{x}) = \frac{\mathbf{F}^2}{24m} (t - t_0)^3 - \mathbf{F}\mathbf{x}(t - t_0) + \frac{\mathbf{x} - \mathbf{x}_0}{t - t_0} \quad (1.6.22)$$

atendendo a que o Lagrangeano é  $L(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}m\dot{\mathbf{x}}^2 - \mathbf{F}\mathbf{x}$  e a que  $\dot{\mathbf{x}}(\tau) = \frac{\mathbf{F}}{2m}(2\tau - t - t_0) - \frac{\mathbf{x} - \mathbf{x}_0}{t - t_0}$ .

♣ **Exemplo 1.14 (Equação H-J para o oscilador harmónico)** ... Continuemos com o exemplo 1.6. A equação de Hamilton-Jacobi para uma função  $S(t, x)$ , correspondente ao Hamiltoniano  $H = \frac{\omega}{2}(x^2 + p^2)$  é:

$$S_t + \frac{\omega}{2}(x^2 + S_x^2) = 0 \quad (1.6.23)$$

A solução geral da equação de Euler-Lagrange é, como vimos antes:

$$x(\tau) = A \cos(\omega\tau + b)$$

onde  $A, b$  são constantes. Se consideramos o feixe de extremais que parte de  $\mathbf{0} = (0, 0)$ , obtemos  $b = \pi/2$ , uma vez que  $0 = x(0) = A \cos b$ . Portanto essa extremais são do tipo  $x(\tau) = A \cos(\omega\tau + \pi/2)$ . Dado um ponto  $(t, x)$ , com  $0 < t < \pi$ , a única extremal que une  $(0, 0)$  a  $(t, x)$  tem por equação:

$$x(\tau) = \frac{x}{\cos(\omega t + \pi/2)} \cos(\omega\tau + \pi/2), \quad 0 \leq \tau \leq t$$

e portanto o campo de inclinações do feixe é:

$$\mathfrak{J}(t, x) = -x\omega \operatorname{tg}(\omega t + \pi/2)$$

Por outro lado, como  $L(x, \dot{x}) = \frac{\dot{x}^2}{2\omega} - \frac{\omega x^2}{2}$ , vem que  $L_{\dot{x}} = \dot{x}/\omega$ , e o campo de momentos do feixe é:

$$p(t, x) = L_{\dot{x}}(t, x, \mathfrak{J}(t, x)) = -x \operatorname{tg}(\omega t + \pi/2)$$

O pull-back da forma de Poincaré-Cartan,  $\psi^*\theta_H$ , é pois dada por:

$$\begin{aligned} \psi^*\theta_H &= p(t, x)dx - H(t, x, p(t, x))dt \\ &= -x \operatorname{tg}(\omega t + \pi/2)dx - \frac{\omega}{2}(x^2 + (-x \operatorname{tg}(\omega t + \pi/2))^2)dt \end{aligned}$$

A acção  $S = S(t, x)$  é dada por:

$$\begin{aligned} S(t, x) &= \int_0^t \left[ \frac{x^2\omega^2}{2\omega \cos^2(\omega\tau + \pi/2)} \sin^2(\omega\tau + \pi/2) - \frac{\omega}{2} \frac{x^2}{\cos^2(\omega\tau + \pi/2)} \cos^2(\omega\tau + \pi/2) \right] d\tau \\ &= -\frac{x^2\omega}{2 \cos^2(\omega t + \pi/2)} \int_0^t \cos(2\omega\tau + \pi) \\ &= -\frac{x^2 \sin(2\omega t + \pi)}{4 \cos^2(\omega t + \pi/2)} \\ &= \frac{x^2}{2} \operatorname{cotg}(\omega t) \end{aligned} \quad (1.6.24)$$

Verifiquemos que  $S$  satisfaz a equação de Hamilton-Jacobi (1.6.23):

$$S_t = -\frac{x^2\omega}{2 \sin^2(\omega t)}, \quad S_x = x \operatorname{cotg}(\omega t)$$

e portanto:

$$\begin{aligned} S_t + \frac{\omega}{2} (x^2 + S_x^2) &= -\frac{x^2\omega}{2\sin^2(\omega t)} + \frac{\omega}{2} (x^2 + x^2 \cotg^2(\omega t)) \\ &= 0 \end{aligned} \tag{1.6.25}$$

como se pretendia.

♣.

## 1.7 Um princípio variacional para sistemas Hamiltonianos. Princípio de Maupertuis

Na secção 1.2, as equações canónicas de Hamilton foram deduzidas a partir do formalismo Lagrangeano através da transformada de Legendre. Vamos nesta secção considerar as equações canónicas como equações básicas e ver como elas podem ser vistas como as equações de Euler-Lagrange de um certo problema variacional.

Recordemos que o funcional de acção, associado a um Lagrangeano  $L = L(t, \mathbf{x}, \dot{\mathbf{x}})$  é:

$$I[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} L(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt$$

Uma extremal deste funcional é, como sabemos, uma solução  $\mathbf{x} = \mathbf{x}(t)$  das equações de Euler-Lagrange. Por outro lado, no formalismo Hamiltoniano, essa extremal corresponde a uma solução  $(\mathbf{x}(t), \mathbf{p}(t))$  das equações canónicas, onde  $\mathbf{p}(t) = L_{\dot{\mathbf{x}}}(t, \mathbf{x}(t), \dot{\mathbf{x}}(t))$ . Como:

$$H(t, \mathbf{x}, \mathbf{p}) = \mathbf{p}\dot{\mathbf{x}} - L(t, \mathbf{x}, \dot{\mathbf{x}}) \quad \Rightarrow \quad L(t, \mathbf{x}, \dot{\mathbf{x}}) = \mathbf{p}\dot{\mathbf{x}} - H(t, \mathbf{x}, \mathbf{p})$$

vemos que:

$$\int_{t_0}^{t_1} L(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt = \int_{t_0}^{t_1} [\mathbf{p}(t)\dot{\mathbf{x}}(t) - H(t, \mathbf{x}(t), \mathbf{p}(t))] dt$$

O segundo integral tem a forma de um integral de linha de uma 1-forma ao longo da curva de fase  $t \mapsto (t, \mathbf{x}(t), \mathbf{p}(t))$ . É pois natural considerar a forma de Poincaré-Cartan:

$$\boldsymbol{\theta}_H \stackrel{\text{def}}{=} \mathbf{p}dx - H(t, \mathbf{x}, \mathbf{p}) dt \tag{1.7.1}$$

definida no **espaço de fases alargado**  $\mathbb{R} \times T^*\mathbb{R}^n$ . Dados dois pontos fixos  $P_0 = (t_0, \mathbf{x}_0)$  e  $P_1 = (t_1, \mathbf{x}_1)$ , com  $t_0 < t_1$ , no espaço de configuração alargado  $\mathbb{R} \times \mathbb{R}^n$ , consideremos o conjunto  $\mathcal{C}$  constituído por todas as curvas:

$$\gamma : t \mapsto (t, \mathbf{x}(t), \mathbf{p}(t))$$

Figure 1.6:

definidas no intervalo (fixo)  $[t_0, t_1]$ , tais que:

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad \text{e} \quad \mathbf{x}(t_1) = \mathbf{x}_1$$

Em particular os valores de  $\mathbf{p}(t_0)$  e  $\mathbf{p}(t_1)$  podem ser arbitrários (ver a figura 1.6).

No conjunto  $\mathcal{C}$ , de todas essas curvas, definimos o funcional seguinte:

$$\begin{aligned} \mathcal{S}_H[\gamma(\cdot)] &\stackrel{\text{def}}{=} \int_{\gamma} \boldsymbol{\theta}_H \\ &= \int_{\gamma} \mathbf{p} d\mathbf{x} - H(t, \mathbf{x}, \mathbf{p}) dt \\ &= \int_{t_0}^{t_1} \left( \mathbf{p}(t) \frac{d\mathbf{x}}{dt} - H(t, \mathbf{x}(t), \mathbf{p}(t)) \right) dt \end{aligned} \quad (1.7.2)$$

Temos então o seguinte teorema:

- ♣ **Teorema 1.1** ... *As equações canônicas de Hamilton:*

$$\begin{cases} \dot{\mathbf{x}} = H_{\mathbf{p}}(t, \mathbf{x}, \mathbf{p}) \\ \dot{\mathbf{p}} = -H_{\mathbf{x}}(t, \mathbf{x}, \mathbf{p}) \end{cases}$$

são as equações de Euler-Lagrange do problema variacional associado ao funcional  $\mathcal{S}_H$ , definido por (1.7.2), na classe de curvas  $\mathcal{C}$ , acima descrita. Mais precisamente, o funcional  $\mathcal{S}_H$  atinge um valor extremal em qualquer solução  $(\mathbf{x}(t), \mathbf{p}(t))$  das equações canônicas, relativamente a todas as variações que deixam as extremidades  $P_0 = (t_0, \mathbf{x}_0 = \mathbf{x}(t_0))$  e  $P_1 = (t_1, \mathbf{x}_1 = \mathbf{x}(t_1))$  fixas, podendo os valores de  $\mathbf{p}(t_0)$  e  $\mathbf{p}(t_1)$  ser arbitrários.

**Dem.:** O Lagrangeano do funcional  $\mathcal{S}_H$  é:

$$\mathcal{L}(t, \mathbf{x}, \mathbf{p}, \dot{\mathbf{x}}, \dot{\mathbf{p}}) = \mathbf{p}\dot{\mathbf{x}} - H(t, \mathbf{x}, \mathbf{p})$$

que é linear em  $\dot{\mathbf{x}}$  e não depende de  $\dot{\mathbf{p}}$ . Portanto:

$$\mathcal{L}_{\dot{\mathbf{x}}} = \mathbf{p}, \quad \mathcal{L}_{\dot{\mathbf{p}}} = \mathbf{0}, \quad \mathcal{L}_{\mathbf{x}} = -H_{\mathbf{x}}, \quad \mathcal{L}_{\mathbf{p}} = \dot{\mathbf{x}} - H_{\mathbf{p}}$$

e as equações de Euler-Lagrange são:

$$\begin{cases} -\frac{d}{dt}\mathcal{L}_{\dot{\mathbf{x}}} + \mathcal{L}_{\mathbf{x}} = \mathbf{0} \\ -\frac{d}{dt}\mathcal{L}_{\dot{\mathbf{p}}} + \mathcal{L}_{\mathbf{p}} = \mathbf{0} \end{cases} \Rightarrow \begin{cases} -\frac{d}{dt}\mathbf{p} - H_{\mathbf{x}} = \mathbf{0} \\ \mathbf{0} + \dot{\mathbf{x}} - H_{\mathbf{p}} = \mathbf{0} \end{cases} \Rightarrow \begin{cases} \dot{\mathbf{p}} = -H_{\mathbf{x}} \\ \dot{\mathbf{x}} = H_{\mathbf{p}} \end{cases}$$

que são exactamente as equações canónicas de Hamilton. Por outro lado, a energia de  $\mathcal{L}$  é:

$$\begin{aligned} E_{\mathcal{L}}(t, \mathbf{x}, \mathbf{p}, \dot{\mathbf{x}}, \dot{\mathbf{p}}) &= \mathcal{L}_{\dot{\mathbf{x}}}\dot{\mathbf{x}} + \mathcal{L}_{\dot{\mathbf{p}}}\dot{\mathbf{p}} - \mathcal{L}(t, \mathbf{x}, \mathbf{p}, \dot{\mathbf{x}}, \dot{\mathbf{p}}) \\ &= \mathbf{p}\dot{\mathbf{x}} - \mathbf{p}\dot{\mathbf{x}} + H(t, \mathbf{x}, \mathbf{p}) \\ &= H(t, \mathbf{x}, \mathbf{p}) \end{aligned} \quad (1.7.3)$$

O espaço de configuração alargado deste problema variacional é o espaço  $\mathbb{R} \times T^*\mathbb{R}^n \cong \mathbb{R}^{2n+1}$ , munido das coordenadas  $(t, \mathbf{x}, \mathbf{p})$ . As restrições são:

$$\begin{aligned} \Phi_0 : \mathbb{R}^{2n+1} &\rightarrow \mathbb{R}^{n+1}, & \Phi_0(t, \mathbf{x}, \mathbf{p}) &= (t - t_0, \mathbf{x} - \mathbf{x}_0) \\ \Phi_1 : \mathbb{R}^{2n+1} &\rightarrow \mathbb{R}^{n+1}, & \Phi_1(t, \mathbf{x}, \mathbf{p}) &= (t - t_1, \mathbf{x} - \mathbf{x}_1) \end{aligned} \quad (1.7.4)$$

e portanto as condições de transversalidade reduzem-se a:

$$\mathcal{L}_{\dot{\mathbf{p}}}\delta\mathbf{p}_0 = 0 = \mathcal{L}_{\dot{\mathbf{p}}}\delta\mathbf{p}_1 \quad (1.7.5)$$

que não impõem qualquer restrição aos valores de  $\mathbf{p}(t_0)$  e  $\mathbf{p}(t_1)$ .

♣.

Suponhamos agora que  $H = H(\mathbf{x}, \mathbf{p})$  não depende explicitamente de  $t$ . Por conservação de energia, se  $t \mapsto (\mathbf{x}(t), \mathbf{p}(t))$  é uma solução das equações canónicas então:

$$H(\mathbf{x}(t), \mathbf{p}(t)) \equiv h \quad (\text{constante})$$

Suponhamos ainda que  $h$  é valor regular de  $H$ , de tal forma que:

$$\Sigma_h \stackrel{\text{def}}{=} \{(\mathbf{x}, \mathbf{p}) \in T^*\mathbb{R}^n : H(\mathbf{x}, \mathbf{p}) \equiv h\} \quad (1.7.6)$$

é uma hipersuperfície de codimensão 1 em  $T^*\mathbb{R}^n$ . Uma solução das equações canónicas que comece em  $\Sigma_h$  permanecerá sempre em  $\Sigma_h$ .

Fixemos dois pontos  $\mathbf{x}_0, \mathbf{x}_1 \in \mathbb{R}^n$  e consideremos o conjunto  $\mathcal{C}_h$  constituído por todas as curvas:

$$\gamma : t \mapsto (\mathbf{x}(t), \mathbf{p}(t))$$

definidas no intervalo  $[t_0, t_1]$  (que agora depende de  $\gamma$ ), tais que:

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad \text{e} \quad \mathbf{x}(t_1) = \mathbf{x}_1 \quad (1.7.7)$$

e que têm energia constante  $h$ , i.e.:

$$H(\mathbf{x}(t), \mathbf{p}(t)) \equiv h \quad (1.7.8)$$

Figure 1.7:

Note que agora os valores de  $t_0, t_1, \mathbf{p}_0 = \mathbf{p}(t_0)$  e  $\mathbf{p}_1 = \mathbf{p}(t_1)$  podem ser arbitrários (ver a figura 1.7). A fixação do nível de energia é de certa forma compensada pela variação do intervalo de parametrização da curva.

Definamos o **funcional de ação reduzida**  $\mathcal{S}_{red}$ , através de:

$$\boxed{\mathcal{S}_{red}[\gamma(\cdot)] \stackrel{\text{def}}{=} \int_{\gamma} \mathbf{p} d\mathbf{x}} \quad (1.7.9)$$

e vamos mostrar que este funcional é estacionário em cada solução  $\gamma(t) = (\mathbf{x}(t), \mathbf{p}(t))$  das equações canônicas, relativamente a variações em  $\mathcal{C}_h$ .

Para isso, consideremos uma família a um parâmetro de curvas  $\gamma_{\alpha} = \gamma(\cdot; \alpha) \in \mathcal{C}_h$ :

$$\gamma_{\alpha}(t) = \gamma(t; \alpha) = (\mathbf{x}(t; \alpha), \mathbf{p}(t; \alpha)), \quad t \in [t_0(\alpha), t_1(\alpha)] \quad (1.7.10)$$

tais que:

$$\mathbf{x}(t_0(\alpha), \alpha) \equiv \mathbf{x}_0, \quad \text{e} \quad \mathbf{x}(t_1(\alpha), \alpha) \equiv \mathbf{x}_1$$

Temos então que:

$$\begin{aligned} \mathcal{S}_{red}(\alpha) &\stackrel{\text{def}}{=} \int_{\gamma_{\alpha}} \mathbf{p} d\mathbf{x} \\ &= \int_{t_0(\alpha)}^{t_1(\alpha)} \left( \mathbf{p}(t; \alpha) \frac{d\mathbf{x}(t; \alpha)}{dt} \right) dt \end{aligned} \quad (1.7.11)$$

Calculemos a derivada em ordem a  $\alpha$ , para  $\alpha = 0$ . Nesse cálculo, usaremos as notações seguintes:

$$\begin{aligned} t_0(0) = t_0, \quad t_1(0) = t_1, \quad \frac{dt_0}{d\alpha}(0) = \delta t_0, \quad \frac{dt_1}{d\alpha}(0) = \delta t_1 \\ \frac{\partial \mathbf{x}}{\partial \alpha}(t; 0) = \boldsymbol{\eta}(t), \quad \frac{\partial \mathbf{p}}{\partial \alpha}(t; 0) = \boldsymbol{\xi}(t) \\ \boldsymbol{\eta}(t_0) = \boldsymbol{\eta}_0, \quad \boldsymbol{\eta}(t_1) = \boldsymbol{\eta}_1, \quad \boldsymbol{\xi}(t_0) = \boldsymbol{\xi}_0, \quad \boldsymbol{\xi}(t_1) = \boldsymbol{\xi}_1 \\ \mathbf{p}(t_0; 0) = \mathbf{p}_0, \quad \mathbf{p}(t_1; 0) = \mathbf{p}_1 \end{aligned}$$

Como estamos a supôr que:

$$\mathbf{x}_0(\alpha) = \mathbf{x}(t_0(\alpha); \alpha) \equiv \mathbf{x}_0, \quad \text{e} \quad \mathbf{x}_1(\alpha) = \mathbf{x}(t_1(\alpha); \alpha) \equiv \mathbf{x}_1$$

isso implica que, para  $\alpha = 0$ :

$$\mathbf{0} = \delta \mathbf{x}_0 \stackrel{\text{def}}{=} \frac{d\mathbf{x}_0}{d\alpha}(0) = \dot{\mathbf{x}}(t_0)\delta t_0 + \boldsymbol{\eta}_0, \quad \text{e} \quad \mathbf{0} = \delta \mathbf{x}_1 \stackrel{\text{def}}{=} \frac{d\mathbf{x}_1}{d\alpha}(0) = \dot{\mathbf{x}}(t_1)\delta t_1 + \boldsymbol{\eta}_1 \quad (1.7.12)$$

Calculando finalmente a derivada de  $\mathcal{S}_{red}(\alpha)$ , em ordem a  $\alpha$ , para  $\alpha = 0$ , vem que:

$$\begin{aligned} \frac{d\mathcal{S}_{red}}{d\alpha}(0) &= [\mathbf{p}_1 \dot{\mathbf{x}}(t_1)] \delta t_1 - [\mathbf{p}_0 \dot{\mathbf{x}}(t_0)] \delta t_0 + \int_{t_0}^{t_1} \frac{d}{d\alpha} \Big|_{\alpha=0} \left( \mathbf{p}(t; \alpha) \frac{d\mathbf{x}(t; \alpha)}{dt} \right) dt \\ &= -\mathbf{p}_1 \boldsymbol{\eta}_1 + \mathbf{p}_0 \boldsymbol{\eta}_0 + \mathbf{p}(t) \boldsymbol{\eta}(t) \Big|_{t_0}^{t_1} + \int_{t_0}^{t_1} [\dot{\mathbf{x}}(t) \boldsymbol{\xi}(t) - \dot{\mathbf{p}}(t) \boldsymbol{\eta}(t)] dt \\ &= \int_{t_0}^{t_1} [\dot{\mathbf{x}}(t) \boldsymbol{\xi}(t) - \dot{\mathbf{p}}(t) \boldsymbol{\eta}(t)] dt \end{aligned} \quad (1.7.13)$$

onde usamos integração por partes e as condições (1.7.12). Suponhamos agora que  $(\mathbf{x}(t), \mathbf{p}(t))$  é solução das equações canónicas, com energia constante igual a  $h$ :

$$H(\mathbf{x}(t; \alpha), \mathbf{p}(t; \alpha)) \equiv h$$

Calculando a derivada em ordem a  $\alpha$ , para  $\alpha = 0$ , vem que:

$$\begin{aligned} 0 &= \frac{d}{d\alpha} \Big|_{\alpha=0} H(\mathbf{x}(t; \alpha), \mathbf{p}(t; \alpha)) \\ &= H_{\mathbf{x}}(\mathbf{x}(t), \mathbf{p}(t)) \boldsymbol{\eta}(t) + H_{\mathbf{p}}(\mathbf{x}(t), \mathbf{p}(t)) \boldsymbol{\xi}(t) \\ &= \dot{\mathbf{p}}(t) \boldsymbol{\eta}(t) - \dot{\mathbf{x}}(t) \boldsymbol{\xi}(t) \end{aligned} \quad (1.7.14)$$

já que estamos a supôr que  $(\mathbf{x}(t), \mathbf{p}(t))$  é solução das equações canónicas. Concluindo:

$$\frac{d\mathcal{S}_{red}}{d\alpha}(0) = 0$$

e, resumindo toda esta discussão, podemos pois enunciar o seguinte teorema:

- **♣ Teorema 1.2 (Princípio de Maupertuis)** ... *Suponhamos que  $H = H(\mathbf{x}, \mathbf{p})$  não depende explicitamente de  $t$ , e que fixamos um certo nível regular de energia constante  $h$ . Então as soluções  $(\mathbf{x}(t), \mathbf{p}(t))$  das equações canónicas são as extremais do funcional de acção reduzida:*

$$\mathcal{S}_{red}[\gamma(\cdot)] = \int_{\gamma} \boldsymbol{\theta} = \int_{\gamma} \mathbf{p} d\mathbf{x} \quad (1.7.15)$$

relativamente à classe  $\mathcal{C}_h$  de todas as curvas situadas em  $\Sigma_h$  (portanto de energia constante  $h$ ), que deixam as extremidades  $\mathbf{x}_0$  e  $\mathbf{x}_1$  fixas (podendo os valores de  $t_0, t_1, \mathbf{p}_0$  e  $\mathbf{p}_1$  ser arbitrários).



## 1.8 Os princípios variacionais de Jacobi e de Fermat. Analogia óptico-mecânica

Consideremos uma partícula de massa  $m$ , movendo-se em  $\mathbb{R}^3$ , sob a acção de um campo de forças conservativo  $\mathbf{F}(\mathbf{x}) = -\nabla V(\mathbf{x})$ . Como já sabemos, o Hamiltoniano é:

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2m} \mathbf{p}^2 + V(\mathbf{x})$$

Consideremos, como na secção anterior, a restrição do funcional de acção reduzida  $\mathcal{S}_{red} = \int \mathbf{p} d\mathbf{x}$ , à classe  $\mathcal{C}_h$  de curvas regulares (com velocidade que nunca se anula), de energia constante  $h$ . Ao longo de cada uma dessas curvas, temos que:

$$H(\mathbf{x}(t), \mathbf{p}(t)) = \frac{1}{2m} \mathbf{p}(t)^2 + V(\mathbf{x}(t)) \equiv h$$

e portanto  $h > V(\mathbf{x}(t))$  e ainda:

$$\|\mathbf{p}(t)\| = \sqrt{2m [h - V(\mathbf{x}(t))]} > 0 \quad (1.8.1)$$

Por outro lado, ao longo de uma extremal, i.e., ao longo de uma solução das equações canónicas, tem-se que  $\dot{\mathbf{x}} = H_{\mathbf{p}} = \frac{\mathbf{p}}{m}$ , isto é,  $\dot{\mathbf{x}}$  e  $\mathbf{p}$  são colineares, e daí que<sup>2</sup>:

$$\mathbf{p}\dot{\mathbf{x}} = \mathbf{p} \cdot \dot{\mathbf{x}} = \|\mathbf{p}\| \|\dot{\mathbf{x}}\|$$

Portanto:

$$\begin{aligned} \mathcal{S}_{red}[\mathbf{x}(\cdot), \mathbf{p}(\cdot)] &= \int \mathbf{p}(t) \dot{\mathbf{x}}(t) dt \\ &= \int \|\mathbf{p}\| \|\dot{\mathbf{x}}\| dt \\ &= \int \sqrt{2m [h - V(\mathbf{x}(t))]} ds \end{aligned} \quad (1.8.2)$$

isto é:

- **♣ Proposição 1.2 (Princípio de Jacobi)** ... As projecções  $\mathbf{x}(t)$ , das soluções regulares de energia constante  $h$ , das equações canónicas, com Hamiltoniano  $H(\mathbf{x}, \mathbf{p}) = \frac{1}{2m} \mathbf{p}^2 + V(\mathbf{x})$ , são as geodésicas (não parametrizadas) da métrica Riemanniana em  $\mathcal{U} = \{\mathbf{x} \in \mathbb{R}^3 : V(\mathbf{x}) < h\}$ :

$$d\sigma \stackrel{def}{=} \sqrt{2m [h - V(\mathbf{x})]} ds \quad (1.8.3)$$

onde  $ds$  é a métrica Euclideana usual em  $\mathbb{R}^3$ .

♣.

<sup>2</sup>pela desigualdade de Cauchy-Schwartz, que, neste caso, é igualdade já que  $\dot{\mathbf{x}} = \frac{\mathbf{p}}{m}$

Consideremos agora o Hamiltoniano  $H(\mathbf{x}, \mathbf{p}) = \frac{\|\mathbf{p}\|}{n(\mathbf{x})}$ , onde  $\mathbf{x} = (x, y, z)$ ,  $\mathbf{p} = (p, q, r)$  e  $\|\mathbf{p}\| = \sqrt{p^2 + q^2 + r^2}$ . Este Hamiltoniano descreve a propagação dos raios de luz num meio isotrópico com índice de refração  $n(\mathbf{x}) > 0$  (ver [9] ou [11]). Pondo  $v(\mathbf{x}) = 1/n(\mathbf{x})$ , o Hamiltoniano escreve-se na forma:

$$H(\mathbf{x}, \mathbf{p}) = v(\mathbf{x}) \|\mathbf{p}\|$$

Consideremos, como antes, a restrição do funcional de acção reduzida  $\mathcal{S}_{red} = \int \mathbf{p} dx$ , à classe de curvas regulares de energia constante  $h = 1$ :

$$1 \equiv H(\mathbf{x}(t), \mathbf{p}(t)) = v(\mathbf{x}) \|\mathbf{p}\| \quad (1.8.4)$$

Ao longo de uma extremal, i.e., ao longo de uma solução das equações canónicas, tem-se que  $\dot{\mathbf{x}} = H_{\mathbf{p}} = v(\mathbf{x}) \frac{\mathbf{p}}{\|\mathbf{p}\|}$ . Em particular,  $\|\dot{\mathbf{x}}\| = v(\mathbf{x})$ , isto é,  $v(\mathbf{x}) = 1/n(\mathbf{x})$  é a velocidade com que o raio de luz passa em  $\mathbf{x}$ .

Como, por (1.8.4),  $v(\mathbf{x})\|\mathbf{p}\| = 1$ , tem-se que  $\|\mathbf{p}\| = 1/v(\mathbf{x})$ . Por outro lado, como  $\dot{\mathbf{x}} = v(\mathbf{x}) \frac{\mathbf{p}}{\|\mathbf{p}\|}$ , vemos que  $\mathbf{p}$  e  $\dot{\mathbf{x}}$  são colineares e portanto  $\mathbf{p}\dot{\mathbf{x}} = \mathbf{p} \cdot \dot{\mathbf{x}} = \|\mathbf{p}\| \|\dot{\mathbf{x}}\|$ . Daí que:

$$\begin{aligned} \mathcal{S}_{red}[\mathbf{x}(\cdot), \mathbf{p}(\cdot)] &= \int \mathbf{p}(t)\dot{\mathbf{x}}(t) dt \\ &= \int \|\mathbf{p}\| \|\dot{\mathbf{x}}\| dt \\ &= \int \frac{1}{v(\mathbf{x})} ds \\ &= \int n(\mathbf{x}) ds \end{aligned} \quad (1.8.5)$$

O integral  $\int_{\gamma} n(\mathbf{x}) ds$  chama-se o **comprimento óptico** do raio  $\gamma$ . Note que esse mesmo integral é igual a  $\int_{\gamma} \frac{ds}{v(\mathbf{x})}$ , e portanto é o **tempo de percurso** da luz, ao longo do raio  $\gamma$ . Obtemos assim o seguinte:

- ♣ **Proposição 1.3 (Princípio de Fermat)** ... *Entre todas as curvas diferenciáveis que unem dois pontos fixos  $\mathbf{x}_0$  e  $\mathbf{x}_1$ , o caminho seguido efectivamente pela luz é aquele em que o tempo de percurso atinge um extremo. Estas curvas (os raios de luz) são as geodésicas da métrica:*

$$d\sigma = n(\mathbf{x}) ds \quad (1.8.6)$$

onde  $ds$  é a métrica Euclideana usual em  $\mathbb{R}^3$ .

♣.

Comparando os dois princípios anteriores - o de Jacobi, no contexto da mecânica clássica (conservativa) com o de Fermat, no contexto da óptica geométrica - mais especificamente, as fórmulas (1.8.3) e (1.8.6), concluimos que, pondo:

$$n(\mathbf{x}) = \sqrt{2m[h - V(\mathbf{x})]} \quad (1.8.7)$$

a mecânica clássica pode ser interpretada como uma óptica geométrica de propagação de raios num meio isotrópico de índice de refração  $n(\mathbf{x}) = \sqrt{2m [h - V(\mathbf{x})]}$ .

Esta analogia este na base dos trabalhos de Hamilton e à sua formulação geométrica da mecânica clássica, hoje chamada mecânica Hamiltoniana. Mais tarde, com Schrödinger, essa mesma analogia esteve também na base da criação da mecânica ondulatória e posteriormente da mecânica quântica.

## 1.9 Feixes de extremais. A iconal

Vamos supôr que, num aberto  $\mathcal{U}$  do espaço de configuração alargado  $\mathbb{R}_{t\mathbf{x}}^{n+1}$ , se verifica a propriedade seguinte: dados dois pontos quaisquer  $P_0 = (t_0, \mathbf{x}_0)$  e  $P_1 = (t_1, \mathbf{x}_1)$ , com  $t_0 < t_1$ , existe uma e uma só extremal  $\mathbf{x}(t)$  (solução das equações de Euler-Lagrange), tal que  $\mathbf{x}(t_0) = \mathbf{x}_0$ ,  $\mathbf{x}(t_1) = \mathbf{x}_1$  e ainda  $(t, \mathbf{x}(t)) \in \mathcal{U}$ ,  $\forall t \in [t_0, t_1]$  (ver a figura 1.8).

Figure 1.8: .

Diz-se então que  $\mathcal{U}$  é coberto por um **feixe de extremais**. A extremal que (cujo gráfico) une  $P_0 = (t_0, \mathbf{x}_0)$  a  $P_1 = (t_1, \mathbf{x}_1)$  será representada por:

$$\mathbf{x} = \mathbf{x}(t; t_0, \mathbf{x}_0; t_1, \mathbf{x}_1), \quad t_0 \leq t \leq t_1 \quad (1.9.1)$$

e o respectivo momento por:

$$\begin{aligned} \mathbf{p} &= \mathbf{p}(t; t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \\ &\stackrel{\text{def}}{=} L_{\dot{\mathbf{x}}}(t, \mathbf{x}, \dot{\mathbf{x}})|_{\mathbf{x}=\mathbf{x}(t; t_0, \mathbf{x}_0; t_1, \mathbf{x}_1)} \end{aligned} \quad (1.9.2)$$

Em particular:

$$\mathbf{x}_0 = \mathbf{x}(t_0; t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \quad \text{e} \quad \mathbf{x}_1 = \mathbf{x}(t_1; t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \quad (1.9.3)$$

Notamos ainda por:

$$\begin{aligned} \dot{\mathbf{x}}_0 &= \dot{\mathbf{x}}_0(t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \\ &= \left. \frac{\partial}{\partial t} \right|_{t=t_0} \mathbf{x}(t; t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \\ \dot{\mathbf{x}}_1 &= \dot{\mathbf{x}}_1(t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \\ &= \left. \frac{\partial}{\partial t} \right|_{t=t_1} \mathbf{x}(t; t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \end{aligned} \quad (1.9.4)$$

as velocidades nas extremidades da extremal, e ainda por:

$$\begin{aligned}\mathbf{p}_0 &= \mathbf{p}_0(t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \\ &= \mathbf{p}(t_0; t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \\ \mathbf{p}_1 &= \mathbf{p}_1(t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \\ &= \mathbf{p}(t_1; t_0, \mathbf{x}_0; t_1, \mathbf{x}_1)\end{aligned}\tag{1.9.5}$$

os momentos nessas mesmas extremidades. Note que  $\dot{\mathbf{x}}_0, \dot{\mathbf{x}}_1, \mathbf{p}_0$  e  $\mathbf{p}_1$  são consideradas como funções das  $2n + 2$  variáveis  $(t_0, \mathbf{x}_0; t_1, \mathbf{x}_1)$ . A estas funções chamamos as **funções de campo** do feixe de extremais considerado.

Se agora substituirmos as funções (1.9.1) e (1.9.2) no funcional canónico de acção  $\mathcal{S} = \mathcal{S}_H$ , definido em (1.7.2), obtemos a seguinte função das  $2n + 2$  variáveis  $(t_0, \mathbf{x}_0; t_1, \mathbf{x}_1)$ :

$$\begin{aligned}\mathcal{S} &= \mathcal{S}(t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \\ &= \int_{t_0}^{t_1} \mathbf{p} d\mathbf{x} - H dt\end{aligned}\tag{1.9.6}$$

a que se chama a **distância geodésica**<sup>3</sup> entre os pontos  $P_0 = (t_0, \mathbf{x}_0)$  e  $P_1 = (t_1, \mathbf{x}_1)$ .

Vamos mostrar que as derivadas parciais de  $\mathcal{S} = \mathcal{S}(t_0, \mathbf{x}_0; t_1, \mathbf{x}_1)$  são dadas por:

$$\begin{aligned}\mathcal{S}_{t_0} &= H_0 = H(t_0, \mathbf{x}_0, \mathbf{p}_0) \\ &= \mathbf{p}_0 \dot{\mathbf{x}}_0 - L(t_0, \mathbf{x}_0, \dot{\mathbf{x}}_0) \\ \mathcal{S}_{\mathbf{x}_0} &= -\mathbf{p}_0 \\ \mathcal{S}_{t_1} &= -H_1 = -H(t_1, \mathbf{x}_1, \mathbf{p}_1) \\ &= L(t_1, \mathbf{x}_1, \dot{\mathbf{x}}_1) - \mathbf{p}_1 \dot{\mathbf{x}}_1 \\ \mathcal{S}_{\mathbf{x}_1} &= \mathbf{p}_1\end{aligned}\tag{1.9.7}$$

É claro que nestas fórmulas  $\dot{\mathbf{x}}_0, \dot{\mathbf{x}}_1, \mathbf{p}_0$  e  $\mathbf{p}_1$  são as funções do feixe (consideradas como funções das  $2n + 2$  variáveis  $(t_0, \mathbf{x}_0; t_1, \mathbf{x}_1)$ ), definidas em (1.9.4) e (1.9.16).

Para provar isto, consideremos uma curva  $\alpha \mapsto (t_0(\alpha), \mathbf{x}_0(\alpha); t_1(\alpha), \mathbf{x}_1(\alpha))$  tal que:

$$\begin{aligned}(t_0(0), \mathbf{x}_0(0); t_1(0), \mathbf{x}_1(0)) &= (t_0, \mathbf{x}_0; t_1, \mathbf{x}_1) \\ (t'_0(0), \mathbf{x}'_0(0); t'_1(0), \mathbf{x}'_1(0)) &= (\delta t_0, \delta \mathbf{x}_0; \delta t_1, \delta \mathbf{x}_1)\end{aligned}$$

e calculemos a derivada da função:

$$\begin{aligned}\mathcal{S}(\alpha) &\stackrel{\text{def}}{=} \mathcal{S}(t_0(\alpha), \mathbf{x}_0(\alpha); t_1(\alpha), \mathbf{x}_1(\alpha)) \\ &= \int_{t_0(\alpha)}^{t_1(\alpha)} [\mathbf{p}(t; t_0(\alpha), \mathbf{x}_0(\alpha); t_1(\alpha), \mathbf{x}_1(\alpha)) \dot{\mathbf{x}}(t; t_0(\alpha), \mathbf{x}_0(\alpha); t_1(\alpha), \mathbf{x}_1(\alpha)) \\ &\quad - H(t, \mathbf{x}(t; t_0(\alpha), \mathbf{x}_0(\alpha); t_1(\alpha), \mathbf{x}_1(\alpha)), \mathbf{p}(t; t_0(\alpha), \mathbf{x}_0(\alpha); t_1(\alpha), \mathbf{x}_1(\alpha))] dt\end{aligned}$$

<sup>3</sup>Outros nomes frequentes para  $\mathcal{S} = \mathcal{S}(t_0, \mathbf{x}_0; t_1, \mathbf{x}_1)$  são a **iconal** (do grego *Eikón*=imagem), em óptica geométrica, e a **função característica pontual** de Hamilton, em mecânica.

em ordem  $\alpha$ , para  $\alpha = 0$ . Lembrando que o último integral é também dado por:

$$\int_{t_0(\alpha)}^{t_1(\alpha)} L dt$$

avaliado ao longo da extremal que une  $(t_0(\alpha), \mathbf{x}_0(\alpha))$  a  $(t_1(\alpha), \mathbf{x}_1(\alpha))$ , podemos aplicar a teoria exposta na secção 1.4, nomeadamente a fórmula (1.4.5), para concluir que:

$$\begin{aligned} \mathcal{S}'(0) &= (\mathbf{p}d\mathbf{x} - Hdt)|_{t_0}^{t_1} \\ &= \mathbf{p}_1\delta\mathbf{x}_1 - \mathbf{p}_0\delta\mathbf{x}_0 - H_1\delta t_1 + H_0\delta t_0 \end{aligned} \quad (1.9.8)$$

o que demonstra as fórmulas (1.9.7). Destas fórmulas deduzimos ainda que a iconal  $\mathcal{S}$  satisfaz as duas equações de tipo Hamilton-Jacobi:

$$\begin{aligned} \mathcal{S}_{t_1} + H(t_1, \mathbf{x}_1, \mathcal{S}_{\mathbf{x}_1}) &= 0 \\ \mathcal{S}_{t_0} - H(t_0, \mathbf{x}_0, -\mathcal{S}_{\mathbf{x}_0}) &= 0 \end{aligned} \quad (1.9.9)$$

♣.

Suponhamos agora que temos uma hipersuperfície regular  $\Sigma \subset \mathbb{R}^{n+1}$ , dada pela equação:

$$\Sigma : \quad \Phi(t, \mathbf{x}) = 0 \quad (1.9.10)$$

Dado um ponto  $P = (t, \mathbf{x}) \in \mathbb{R}^{n+1}$  determinemos um ponto  $P_0 = (t_0, \mathbf{x}_0) \in \Sigma$  tal que a distância geodésica  $\mathcal{S}(P_0, P)$  seja estacionária, sob variações do ponto  $P_0$  em  $\Sigma$  (ver a figura 1.9). Este problema foi discutido na secção 1.5. Vimos então que uma extremal que une  $P_0$  a  $P$  terá de verificar a condição de transversalidade seguinte:

$$\mathbf{p}_0\delta\mathbf{x}_0 - H_0\delta t_0 = 0 \quad (1.9.11)$$

para todo o vector tangente  $(\delta t_0, \delta\mathbf{x}_0) \in T_{(t_0, \mathbf{x}_0)}\Sigma$ , isto é, para todo o vector  $(\delta t_0, \delta\mathbf{x}_0)$  tal que  $0 = d\Phi_{(t_0, \mathbf{x}_0)}(\delta t_0, \delta\mathbf{x}_0) = (\Phi_t dt + \Phi_{\mathbf{x}} d\mathbf{x})|_{(t_0, \mathbf{x}_0)}(\delta t_0, \delta\mathbf{x}_0)$ , ou ainda:

$$\Phi_t(t_0, \mathbf{x}_0)\delta t_0 + \Phi_{\mathbf{x}}(t_0, \mathbf{x}_0)\delta\mathbf{x}_0 = 0 \quad (1.9.12)$$

Notemos que as duas condições (1.9.11) e (1.9.12), significam que o vector  $(-H_0, \mathbf{p}_0)$  é colinear com o gradiente de  $\Phi$  em  $(t_0, \mathbf{x}_0)$ :

$$-H_0 = \lambda \Phi_t(t_0, \mathbf{x}_0) \quad (1.9.13)$$

$$\mathbf{p}_0 = \lambda \Phi_{\mathbf{x}}(t_0, \mathbf{x}_0) \quad (1.9.14)$$

ou, por outras palavras, o vector  $(-H_0, \mathbf{p}_0)$  é ortogonal a  $\Sigma$  no ponto  $P_0 = (t_0, \mathbf{x}_0)$ . Uma extremal que verifique as duas condições (1.9.13) e (1.9.14) diz-se, por isso, uma **extremal ortogonal a  $\Sigma$** , em  $P_0 = (t_0, \mathbf{x}_0)$ . Se, a cada ponto de  $\Sigma$ , associarmos uma extremal ortogonal a  $\Sigma$  nesse ponto, obtemos uma família a  $n$  parâmetros de extremais.

Vamos agora supôr que, num certo aberto  $\mathcal{U}$  de  $\mathbb{R}^{n+1}$ , que contem  $\Sigma$ , se verifica a seguinte propriedade: dado um ponto qualquer  $P = (t, \mathbf{x}) \in \mathcal{U}$ , existe uma e uma só

Figure 1.9: .

extremal que passa em  $P$  e que é ortogonal a  $\Sigma$  (ver a figura 1.9). Diz-se então que, em  $\mathcal{U}$ , está definido um **feixe de extremais ortogonais** a  $\Sigma$ .

Portanto, para cada ponto  $P = (t, \mathbf{x}) \in \mathcal{U}$ , podemos determinar um único ponto:

$$P_0 = (t_0(t, \mathbf{x}), \mathbf{x}_0(t, \mathbf{x})) \in \Sigma \quad (1.9.15)$$

para o qual a distância geodésica  $\mathcal{S}(P_0, P)$  é estacionária, sob variações do ponto  $P_0$  em  $\Sigma$ , e, por isso, as funções de campo:

$$\begin{aligned} \dot{\mathbf{x}}(t, \mathbf{x}) &= \dot{\mathbf{x}}(t_0(t, \mathbf{x}), \mathbf{x}_0(t, \mathbf{x}); t, \mathbf{x}) \\ \mathbf{p}(t, \mathbf{x}) &= \mathbf{p}(t_0(t, \mathbf{x}), \mathbf{x}_0(t, \mathbf{x}); t, \mathbf{x}) \end{aligned} \quad (1.9.16)$$

são agora funções bem definidas em  $\mathcal{U}$ . À função  $\mathcal{S}_\Sigma : \mathcal{U} \rightarrow \mathbb{R}$ , definida por:

$$\mathcal{S}_\Sigma(t, \mathbf{x}) \stackrel{\text{def}}{=} \mathcal{S}(t_0(t, \mathbf{x}), \mathbf{x}_0(t, \mathbf{x}); t, \mathbf{x}) \quad (1.9.17)$$

chama-se a **distância geodésica** relativamente à hipersuperfície  $\Sigma$ . A família de hipersuperfícies:

$$\Sigma_c \stackrel{\text{def}}{=} \{(t, \mathbf{x}) \in \mathcal{U} : \mathcal{S}_\Sigma(t, \mathbf{x}) \equiv c\} \quad (1.9.18)$$

diz-se a família de **superfícies paralelas** do problema variacional. Nestas condições, podemos enunciar a seguinte proposição:

- **♣ Proposição 1.4** ... Dado um tal feixe de extremais em  $\mathcal{U} \subseteq \mathbb{R}^{n+1}$ , ortogonais à hipersuperfície  $\Sigma$ , a distância geodésica  $\mathcal{S}_\Sigma$ , relativa a  $\Sigma$ , satisfaz as equações:

$$\begin{aligned} (\mathcal{S}_\Sigma)_t &= -H(t, \mathbf{x}, \mathbf{p}(t, \mathbf{x})) \\ (\mathcal{S}_\Sigma)_{\mathbf{x}} &= \mathbf{p}(t, \mathbf{x}) \end{aligned} \quad (1.9.19)$$

e portanto  $\mathcal{S}_\Sigma$ , satisfaz a equação de Hamilton-Jacobi:

$$(\mathcal{S}_\Sigma)_t + H(t, \mathbf{x}, (\mathcal{S}_\Sigma)_{\mathbf{x}}) = 0 \quad (1.9.20)$$

**Dem.:** Derivando (1.9.17), respectivamente em ordem a  $t$  e a  $\mathbf{x}$ , e usando a regra da cadeia e ainda as relações (1.9.7), obtemos:

$$\begin{aligned}
 (\mathcal{S}_\Sigma)_t &= \mathcal{S}_{t_0} \frac{\partial t_0}{\partial t} + \mathcal{S}_{\mathbf{x}_0} \frac{\partial \mathbf{x}_0}{\partial t} + \mathcal{S}_t \\
 &= H_0 \frac{\partial t_0}{\partial t} - \mathbf{p}_0 \frac{\partial \mathbf{x}_0}{\partial t} - H \\
 (\mathcal{S}_\Sigma)_{\mathbf{x}} &= \mathcal{S}_{t_0} \frac{\partial t_0}{\partial \mathbf{x}} + \mathcal{S}_{\mathbf{x}_0} \frac{\partial \mathbf{x}_0}{\partial \mathbf{x}} + \mathcal{S}_{\mathbf{x}} \\
 &= H_0 \frac{\partial t_0}{\partial \mathbf{x}} - \mathbf{p}_0 \frac{\partial \mathbf{x}_0}{\partial \mathbf{x}} + \mathbf{p}
 \end{aligned} \tag{1.9.21}$$

Mas:

$$\Phi(t_0(t, \mathbf{x}), \mathbf{x}_0(t, \mathbf{x})) \equiv 0$$

e derivando esta equação em ordem a  $t$  e a  $\mathbf{x}$ , respectivamente, deduzimos que:

$$\begin{aligned}
 0 &= \Phi_t(t_0, \mathbf{x}_0) \frac{\partial t_0}{\partial t} + \Phi_{\mathbf{x}}(t_0, \mathbf{x}_0) \frac{\partial \mathbf{x}_0}{\partial t} \\
 &= -H_0 \frac{\partial t_0}{\partial t} + \mathbf{p}_0 \frac{\partial \mathbf{x}_0}{\partial t} \\
 0 &= \Phi_t(t_0, \mathbf{x}_0) \frac{\partial t_0}{\partial \mathbf{x}} + \Phi_{\mathbf{x}}(t_0, \mathbf{x}_0) \frac{\partial \mathbf{x}_0}{\partial \mathbf{x}} \\
 &= -H_0 \frac{\partial t_0}{\partial \mathbf{x}} + \mathbf{p}_0 \frac{\partial \mathbf{x}_0}{\partial \mathbf{x}}
 \end{aligned} \tag{1.9.22}$$

onde usamos as condições de ortogonalidade (1.9.13) e (1.9.14). Substituindo estas relações em (1.9.21), obtemos finalmente:

$$(\mathcal{S}_\Sigma)_t = -H, \quad \text{e} \quad (\mathcal{S}_\Sigma)_{\mathbf{x}} = \mathbf{p}$$

como se pretendia. ♣.

Reciprocamente, temos a seguinte proposição:

- **♣ Proposição 1.5** ... Se  $\mathcal{S} = \mathcal{S}(t, \mathbf{x})$  é uma solução da equação de Hamilton-Jacobi:

$$\mathcal{S}_t + H(t, \mathbf{x}, \mathcal{S}_{\mathbf{x}}) = 0 \tag{1.9.23}$$

de classe  $C^2$ , então existe um feixe de extremais, ortogonais a todas as hipersuperfícies  $\mathcal{S}(t, \mathbf{x}) \equiv c$  (constante), e  $\mathcal{S}$  é a distância geodésica relativa à hipersuperfície  $\mathcal{S} \equiv 0$ .

**Dem.:** Seja  $\mathcal{S} = \mathcal{S}(t, \mathbf{x})$  uma solução da equação de Hamilton-Jacobi (1.9.23), e definamos o campo de momentos  $\mathbf{p} = \mathbf{p}(t, \mathbf{x})$ , através de:

$$\mathbf{p}(t, \mathbf{x}) \stackrel{\text{def}}{=} \mathcal{S}_{\mathbf{x}}(t, \mathbf{x}) \tag{1.9.24}$$

Consideremos agora o seguinte sistema (não autónomo) de  $n$  ODE's de primeira ordem:

$$\dot{\mathbf{x}} = H_{\mathbf{p}}(t, \mathbf{x}, \mathbf{p}(t, \mathbf{x})) \quad (1.9.25)$$

Este sistema define uma família a  $n$  parâmetros de curvas. Ao longo de cada uma dessas curvas, temos que  $\mathbf{p}(t, \mathbf{x}(t)) = \mathcal{S}_{\mathbf{x}}(t, \mathbf{x}(t))$  e, derivando em ordem a  $t$ , obtemos:

$$\begin{aligned} \dot{\mathbf{p}} &= \mathcal{S}_{\mathbf{x}t} + \mathcal{S}_{\mathbf{x}\mathbf{x}}\dot{\mathbf{x}}, & \text{isto é} & & \dot{p}_i &= \mathcal{S}_{x^i t} + \mathcal{S}_{x^i x^j}\dot{x}^j \\ &= \mathcal{S}_{\mathbf{x}t} + \mathcal{S}_{\mathbf{x}\mathbf{x}}H_{\mathbf{p}} \end{aligned} \quad (1.9.26)$$

onde usamos (1.9.25) na última igualdade. Por outro lado, derivando (1.9.23) em ordem a  $\mathbf{x}$ , obtemos:

$$\mathcal{S}_{\mathbf{x}t} + H_{\mathbf{x}} + H_{\mathbf{p}}\mathcal{S}_{\mathbf{x}\mathbf{x}} = 0, \quad \text{isto é} \quad \mathcal{S}_{x^i t} + H_{x^i} + H_{p_j}\mathcal{S}_{x^j x^i} = 0$$

Comparando (1.9.26) com esta última igualdade, deduzimos que:

$$\dot{\mathbf{p}} = -H_{\mathbf{x}} \quad (1.9.27)$$

Concluindo - as equações canónicas:

$$\begin{cases} \dot{\mathbf{x}} = H_{\mathbf{p}} \\ \dot{\mathbf{p}} = -H_{\mathbf{x}} \end{cases}$$

caracterizam a referida família de curvas, como uma família a  $n$  parâmetros de extremais. Resta provar que estas extremais são ortogonais a todas as hipersuperfícies  $\mathcal{S}(t, \mathbf{x}) \equiv c$  (constante), isto é, que elas verificam as condições de ortogonalidade (1.9.13) e (1.9.14), com  $\Phi = \mathcal{S}$ :

$$-H = \lambda \mathcal{S}_t(t, \mathbf{x}), \quad \text{e} \quad \mathbf{p} = \lambda \mathcal{S}_{\mathbf{x}}(t, \mathbf{x})$$

o que é óbvio, com  $\lambda = 1$ , atendendo à equação de Hamilton-Jacobi e à definição de  $\mathbf{p}$ , respectivamente, (1.9.23) e (1.9.24).

♣.

Notemos que a proposição 1.4 diz, grosso modo, que, se soubermos determinar as extremais ortogonais a uma certa hipersuperfície  $\Sigma = \{\Phi = 0\}$ , então sabemos determinar soluções da equação de Hamilton-Jacobi, dependendo de uma certa função inicial  $\Phi$ . Reciprocamente, a proposição 1.5 diz, grosso modo, que se soubermos determinar soluções da equação de Hamilton-Jacobi, então sabemos determinar uma família a  $n$  parâmetros de soluções das equações canónicas (isto é, sabemos determinar uma família a  $n$  parâmetros de extremais).

♣ **Exemplo 1.15** ... Consideremos o funcional:

$$I[x(\cdot)] = \int \left( 2t\dot{x} - \frac{1}{2}\dot{x}^2 \right) dt$$

Como  $L(t, x, \dot{x}) = 2t\dot{x} - \frac{1}{2}\dot{x}^2$  e  $L_{\dot{x}\dot{x}} = -1 \neq 0$ , o Lagrangeano é hiperregular. A transformada de Legendre é:

$$(t, x, \dot{x}) \longmapsto (t, x, p = L_{\dot{x}} = 2t - \dot{x})$$

o Hamiltoniano é igual a:

$$\begin{aligned} H(t, x, p) &= p\dot{x} - L|_{\dot{x}=2t-p} = p(2t-p) - 2t(2t-p) + \frac{1}{2}(2t-p)^2 \\ &= -\frac{1}{2}p^2 + 2tp - 2t^2 \end{aligned}$$

e as equações canónicas são:

$$\begin{cases} \dot{x} = H_p = -p + 2t \\ \dot{p} = -H_x = 0 \end{cases} \quad \begin{cases} x(t) = t^2 - at + b \\ p(t) \equiv a \end{cases}, \quad a, b \text{ constantes}$$

A família a 2 parâmetros de extremais  $x(t) = t^2 - at + b$  constitui um feixe de extremais em  $\mathbb{R}_{tx}^2$  - dados dois pontos  $P_0(t_0, x_0), P_1(t_1, x_1)$ , com  $t_0 < t_1$ , existe uma e uma só extremal que os une.

Consideremos o feixe de extremais ortogonais a  $\Sigma = \{\Phi(t, x) = t = 0\}$  (o eixo dos  $x$ 's). A condição de ortogonalidade num ponto  $(0, x_0) \in \Sigma$  é que, para todo o vector tangente  $(\delta t_0, \delta x_0) \in T_{(0, x_0)}\Sigma$ , se verifique  $p_0\delta x_0 - H\delta t_0 = 0$ . Como  $\delta t_0 = 0$ , isto traduz-se em que:

$$p_0\delta x_0 = 0, \quad \forall \delta x_0 \in \mathbb{R} \quad \Rightarrow \quad p_0 = 2 \cdot 0 - \dot{x}_0 = -a = 0$$

e portanto o feixe de extremais ortogonais a  $\Sigma = \{\Phi(t, x) = t = 0\}$  é a família a 1 parâmetro  $x(t) = t^2 + x_0$ . Dado um ponto qualquer  $P = (t, x)$  em  $\mathbb{R}^2$ , com  $t \neq 0$ , a única extremal que passa em  $P$  e é ortogonal a  $\Sigma$  é:

$$x(\tau) = \tau^2 - t^2 + x, \quad 0 \leq \tau \leq t$$

que intersecta  $\Sigma$  no ponto  $P_0 = (0, x_0 = x - t^2)$ . A distância geodésica  $\mathcal{S}_\Sigma$ , relativa a  $\Sigma$ , é dada por:

$$\mathcal{S}_\Sigma(t, x) = \int_0^t L(\tau, x(\tau), \dot{x}(\tau)) d\tau = \int_0^t 2\tau^2 d\tau = \frac{2}{3}t^3$$

(não depende de  $x$ !). As superfícies paralelas a  $\Sigma$  são as rectas verticais  $t \equiv \text{constante}$ .

Note que, neste exemplo, o campo de momentos é sempre nulo:  $p(t, x) = 2t\dot{x}(t) = 2t - 2t \equiv 0$ .

A equação de Hamilton-Jacobi para uma função  $S = S(t, x)$ , é:

$$S_t + H(t, x, S_x) = 0, \quad \text{isto é} \quad S_t - \frac{1}{2}S_x^2 + 2tS_x - 2t^2 = 0$$

e é imediato que  $\mathcal{S}_\Sigma$  é solução. Aliás qualquer função do tipo  $\mathcal{S}_\Sigma + \text{constante}$ , é também solução.



## 1.10 O integral invariante de Hilbert

Consideremos mais uma vez uma solução  $\mathcal{S} = \mathcal{S}(t, \mathbf{x})$  da equação de Hamilton-Jacobi  $\mathcal{S}_t + H(t, \mathbf{x}, \mathcal{S}_\mathbf{x}) = 0$ , o feixe de extremais ortogonais a  $\mathcal{S} \equiv 0$ , construído na proposição 1.5, e as funções de feixe correspondentes (definidas num certo aberto  $\mathcal{U} \subseteq \mathbb{R}^{n+1}$  que contem  $\{\mathcal{S} = 0\}$ ). Em particular, o campo de momentos  $\mathbf{p} = \mathbf{p}(t, \mathbf{x})$ , é dado por:

$$\mathbf{p}(t, \mathbf{x}) = \mathcal{S}_\mathbf{x}(t, \mathbf{x})$$

Se  $P_0 = (t_0, \mathbf{x}_0)$  e  $P = (t, \mathbf{x})$  são dois pontos, em  $\mathcal{U}$ , o integral:

$$\mathcal{S}(P) - \mathcal{S}(P_0) = \int_{P_0}^P d\mathcal{S}$$

óbviamente que não depende da curva que une  $P_0$  a  $P$ . Em particular, calculando-o ao longo de uma curva qualquer  $\tau \mapsto (\tau, \phi(\tau))$ ,  $t_0 \leq \tau \leq t$ , que une  $P_0$  a  $P$ , obtemos sucessivamente o seguinte:

$$\begin{aligned} \mathcal{S}(P) - \mathcal{S}(P_0) &= \int_{P_0}^P d\mathcal{S} \\ &= \int_{t_0}^t \frac{d}{d\tau} \mathcal{S}(\tau, \phi(\tau)) d\tau \\ &= \int_{t_0}^t \left[ \mathcal{S}_\mathbf{x}(\tau, \phi(\tau)) \dot{\phi}(\tau) + \mathcal{S}_\tau(\tau, \phi(\tau)) \right] d\tau \\ &= \int_{t_0}^t \left[ \mathbf{p}(\tau, \phi(\tau)) \dot{\phi}(\tau) - H(\tau, \phi(\tau), \mathbf{p}(\tau, \phi(\tau))) \right] d\tau \\ &= \int_{P_0}^P \mathbf{p}d\mathbf{x} - Hdt \end{aligned} \tag{1.10.1}$$

onde usamos o facto de  $\mathcal{S} = \mathcal{S}(t, \mathbf{x})$  ser solução da equação de Hamilton-Jacobi, e a definição  $\mathbf{p} = \mathcal{S}_\mathbf{x}$ . Obtemos, desta forma, uma representação integral para a distância geodésica entre os pontos  $P_0 = (t_0, \mathbf{x}_0)$  e  $P = (t, \mathbf{x})$ . O integral (1.10.1) chama-se o **integral invariante de Hilbert**.

Reciprocamente, suponhamos que, num certo aberto  $\mathcal{U} \subseteq \mathbb{R}^{n+1}$ , se define uma função vectorial  $\mathbf{p} : \mathcal{U} \rightarrow \mathbb{R}^n$ :

$$\mathbf{p} = \mathbf{p}(t, \mathbf{x}) \tag{1.10.2}$$

para a qual o integral (1.10.1) não depende da curva  $\mathbf{x}(\tau)$  que une  $P_0 = (t_0, \mathbf{x}_0)$  a  $P = (t, \mathbf{x})$  em  $\mathcal{U}$ . Então  $\mathbf{p} = \mathbf{p}(t, \mathbf{x})$  é o campo de momentos de um certo feixe de extremais em  $\mathcal{U}$ .

De facto, a invariância do integral de Hilbert permite definir uma função  $\mathcal{S} : \mathcal{U} \rightarrow \mathbb{R}$  através de:

$$\mathcal{S}(P) = \mathcal{S}(P_0) + \int_{P_0}^P \mathbf{p}d\mathbf{x} - Hdt \tag{1.10.3}$$

É um facto geral que, como este integral não depende da curva que une  $P_0$  a  $P$ , devemos ter  $d\mathcal{S} = \mathbf{p}d\mathbf{x} - Hdt$ , isto é:

$$\mathcal{S}_{\mathbf{x}} = \mathbf{p}, \quad \text{e} \quad \mathcal{S}_t = -H(t, \mathbf{x}, \mathbf{p}) \quad (1.10.4)$$

e portanto  $\mathcal{S}$  é solução da equação de Hamilton-Jacobi. Pela proposição 1.5, qualquer solução da equação de Hamilton-Jacobi é a distância geodésica de um certo feixe de extremais (soluções das equações canónicas), e  $\mathbf{p} = \mathcal{S}_{\mathbf{x}}$ .

Resumindo toda a discussão anterior, podemos afirmar que a equação de Hamilton-Jacobi, a construção de feixes extremais e da correspondente distância geodésica, e a invariância do integral de Hilbert, são três aspectos equivalentes de uma mesma situação.

♣ **Exemplo 1.16** ... Consideremos de novo o funcional:

$$I[y(\cdot)] = \frac{1}{2} \int (y'^2 - y^2) dx$$

estudado no exemplo 1.10. Aí vimos que a equação de Euler-Lagrange é  $y'' + y = 0$ , cuja solução geral é:

$$y(x) = a \cos x + b \sin x$$

Consideremos o feixe de extremais ortogonais a  $\Sigma = \{\Phi(x, y) = x = 0\}$  (o eixo dos  $y$ 's). A condição de ortogonalidade num ponto  $(0, y_0) \in \Sigma$  é que, para todo o vector tangente  $(\delta x_0, \delta y_0) \in T_{(0, y_0)}\Sigma$ , se verifique  $p_0 \delta y_0 - H \delta x_0 = 0$ . Como  $\delta x_0 = 0$ , isto traduz-se em que:

$$p_0 \delta y_0 = 0, \quad \forall \delta y_0 \in \mathbb{R} \quad \Rightarrow \quad p_0 = L_{y'}(0, y(0), y'(0)) = y'(0) = b = 0$$

e portanto o feixe de extremais ortogonais a  $\Sigma$  é a família a 1 parâmetro  $y(x) = a \cos x$ . Dado um ponto qualquer  $P = (x, y)$  em  $\mathcal{U} \subset \mathbb{R}^2$ , com  $x \neq 0$ , a única extremal que passa em  $P$  e é ortogonal a  $\Sigma$  é:

$$y(\tau) = \frac{y}{\cos x} \cos \tau, \quad 0 \leq \tau \leq x$$

que intersecta  $\Sigma$  no ponto  $P_0 = (0, y_0 = y/\cos x)$ . A distância geodésica  $\mathcal{S}_{\Sigma}$ , relativa a  $\Sigma$ , é dada por:

$$\begin{aligned} \mathcal{S}_{\Sigma}(x, y) &= \int_0^x L(\tau, y(\tau), y'(\tau)) d\tau \\ &= -\frac{1}{2} \frac{y^2}{\cos^2 x} \int_0^x \cos 2\tau d\tau \\ &= -\frac{1}{2} y^2 \operatorname{tg} x \end{aligned} \quad (1.10.5)$$

As linhas paralelas a  $\Sigma$  são as curvas  $y^2 \operatorname{tg} x \equiv c$  (constante). O campo de momentos é:

$$p(x, y) = L_{y'}(x, y(x), y'(x)) = y'(x) = -y \operatorname{tg} x = S_y(x, y)$$

e o Hamiltoniano é:

$$H(x, y, p(x, y)) = \frac{1}{2}(p(x, y)^2 + y^2) = \frac{y^2}{2 \cos^2 x} = -S_x(x, y)$$

Portanto  $S$  satisfaz a equação de Hamilton-Jacobi:

$$S_x + H(x, y, S_y(x, y)) = 0$$

Por outro lado:

$$dS = S_y dy + S_x dx = p dy - H dx$$

e o integral invariante de Hilbert é:

$$\int p dy - H dx = \int -y \operatorname{tg} x dy - \frac{y^2}{2 \cos^2 x} dx$$

de tal forma que a distância geodésica entre dois pontos  $P_0 = (x_0, y_0), P_1 = (x_1, y_1) \in \mathcal{U}$ , com  $x_0 < x_1$ , é dada por:

$$S(P_1) - S(P_0) = \int_{P_0}^{P_1} -y \operatorname{tg} x dy - \frac{y^2}{2 \cos^2 x} dx$$

onde o integral é calculado ao longo de uma qualquer curva que una  $P_0$  a  $P_1$ , em  $\mathcal{U}$ .

Note finalmente que o feixe de extremais  $y(x) = a \cos x$ , é a solução geral da ODE de primeira ordem:

$$y' = H_p(x, y, p(x, y)) = p(x, y) = -y \operatorname{tg} x$$

♣.

## 1.11 Transformações canónicas. Método de Hamilton

Para motivar o conceito de transformação canónica e respectivas funções geradoras, vamos, nesta secção, descrever a abordagem de Hamilton, baseada em argumentos muito simples de óptica geométrica.

Suponhamos então que temos um sistema óptico, onde se propagam os raios de luz. Estes partem de um plano  $\mathcal{P}_0$  (o plano objecto), atravessam o sistema óptico, e atingem um outro plano  $\mathcal{P}_1$  (o plano imagem). Supômos que esse sistema óptico admite um eixo a que, como é tradicional, chamamos o eixo dos  $t$ 's, e que os raios luminosos  $\boldsymbol{\rho}$  se projectam difeomòrficamente sobre esse eixo, de tal forma que podem ser parametrizados na forma  $t \mapsto \boldsymbol{\rho}(t) = (t, x(t), y(t)) = (t, \mathbf{x}(t))$ . Os planos objecto  $\mathcal{P}_0$  e imagem  $\mathcal{P}_1$ , correspondem a  $t = t_0$  e a  $t = t_1$ , respectivamente. Adoptamos coordenadas  $(\mathbf{x}_0, \dot{\mathbf{x}}_0) = (x_0, y_0, \dot{x}_0, \dot{y}_0)$  para  $T\mathcal{P}_0$  e  $(\mathbf{x}_1, \dot{\mathbf{x}}_1) = (x_1, y_1, \dot{x}_1, \dot{y}_1)$  para  $T\mathcal{P}_1$  e ainda coordenadas canónicas  $(\mathbf{x}_0, \mathbf{p}_0) = (x_0, y_0, p_0, q_0)$  para  $T^*\mathcal{P}_0$  e  $(\mathbf{x}_1, \mathbf{p}_1) = (x_1, y_1, p_1, q_1)$  para  $T^*\mathcal{P}_1$ .

Figure 1.10: .

O nosso objectivo é mostrar que, após aplicar transformações de Legendre, os raios luminosos definem uma transformação  $F_{t_0, t_1} : T^*\mathcal{P}_0 \rightarrow T^*\mathcal{P}_1$ , que é canónica, isto é, preserva as estruturas simplécticas (figura 1.10).

Qualquer raio  $\rho$ , dado por  $t \mapsto \rho(t) = (t, \mathbf{x}(t))$ ,  $t \in [t_0, t_1]$ , tem um comprimento óptico dado por:

$$\mathcal{V}[\rho] = \int_{t_0}^{t_1} L(t, \mathbf{x}, \dot{\mathbf{x}}) dt \quad (1.11.1)$$

onde:

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

Façamos uma transformação de Legendre:

$$\mathbf{p} = L_{\dot{\mathbf{x}}} = \frac{n \dot{\mathbf{x}}}{\sqrt{1 + \dot{\mathbf{x}}^2}} \quad (1.11.2)$$

Portanto:

$$H = \mathbf{p}\dot{\mathbf{x}} - L = \frac{-n}{\sqrt{1 + \dot{\mathbf{x}}^2}} = -\sqrt{n^2 - \mathbf{p}^2} \quad (1.11.3)$$

onde  $\mathbf{p} = (p, q)$  e  $\mathbf{p}^2 = \mathbf{p} \cdot \mathbf{p} = p^2 + q^2$ . Note que, com  $\mathbf{x} = (x, y)$ :

$$\begin{aligned} \cos a &= \frac{\dot{x}}{\sqrt{1 + \dot{x}^2 + \dot{y}^2}} \\ \cos b &= \frac{\dot{y}}{\sqrt{1 + \dot{x}^2 + \dot{y}^2}} \\ \cos c &= \frac{1}{\sqrt{1 + \dot{x}^2 + \dot{y}^2}} \end{aligned} \quad (1.11.4)$$

são os cossenos directores do raio de luz, relativamente aos eixos  $x, y$  e  $t$ . A:

$$\begin{aligned} p &= n \cos a \\ q &= n \cos b \end{aligned} \quad (1.11.5)$$

chamamos por isso os **cossenos directores ópticos** do raio. Note que o Hamiltoniano é  $H = -n \cos c$ . Podemos ainda exprimir o comprimento óptico  $\mathcal{V}[\rho]$  em termos de  $\mathbf{x}(t)$

e  $\mathbf{p}(t)$ . De facto, usando as fórmulas anteriores, podemos deduzir que:

$$\begin{aligned} \mathcal{V}[\boldsymbol{\rho}] &= \int_{t_0}^{t_1} [\mathbf{p}(t)\dot{\mathbf{x}}(t) - H(t, \mathbf{x}(t), \mathbf{p}(t))] dt \\ &= \int_{\boldsymbol{\rho}} \mathbf{p}d\mathbf{x} - H \end{aligned} \quad (1.11.6)$$

O raio incidente é completamente determinado pelo seu ponto  $\mathbf{x}_0 = (x_0, y_0)$  de intersecção com o plano objecto  $\mathcal{P}_0$ , e pelos seus cossenos directores ópticos  $\mathbf{p}_0 = (n_0 \cos a_0, n_0 \cos b_0)$ . A intersecção  $\mathbf{x}_1 = (x_1, y_1)$ , do raio refractado com o plano imagem  $\mathcal{P}_1$ , e os respectivos cossenos directores ópticos  $\mathbf{p}_1$ , são funções dos dados iniciais em  $\mathcal{P}_0$ :

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{x}(t_0, t_1; \mathbf{x}_0, \mathbf{p}_0) \\ \mathbf{p}_1 &= \mathbf{p}(t_0, t_1; \mathbf{x}_0, \mathbf{p}_0) \end{aligned} \quad (1.11.7)$$

e são estas funções que definem a transformação canónica:

$$F = F_{t_0, t_1} : T^*\mathcal{P}_0 \rightarrow T^*\mathcal{P}_1 \quad (1.11.8)$$

Mantemos a dependência explícita de  $t_0$  e  $t_1$ , isto é, dos planos objecto e imagem, pois essa dependência é importante no projecto de sistemas ópticos.

A imagem óptica do plano objecto  $\mathcal{P}_0 = \{t = t_0\}$  no plano imagem  $\mathcal{P}_1 = \{t = t_1\}$  diz-se **perfeita** se a primeira equação em (1.11.7) se reduz a:

$$\mathbf{x}_1 = c\mathbf{x}_0 \quad (1.11.9)$$

onde  $c$  é uma constante igual para todos os pontos  $\mathbf{x}_0 \in \mathcal{P}_0$  e todas as direcções  $\mathbf{p}_0$ . O problema principal da concepção de instrumentos ópticos é o de determinar uma distribuição dos meios ópticos  $n = n(x, y, t)$  tais que (1.11.9) seja verificada. Os desvios:

$$\Delta\mathbf{x}_1 = \mathbf{x}_1 - c\mathbf{x}_0 \quad (1.11.10)$$

dizem-se as **aberrações** do sistema óptico (para um tratamento detalhado deste assunto ver [11]).

O resultado principal da abordagem de Hamilton ao problema anterior, é que é possível reduzir o problema de calcular as (quatro) funções (1.11.7), que definem a transformação canónica (1.11.8), ao problema de calcular apenas uma! função. Esta função dir-se-á por isso a **função geradora** da transformação canónica  $F = F_{t_0, t_1} : T^*\mathcal{P}_0 \rightarrow T^*\mathcal{P}_1$ . As funções (1.11.7) são então calculadas a partir desta função geradora, usando apenas as operações de derivação e eliminação! Vejamos como.

Em primeiro lugar, as funções (1.11.7) são calculadas directamente a partir das equações canónicas. Mais detalhadamente - suponhamos que:

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{x}(t_0, t; \mathbf{x}_0, \mathbf{p}_0) \\ \mathbf{p}(t) &= \mathbf{p}(t_0, t; \mathbf{x}_0, \mathbf{p}_0) \end{aligned} \quad (1.11.11)$$

é a solução das equações canónicas:

$$\begin{cases} \dot{\mathbf{x}} = H_{\mathbf{p}} \\ \dot{\mathbf{p}} = -H_{\mathbf{x}} \end{cases} \quad (1.11.12)$$

que, para  $t = t_0$ , tem os valores iniciais  $\mathbf{x}_0, \mathbf{p}_0$ , isto é:

$$\begin{aligned} \mathbf{x}_0 &= \mathbf{x}(t_0) = \mathbf{x}(t_0, t_0; \mathbf{x}_0, \mathbf{p}_0) \\ \mathbf{p}_0 &= \mathbf{p}(t_0) = \mathbf{p}(t_0, t_0; \mathbf{x}_0, \mathbf{p}_0) \end{aligned} \quad (1.11.13)$$

No plano imagem  $\mathcal{P}_1$  as funções (1.11.11) tomam os valores:

$$\mathbf{x}_1 = \mathbf{x}(t_1) = \mathbf{x}(t_0, t_1; \mathbf{x}_0, \mathbf{p}_0) \quad (1.11.14)$$

$$\mathbf{p}_1 = \mathbf{p}(t_1) = \mathbf{p}(t_0, t_1; \mathbf{x}_0, \mathbf{p}_0) \quad (1.11.15)$$

que são os valores que definem as funções (1.11.7) e, portanto, a transformação canónica (1.11.8).

Suponhamos agora que o Jacobiano  $\frac{\partial(\mathbf{x}_1)}{\partial(\mathbf{p}_0)}$ , da aplicação (1.11.14), é não nulo:

$$\frac{\partial(\mathbf{x}_1)}{\partial(\mathbf{p}_0)} = \frac{\partial(x_1, y_1)}{\partial(p_0, q_0)} \neq 0 \quad (1.11.16)$$

Neste caso, podemos calcular (localmente)  $\mathbf{p}_0$ , a partir das equações (1.11.14), como função de  $(t_0, t_1; \mathbf{x}_0, \mathbf{x}_1)$ . Substituindo então a função  $\mathbf{p}_0$ , assim obtida, em (1.11.11), obtemos 4 funções  $\mathbf{x}, \mathbf{p}$  das variáveis  $(t_0, t_1; \mathbf{x}_0, \mathbf{x}_1)$  e ainda  $t$ , que notamos por:

$$\begin{aligned} \mathbf{x} &= \mathbf{x}(t_0, t_1; \mathbf{x}_0, \mathbf{x}_1; t) \\ \mathbf{p} &= \mathbf{p}(t_0, t_1; \mathbf{x}_0, \mathbf{x}_1; t) \end{aligned} \quad (1.11.17)$$

que representam o raio de luz  $\rho(P_0, P_1)$  que passa nos pontos  $P_0 = (t_0, \mathbf{x}_0) \in \mathcal{P}_0$  e  $P_1 = (t_1, \mathbf{x}_1) \in \mathcal{P}_1$ . Introduzamos agora estas funções no integral (1.11.6), para o comprimento óptico, para obter uma função  $\mathcal{V} = \mathcal{V}(t_0, t_1; \mathbf{x}_0, \mathbf{x}_1)$ :

$$\mathcal{V}(t_0, t_1; \mathbf{x}_0, \mathbf{x}_1) = \int_{\rho(P_0, P_1)} \mathbf{p} d\mathbf{x} - H \quad (1.11.18)$$

que dá a distância óptica entre os pontos  $P_0$  e  $P_1$ .

A esta função dá-se o nome de **iconal** ou de **função característica pontual de Hamilton**. Trata-se exactamente da função distância geodésica, que foi tratada numa secção anterior, e que aí foi notada por  $\mathcal{S}$ . Como vimos nessa secção, a diferencial  $d\mathcal{V}$  é dada por:

$$d\mathcal{V} = \mathbf{p}_1 d\mathbf{x}_1 - \mathbf{p}_0 d\mathbf{x}_0 - H_1 dt_1 + H_0 dt_0 \quad (1.11.19)$$

onde os coeficientes  $\mathbf{p}_0, \mathbf{p}_1$  são as funções de  $(t_0, t_1; \mathbf{x}_0, \mathbf{x}_1)$ , obtidas fazendo, respectivamente,  $t = t_0$  e  $t = t_1$ , em (1.11.17), e:

$$\begin{aligned} H_0 &= H(t_0, \mathbf{x}_0, \mathbf{p}_0) \\ H_1 &= H(t_1, \mathbf{x}_1, \mathbf{p}_1) \end{aligned} \quad (1.11.20)$$

Sendo assim, deduzimos de (1.11.19) as seguintes fórmulas fundamentais:

$$\boxed{\mathbf{p}_0 = -\mathcal{V}_{\mathbf{x}_0}, \quad \mathbf{p}_1 = \mathcal{V}_{\mathbf{x}_1}, \quad H_0 = \mathcal{V}_{t_0}, \quad H_1 = -\mathcal{V}_{t_1}} \quad (1.11.21)$$

As duas primeiras equações são equivalentes às equações (1.11.14) e (1.11.15), e demonstram o facto mencionado no início desta secção de que as funções (1.11.14) e (1.11.15) (que definem a transformação canónica  $F$ ), podem ser calculadas a partir de uma única - a função geradora  $\mathcal{V}(t_0, t_1; \mathbf{x}_0, \mathbf{x}_1)$  - por derivação e eliminação. Quanto às duas últimas equações, elas representam duas PDE's obtidas introduzindo as derivadas parciais de  $\mathcal{V}$ , dadas por (1.11.21), em  $H_0$  e  $H_1$ . Obtemos então duas equações de tipo Hamilton-Jacobi:

$$\mathcal{V}_{t_0} - H(t_0, \mathbf{x}_0; -\mathcal{V}_{\mathbf{x}_0}) = 0$$

$$\mathcal{V}_{t_1} + H(t_1, \mathbf{x}_1; \mathcal{V}_{\mathbf{x}_1}) = 0$$

### Interpretações das fórmulas (1.11.21):

1. Vamos fixar o valor de  $t_0$  e adoptar as notações mais usuais:

$$\mathbf{a} = \mathbf{x}_0, \quad \mathbf{b} = \mathbf{p}_0, \quad t = t_1, \quad \mathbf{x} = \mathbf{x}_1, \quad \mathbf{p} = \mathbf{p}_1$$

Definamos ainda:

$$S(t, \mathbf{x}, \mathbf{a}) = \mathcal{V}(t_0, t_1 = t; \mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1 = \mathbf{x}) \quad (1.11.22)$$

Então  $S$  é uma solução da equação de Hamilton-Jacobi:

$$S_t + H(t, \mathbf{x}, S_{\mathbf{x}}) = 0$$

que depende de  $n$  parâmetros  $\mathbf{a} = (a^1, \dots, a^n)$ . As duas primeiras equações em (1.11.21) têm agora a forma:

$$\begin{aligned} \mathbf{b} &= -S_{\mathbf{a}}(t, \mathbf{x}, \mathbf{a}) \\ \mathbf{p} &= S_{\mathbf{x}}(t, \mathbf{x}, \mathbf{a}) \end{aligned}$$

De acordo com a interpretação óptica, acima discutida, estas fórmulas definem, para cada  $t$  fixo, uma transformação canónica  $(\mathbf{a}, \mathbf{b}) \mapsto (\mathbf{x}, \mathbf{p})$ , gerada pela iconal  $S$ , e que é obtida usando a primeira equação em (1.11.23), para exprimir  $\mathbf{x}$  como função de  $\mathbf{a}$  e  $\mathbf{b}$ :

$$\mathbf{b} = -S_{\mathbf{a}}(\mathbf{x}, \mathbf{a}) \quad \longrightarrow \quad \mathbf{x} = \mathbf{x}(\mathbf{a}, \mathbf{b})$$

e depois substituindo este valor de  $\mathbf{x}$  na segunda equação em (1.11.23), para obter  $\mathbf{p}$  como função de  $\mathbf{a}$  e  $\mathbf{b}$ :

$$\mathbf{p} = \mathbf{p}(\mathbf{a}, \mathbf{b}) = S_{\mathbf{x}}(\mathbf{x}, \mathbf{a})|_{\mathbf{x}=\mathbf{x}(\mathbf{a}, \mathbf{b})}$$

(omitimos a dependência de  $t_0$  e  $t$ ).

2. Vamos agora dar uma segunda interpretação das fórmulas (1.11.23). Desta vez fixamos o plano objecto  $\mathcal{P}_0$ , mas variamos o plano imagem  $\mathcal{P}(t)$ . Como vimos, as fórmulas (1.11.23) estabelecem uma correspondência entre os elementos ópticos  $(t_0, \mathbf{a}, \mathbf{b})$  do plano objecto e os elementos ópticos  $(t, \mathbf{x}, \mathbf{p})$  do plano imagem  $\mathcal{P}(t)$ . Fixando  $\mathbf{a}$  e  $\mathbf{b}$ , e variando  $t$ , obtemos um raio:

$$\rho(t) = \left( t, \mathbf{x}(t, \mathbf{a}, \mathbf{b}), \mathbf{p}(t, \mathbf{a}, \mathbf{b}) \right)$$

que satisfaz as equações canónicas:

$$\dot{\mathbf{x}} = H_{\mathbf{p}}(t, \mathbf{x}, \mathbf{p}), \quad \dot{\mathbf{p}} = -H_{\mathbf{x}}(t, \mathbf{x}, \mathbf{p})$$

e que depende dos  $2n$  parâmetros  $\mathbf{a}, \mathbf{b}$ . Analiticamente, esta solução das equações canónicas obtem-se do seguinte modo - primeiro usamos a equação:

$$S_{\mathbf{a}}(t, \mathbf{x}, \mathbf{a}) = -\mathbf{b}$$

para exprimir  $\mathbf{x}$  como uma função  $\mathbf{x}(t, \mathbf{a}, \mathbf{b})$ , de  $t$  e dos  $2n$  parâmetros  $\mathbf{a}, \mathbf{b}$ . Em seguida inserimos esta função em:

$$\mathbf{p} = S_{\mathbf{x}}(t, \mathbf{x}, \mathbf{a})|_{\mathbf{x}=\mathbf{x}(t, \mathbf{a}, \mathbf{b})}$$

para obter  $\mathbf{p}$  como uma função  $\mathbf{p}(t, \mathbf{a}, \mathbf{b})$ , de  $t$  e ainda dos  $2n$  parâmetros  $\mathbf{a}, \mathbf{b}$ . Esta é, essencialmente, a ideia base do método de Jacobi para integrar as equações canónicas, que vamos discutir na próxima secção.

## 1.12 Método de Jacobi para integrar as equações canónicas de Hamilton. Teorema de Jacobi

Como vimos na prova da proposição 1.5, uma solução  $S(t, \mathbf{x})$  da equação de Hamilton-Jacobi determina uma família a  $n$  parâmetros de soluções das equações canónicas de Hamilton, obtida resolvendo o sistema (não autónomo) de  $n$  ODE's de primeira ordem (1.9.25):

$$\dot{\mathbf{x}} = H_{\mathbf{p}}(t, \mathbf{x}, \mathbf{p}(t, \mathbf{x})) \tag{1.12.1}$$

(uma solução para cada "parâmetro"  $\mathbf{x} \in \mathbb{R}^n$ , como condição inicial para esse sistema de ODE's). Se  $t \mapsto (t, \mathbf{x}(t))$  é uma solução do sistema (1.12.1), então a curva:

$$t \mapsto \left( t, \mathbf{x}(t), \mathbf{p}(t) = S_{\mathbf{x}}(t, \mathbf{x}(t)) \right)$$

é solução das equações canónicas, como se viu antes.

Portanto uma família a  $n$  parâmetros  $S(t, \mathbf{x}; \mathbf{a})$  de soluções da equação de Hamilton-Jacobi, que dependa "essencialmente" dos  $n$  parâmetros  $\mathbf{a} = (a^1, \dots, a^n) \in \mathbb{R}^n$ , no sentido em que:

$$\det [S_{\mathbf{x}\mathbf{a}}] = \det [S_{x^i a^k}] \neq 0 \tag{1.12.2}$$

deve, em princípio, determinar toda a família a  $2n$  parâmetros de soluções das equações de Hamilton.

Uma família a  $n$  parâmetros  $S(t, \mathbf{x}; \mathbf{a})$  de soluções da equação de Hamilton-Jacobi, que satisfaça a condição (1.12.2), diz-se um **integral completo** da equação de Hamilton-Jacobi. O método de Jacobi para obter a solução geral  $\mathbf{x}(t, \mathbf{a}, \mathbf{b})$ ,  $\mathbf{p}(t, \mathbf{a}, \mathbf{b})$  das equações canônicas:

$$\begin{cases} \dot{\mathbf{x}} = H_{\mathbf{p}}(t, \mathbf{x}, \mathbf{p}) \\ \dot{\mathbf{p}} = -H_{\mathbf{x}}(t, \mathbf{x}, \mathbf{p}) \end{cases} \quad (1.12.3)$$

que dependa dos  $2n$  parâmetros  $\mathbf{a} = (a^1, \dots, a^n)$ ,  $\mathbf{b} = (b_1, \dots, b_n)$ , a partir de uma solução completa  $S(t, \mathbf{x}, \mathbf{a})$  da equação de Hamilton-Jacobi, consiste nos passos seguintes:

1. Primeiro resolvem-se as  $n$  equações implícitas seguintes:

$$S_{a^i}(t, \mathbf{x}; \mathbf{a}) = -b_i, \quad i = 1, \dots, n \quad (1.12.4)$$

em ordem a  $\mathbf{x}$ , para obter uma solução do tipo:

$$\mathbf{x} = \mathbf{x}(t, \mathbf{a}, \mathbf{b}) \quad (1.12.5)$$

2. Em seguida complementamos esta função  $\mathbf{x} = \mathbf{x}(t, \mathbf{a}, \mathbf{b})$ , por uma outra função  $\mathbf{p} = \mathbf{p}(t, \mathbf{a}, \mathbf{b})$ , definida como habitualmente por:

$$\mathbf{p} = \mathbf{p}(t, \mathbf{a}, \mathbf{b}) = S_{\mathbf{x}}(t, \mathbf{x}(t, \mathbf{a}, \mathbf{b}), \mathbf{a}) \quad (1.12.6)$$

Antes de demonstrar porque é que este método funciona, vejamos um exemplo concreto:

♣ **Exemplo 1.17 (Oscilador harmónico)** ... Como já vimos, o oscilador harmónico é descrito pelo seguinte Hamiltoniano:

$$H(x, p) = \frac{\omega}{2}(x^2 + p^2) \quad (1.12.7)$$

As equações canônicas são:

$$\begin{cases} \dot{x} = H_p = \omega p \\ \dot{p} = -H_x = -\omega x \end{cases} \quad (1.12.8)$$

cuja solução geral é:

$$\begin{cases} x(t) = A \cos(\omega t + b) \\ p(t) = -A \sin(\omega t + b) \end{cases}, \quad A \text{ e } b \text{ constantes} \quad (1.12.9)$$

A equação de Hamilton-Jacobi para uma função  $S(t, x)$ , correspondente ao Hamiltoniano  $H = \frac{\omega}{2}(x^2 + p^2)$  é:

$$S_t + \frac{\omega}{2}(x^2 + S_x^2) = 0 \quad (1.12.10)$$

Vamos tentar encontrar um integral completo  $S(t, x, a)$ , pelo método de separação de variáveis, com o *ansatz* da forma:

$$S(t, x) = f(t) + \psi(x)$$

Com este  $S$ , (1.12.10) escreve-se na forma:

$$\dot{f}(t) + \frac{\omega}{2} (x^2 + \psi'(x)^2) = 0$$

o que implica que:

$$\dot{f}(t) = -\frac{\omega}{2} (x^2 + \psi'(x)^2) = \text{constante} = -a$$

e portanto:

$$\dot{f}(t) = -a, \quad \psi'(x) = \pm \sqrt{\frac{2a}{\omega} - x^2}$$

Concluimos portanto que:

$$S(t, x, a) = -at + \int_0^x \sqrt{\frac{2a}{\omega} - \tau^2} d\tau \quad (1.12.11)$$

é uma solução da equação de Hamilton-Jacobi (1.6.23), que depende de um parâmetro  $a$ . Não é necessário calcular este integral, já que o nosso objectivo é calcular:

$$S_a(t, x, a) = -b$$

o que é equivalente a:

$$-t + \frac{1}{\omega} \int_0^x \frac{d\tau}{\sqrt{\frac{2a}{\omega} - \tau^2}} = -b$$

Pondo  $\beta = -\omega b - \arccos A$ , com  $A = \sqrt{\frac{2a}{\omega}}$ , vem que:

$$-\arccos(x/A) = \omega t + \beta$$

donde se deduz a solução usual:

$$x(t) = A \cos(\omega t + \beta)$$

De:

$$p = S_x(t, x, a) = \sqrt{A^2 - x^2}$$

deduzimos ainda que:

$$p(t) = \pm A \sin(\omega t + \beta)$$

e como  $x(t), p(t)$  satisfazem as equações canônicas:

$$\dot{x} = H_p = \omega p, \quad \dot{p} = -H_x = -\omega x$$

obtemos:

$$p(t) = -A \sin(\omega t + \beta)$$

Além disso:

$$a = -S_t = H(x, S_x)$$

e, para  $x = x(t)$ , vem que:

$$a = H(x(t), p(t))$$

isto é,  $a$  é o nível de energia da trajectória:

$$x(t) = A \cos(\omega t + \beta), \quad p(t) = -A \sin(\omega t + \beta)$$

Finalmente, (1.12.11) dá que:

$$S(t, x, a) = \frac{A^2}{2} \arcsin \frac{x}{A} + \frac{1}{2} x \sqrt{A^2 - x^2} - at, \quad A = \sqrt{\frac{2a}{\omega}}$$



Antes de enunciar e demonstrar o Teorema de Jacobi, vejamos uma proposição prévia:

- **♣ Proposição 1.6** ... Seja  $S = S(t, \mathbf{x}, \mathbf{a})$  uma solução da equação de Hamilton-Jacobi, que depende dos parâmetros  $\mathbf{a} = (a^i) \in \mathbb{R}^m$ . Então cada derivada  $S_{a^i}$ ,  $i = 1, \dots, m$ , é um integral primeiro das equações canônicas, isto é,  $S_{a^i} \equiv \text{constante}$ , ao longo de cada extremal.

**Dem.:** Temos que mostrar que  $\frac{d}{dt} S_{a^i} = 0$ , ao longo de cada extremal. Em primeiro lugar, temos que:

$$S_t(t, \mathbf{x}, \mathbf{a}) + H(t, \mathbf{x}, S_{\mathbf{x}}(t, \mathbf{x}, \mathbf{a})) = 0$$

já que, por hipótese,  $S = S(t, \mathbf{x}, \mathbf{a})$  é uma solução da equação de Hamilton-Jacobi. Derivando esta relação em ordem a  $a^i$ , vem que:

$$S_{a^i t} + H_{\mathbf{p}} S_{a^i \mathbf{x}} = 0 \tag{1.12.12}$$

Vem então que:

$$\begin{aligned} \frac{d}{dt} S_{a^i}(t, \mathbf{x}(t), \mathbf{a}) &= S_{ta^i} + S_{\mathbf{x}a^i} \dot{\mathbf{x}} \\ &= -H_{\mathbf{p}} S_{a^i \mathbf{x}} + S_{\mathbf{x}a^i} \dot{\mathbf{x}}, && \text{por (1.12.12)} \\ &= (\dot{\mathbf{x}} - H_{\mathbf{p}}) S_{a^i \mathbf{x}} \\ &= 0 \end{aligned}$$

porque,  $\dot{\mathbf{x}} = H_{\mathbf{p}}$ , ao longo de cada extremal.



- ♣ **Teorema 1.3 (Teorema de Jacobi)** ... Seja  $S(t, \mathbf{x}; \mathbf{a})$  uma solução completa da equação de Hamilton-Jacobi  $S_t + H(t, \mathbf{x}, S_{\mathbf{x}}) = 0$ , e suponhamos que  $\mathbf{x} = \mathbf{x}(t, \mathbf{a}, \mathbf{b})$  e  $\mathbf{p} = \mathbf{p}(t, \mathbf{a}, \mathbf{b})$  são funções que satisfazem as equações seguintes:

$$\begin{aligned} S_{\mathbf{a}}(t, \mathbf{x}(t, \mathbf{a}, \mathbf{b}), \mathbf{a}) &= -\mathbf{b} \\ \mathbf{p}(t, \mathbf{a}, \mathbf{b}) &= S_{\mathbf{x}}(t, \mathbf{x}(t, \mathbf{a}, \mathbf{b}), \mathbf{a}) \end{aligned} \quad (1.12.13)$$

Então  $t \mapsto (\mathbf{x}(t, \mathbf{a}, \mathbf{b}), \mathbf{p}(t, \mathbf{a}, \mathbf{b}))$  é uma solução das equações canônicas (1.12.3), que depende dos  $2n$  parâmetros  $\mathbf{a}$  e  $\mathbf{b}$ .

**Dem.:** Derivando a primeira equação em (1.12.13), em ordem a  $t$ , obtemos:

$$\begin{aligned} \mathbf{0} &= S_{t\mathbf{a}} + S_{\mathbf{x}\mathbf{a}}\dot{\mathbf{x}} \\ &= S_{\mathbf{x}\mathbf{a}}(\dot{\mathbf{x}} - H_{\mathbf{p}}) \end{aligned} \quad (1.12.14)$$

onde na última igualdade usamos (1.12.14). Como estamos a supôr que  $\det S_{\mathbf{x}\mathbf{a}} \neq 0$ , a igualdade (1.12.14) permite deduzir que:

$$\dot{\mathbf{x}} = H_{\mathbf{p}}$$

que é a primeira equação canônica. Para obter a segunda, derivamos a segunda equação em (1.12.13), em ordem a  $t$ :

$$\begin{aligned} \dot{\mathbf{p}} &= S_{t\mathbf{x}} + S_{\mathbf{x}\mathbf{x}}\dot{\mathbf{x}} \\ &= S_{t\mathbf{x}} + S_{\mathbf{x}\mathbf{x}}H_{\mathbf{p}} \\ &= -H_{\mathbf{x}} \end{aligned} \quad (1.12.15)$$

onde na última igualdade usamos  $S_{\mathbf{x}t} + H_{\mathbf{x}} + H_{\mathbf{p}}S_{\mathbf{x}\mathbf{x}} = \mathbf{0}$ , que se obtém, derivando em ordem a  $\mathbf{x}$ , a equação de Hamilton-Jacobi.

♣

- ♣ **Exemplo 1.18** ... Consideremos o funcional:

$$J[y(\cdot)] = \int xy\sqrt{y'} dx$$

Como  $L(x, y, y') = xy\sqrt{y'}$ , vem que  $p = L_{y'} = \frac{xy}{2\sqrt{y'}} \Rightarrow y' = \frac{x^2y^2}{4p^2}$ , e o Hamiltoniano é dado por:

$$H(x, y, p) = py' - L = -\frac{x^2y^2}{4p}$$

A equação de Hamilton-Jacobi, para uma função  $S = S(x, y)$ , é:

$$S_x - \frac{x^2y^2}{4S_y} = 0$$

Separando variáveis:

$$S(x, y) = f(x) + g(y)$$

vem que:

$$f'(x) - \frac{x^2 y^2}{4g'(y)} = 0 \quad \Rightarrow \quad \frac{f'(x)}{x^2} = \frac{y^2}{4g'(y)} \equiv a$$

e portanto:

$$f(x) = \frac{ax^3}{3} \quad \text{e} \quad g(y) = \frac{y^3}{12a}$$

O integral completo é pois:

$$S(x, y, a) = \frac{ax^3}{3} + \frac{y^3}{12a}$$

Agora pômos:

$$S_a = \frac{x^3}{3} - \frac{y^3}{12a^2} = -b$$

e resolvemos isto em ordem a  $y = y(x, a, b)$ , para obter uma solução geral, em forma implícita, do tipo:

$$y^3 = cx^3 + d$$

♣.

♣ **Exemplo 1.19** ... Consideremos o funcional:

$$J[x(\cdot)] = \int \sqrt{t^2 + x^2} \sqrt{1 + \dot{x}^2} dt$$

Este funcional é de tipo óptico, com índice de refração  $n(t, x) = \sqrt{t^2 + x^2}$ . O Hamiltoniano é dado por (2.2.32), isto é:

$$H(t, x, p) = -\sqrt{t^2 + x^2 - p^2}$$

A equação de Hamilton-Jacobi, para uma função  $S = S(t, x)$ , é:

$$S_t - \sqrt{t^2 + x^2} - S_x^2 = 0$$

que é do tipo  $\|\nabla S\|^2 = n^2$ , isto é:

$$S_t^2 + S_x^2 = t^2 + x^2$$

Separando variáveis, vem que:

$$S_t^2 - t^2 \equiv -a \quad \text{e} \quad S_x^2 - x^2 \equiv a$$

e portanto:

$$S_t = \sqrt{t^2 - a} \quad \text{e} \quad S_x = \sqrt{x^2 + a}$$

O integral completo é pois:

$$S(t, x, a) = \int \sqrt{t^2 - a} dt + \int \sqrt{x^2 + a} dx$$

♣.

## 1.13 Invariantes integrais

### 1.13.1 Preliminares de álgebra linear

Seja  $\omega : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$  uma 2-forma exterior (i.e., uma forma bilinear alternada) num espaço vectorial real de dimensão finita. O núcleo  $\ker \omega$  é constituído por todos os vectores  $\mathbf{u} \in \mathcal{V}$  que são  $\omega$ -ortogonais a todos os vectores de  $\mathcal{V}$ :

$$\ker \omega = \{\mathbf{u} \in \mathcal{V} : \omega(\mathbf{u}, \mathbf{v}) = 0, \forall \mathbf{v} \in \mathcal{V}\}$$

É claro que  $\ker \omega$  coincide com o núcleo da aplicação linear:

$$\begin{aligned} \Phi_\omega : \mathcal{V} &\longrightarrow \mathcal{V}^* \\ \mathbf{u} &\longmapsto i_{\mathbf{u}}\omega \end{aligned} \quad (1.13.1)$$

onde  $i_{\mathbf{u}}\omega$  é a forma linear definida por  $(i_{\mathbf{u}}\omega)(\mathbf{v}) = \omega(\mathbf{u}, \mathbf{v})$ . Portanto  $\ker \omega$  é um subespaço de  $\mathcal{V}$ . Quando  $\ker \omega = \{\mathbf{0}\}$ , a aplicação  $\Phi_\omega$  é um isomorfismo e a forma  $\omega$  diz-se **não degenerada** ou uma **forma simpléctica** em  $\mathcal{V}$ . Um **espaço vectorial simpléctico** é um par  $(\mathcal{V}, \omega)$ , onde  $\mathcal{V}$  é um espaço vectorial<sup>4</sup> e  $\omega$  uma 2-forma exterior não degenerada. Em breve veremos que a dimensão de  $\mathcal{V}$  tem que ser par.

Seja  $\{\mathbf{e}_i\}$  uma base para  $\mathcal{V}$  e  $\{\mathbf{e}^i\}$  a correspondente base dual para  $\mathcal{V}^*$ , de tal forma que  $\mathbf{e}^i(\mathbf{e}_j) = \delta_j^i$ . Então a matriz de  $\Phi_\omega$  relativamente a essas bases é a matriz de Gram  $[\omega_{ij}]$ , onde  $\omega_{ij} = \omega(\mathbf{e}_i, \mathbf{e}_j)$ . De facto  $\Phi_\omega(\mathbf{e}_i) = \omega_{ij}\mathbf{e}^j$ . O **rank** da forma  $\omega$ ,  $\text{rank } \omega$ , é, por definição, o rank da matriz  $[\omega_{ij}]$ , isto é, a dimensão de  $\text{im } \Phi_\omega = \Phi_\omega(\mathcal{V})$ :

$$\text{rank } \omega = \dim \Phi_\omega(\mathcal{V}) \quad (1.13.2)$$

Portanto  $\omega$  é não degenerada sse  $\text{rank } \omega$  é máximo, isto é, igual à  $\dim \mathcal{V}^* = \dim \mathcal{V}$ .

♣ **Proposição 1.7** ... *Seja  $\mathcal{V}$  um espaço vectorial real de dimensão  $N$ , e  $\omega$  uma 2-forma exterior em  $\mathcal{V}$ . Então  $\text{rank } \omega = 2n$  para algum inteiro  $n$  e existe uma base ordenada  $\{\mathbf{e}_i\}_{i=1, \dots, N}$  para  $\mathcal{V}$ , com base dual  $\{\mathbf{e}^i\}_{i=1, \dots, N}$  para  $\mathcal{V}^*$ , tal que:*

$$\omega = \sum_{i=1}^n \mathbf{e}^i \wedge \mathbf{e}^{n+i} \quad (1.13.3)$$

ou, de forma equivalente, relativamente à qual a matriz de Gram de  $\omega$  é:

$$\begin{bmatrix} \mathbf{0} & \mathbf{I}_n & \mathbf{0} \\ -\mathbf{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (1.13.4)$$

**Dem.:** Escolhamos vectores não nulos  $\mathbf{e}_1, \mathbf{e}_{n+1} \in \mathcal{V}$ , tais que  $\omega(\mathbf{e}_1, \mathbf{e}_{n+1}) \neq 0$ , o que é possível se  $\omega \neq 0$ . Multiplicando  $\mathbf{e}_1$  por um escalar podemos supôr que  $\omega(\mathbf{e}_1, \mathbf{e}_{n+1}) = 1$ .

<sup>4</sup>Neste curso, apenas consideramos espaços vectoriais reais de dimensão finita.

Como  $\omega(\mathbf{e}_1, \mathbf{e}_1) = 0 = \omega(\mathbf{e}_{n+1}, \mathbf{e}_{n+1})$ , a matriz de  $\omega$ , restrita ao plano  $\mathcal{S} = \text{span}\{\mathbf{e}_1, \mathbf{e}_{n+1}\}$  é:

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Consideremos agora o  $\omega$ -ortogonal  $\mathcal{S}^\perp$ , de  $\mathcal{S}$ :

$$\mathcal{S}^\perp = \{\mathbf{v} \in \mathcal{V} : \omega(\mathbf{v}, \mathbf{s}) = 0, \quad \forall \mathbf{s} \in \mathcal{S}\}$$

É claro que  $\mathcal{S}^\perp \cap \mathcal{S} = \{\mathbf{0}\}$ . Por outro lado,  $\mathcal{S}^\perp + \mathcal{S} = \mathcal{V}$ . De facto, se  $\mathbf{v} \in \mathcal{V}$ , então:

$$\mathbf{v} - \omega(\mathbf{v}, \mathbf{e}_{n+1})\mathbf{e}_1 + \omega(\mathbf{v}, \mathbf{e}_1)\mathbf{e}_{n+1} \in \mathcal{S}^\perp$$

Portanto  $\mathcal{S}^\perp \oplus \mathcal{S} = \mathcal{V}$ . Podemos então repetir o processo para  $\mathcal{S}^\perp$ , escolhendo  $\mathbf{e}_2$  e  $\mathbf{e}_{n+2}$  tais que  $\omega(\mathbf{e}_2, \mathbf{e}_{n+2}) = 1$ , e continuar assim indutivamente.

♣.

Se  $\mathbf{v} \in \mathcal{V}$  se escreve como combinação linear na base  $\{\mathbf{e}_i\}$ , referida no teorema:

$$\mathbf{v} = x^1\mathbf{e}_1 + \cdots + x^n\mathbf{e}_n + y^1\mathbf{e}_{n+1} + \cdots + y^n\mathbf{e}_{2n} + z^{2n+1}\mathbf{e}_{2n+1} + \cdots + z^N\mathbf{e}_N$$

e análogamente para  $\mathbf{v}' \in \mathcal{V}$ , então:

$$\omega(\mathbf{v}, \mathbf{v}') = \sum_{i=1}^n (x^i y'^i - y^i x'^i) \quad (1.13.5)$$

Quando  $\omega$  é não degenerada,  $\text{rank } \omega$  é máximo e igual à dimensão de  $\mathcal{V}$ , e portanto neste caso  $\dim \mathcal{V}$  tem que ser par. Em particular, a dimensão de um espaço vectorial simpléctico é par.

Quando  $\dim \mathcal{V}$  é ímpar então  $\ker \omega \neq \{\mathbf{0}\}$ . Um vector não nulo  $\mathbf{u} \in \ker \omega$  diz-se um **vector característico** da forma  $\omega$ . Neste caso  $\omega$  diz-se **não singular**, se  $\dim \ker \omega$  é a menor possível, isto é, igual a 1. Portanto, se  $\omega$  é uma 2-forma exterior não singular num espaço vectorial de dimensão ímpar, todos os seus vectores característicos pertencem a uma recta, unívocamente determinada pela forma  $\omega$ , a que chamamos a **recta característica** de  $\omega$ .

### 1.13.2 Subvariedades integrais. Teorema de Darboux

♣ **Definição 1.1** ... Seja  $M$  uma variedade de dimensão  $m$  e  $\omega \in \Omega^2(M)$  uma 2-forma fechada. Uma subvariedade  $\varphi : N \rightarrow M$  diz-se uma **subvariedade integral** de  $\omega$  se  $\varphi^*\omega = 0$ .

♣ **Teorema 1.4** ... Seja  $M$  uma variedade de dimensão  $m$ ,  $\omega \in \Omega^2(M)$  uma 2-forma fechada e  $\varphi : N \rightarrow M$  uma subvariedade integral de  $\omega$ .

Seja  $X \in \mathfrak{X}(M)$  uma campo de vectores característicos, i.e.,  $X_p \in \ker \omega_p, \forall p \in M$ , transversal a  $\varphi(N) \subset M$ , isto é,  $X_p \notin T_p\varphi(N), \forall p$ . Defina-se, para  $t$  suficientemente pequeno,  $t \in I$ , uma aplicação:

$$\Phi(t, n) \stackrel{\text{def}}{=} Fl_t^X(\varphi(n)), \quad (t, n) \in I \times N$$

Então  $\Phi : I \times N \rightarrow M$  é ainda uma subvariedade integral de  $\omega$ .

**Dem.:** Notemos em primeiro lugar que a derivada de Lie  $L_X\omega = 0$ . De facto:

$$L_X\omega = X\lrcorner d\omega + d(X\lrcorner\omega) = 0 \quad (1.13.6)$$

uma vez que  $X$  é característico ( $X\lrcorner\omega = 0$ ) e  $\omega$  é fechada ( $d\omega = 0$ ).

Como o teorema é local, vamos supôr que escolhemos coordenadas locais  $(x^i)$  para  $M$  tais que  $X = \frac{\partial}{\partial x^1}$  e  $\omega = \omega_{ij}dx^i \wedge dx^j$ . Por (1.13.6) vem então que:

$$0 = L_X\omega = (X\omega_{ij})dx^i \wedge dx^j = \frac{\partial\omega_{ij}}{\partial x^1}dx^i \wedge dx^j \Rightarrow \frac{\partial\omega_{ij}}{\partial x^1} = 0$$

e os  $\omega_{ij}$  não dependem de  $x^1$ . Por outro lado:

$$0 = X\lrcorner\omega = \omega_{1j}dx^j \Rightarrow \omega_{1j} = 0$$

Portanto:

$$\omega = \sum_{i,j \geq 2} \omega_{ij}(x^2, \dots, x^m)dx^i \wedge dx^j$$

Mas, por construção, e atendendo a que  $X = \frac{\partial}{\partial x^1}$ , os  $x^i$ , para  $i \geq 2$  são constantes ao longo das curvas integrais de  $X$ , isto é:

$$x^i \circ \Phi = x^i \circ \varphi, \quad i \geq 2$$

Portanto:

$$\Phi^*\omega = \varphi^*\omega = 0$$

♣.

♣ **Definição 1.2** ... Seja  $M$  uma variedade de dimensão  $m$  e  $\omega \in \Omega^2(M)$  uma 2-forma fechada de rank constante. Define-se o **fibrado Êcaracterístico**  $\mathfrak{C}_\omega$  de  $\omega$  através de:

$$\mathfrak{C}_\omega \stackrel{\text{def}}{=} \{X \in \mathfrak{X}(M) : X\lrcorner\omega = 0\} \quad (1.13.7)$$

O fibrado Êcaracterístico  $\mathfrak{C}_\omega$ , é portanto igual ao  $\ker \omega_p$ , em cada ponto  $p \in M$ . Um **campo de vectores característico** é um campo  $X \in \mathfrak{X}(M)$  tal que  $X\lrcorner\omega = 0$ , isto é, tal que  $X(p) \in \mathfrak{C}_\omega(p) = \ker \omega_p$ ,  $\forall p$ .

Se o rank de  $\omega$  é constante, então  $\mathfrak{C}_\omega$  é um subfibrado de  $TM$ , ou, por outras palavras, é uma distribuição em  $M$ , chamada a **distribuição Êcaracterística** de  $\omega$ .

♣ **Proposição 1.8** ... A distribuição Êcaracterística  $\mathfrak{C}_\omega$  de uma 2-forma fechada  $\omega \in \Omega^2(M)$  de rank constante, é involutiva:

$$[\mathfrak{C}_\omega, \mathfrak{C}_\omega] \subset \mathfrak{C}_\omega$$

**Dem.:** Como  $\omega$  tem rank constante, a dimensão de  $\mathfrak{C}_\omega(p)$  é também constante,  $\forall p \in M$ , e portanto  $\mathfrak{C}_\omega$  é uma distribuição. Sejam  $X, Y$  campos de vetores característicos. Então:

$$\begin{aligned} [X, Y] \lrcorner \omega &= L_X(Y \lrcorner \omega) - Y \lrcorner (L_X \omega) \\ &= Y \lrcorner (L_X \omega) \\ &= Y \lrcorner (X \lrcorner d\omega + d(X \lrcorner \omega)) = 0 \end{aligned} \quad (1.13.8)$$

e portanto  $[X, Y]$  é também um campo de vetores característico.

♣.

♣ **Teorema 1.5 (Darboux)** ... Seja  $M$  uma variedade de dimensão  $2n + k$  e  $\omega \in \Omega^2(M)$  uma 2-forma fechada de rank constante igual a  $2n$ . Então, em torno de cada ponto  $p \in M$ , podemos escolher coordenadas locais:

$$(x^i, y^i, z^\ell) = (x^1, \dots, x^n, y^1, \dots, y^n, z^1, \dots, z^k)$$

tais que:

$$\omega = \sum_{i=1}^n dx^i \wedge dy^i \quad (1.13.9)$$

**Dem.:** Como o teorema é puramente local, podemos supôr que  $M = \mathbb{R}^{2n+k}$ , e que  $p = \mathbf{0}$  é a origem. A distribuição Êcaracterística  $\mathfrak{C}_\omega$  sendo integrável, pelo teorema de Frobenius, podemos escolher coordenadas  $(x^i, y^i, z^\ell)$ , em torno de  $\mathbf{0}$ , tais que as folhas de  $\mathfrak{C}_\omega$ , que têm dimensão  $k$ , sejam dadas por  $x^i \equiv c^i$ ,  $y^i \equiv d^i$ , onde  $c^i, d^i$  são constantes. Em particular a folha que passa em  $\mathbf{0}$  é o subespaço  $\mathbf{0}_{2n} \times \mathbb{R}^k \subset \mathbb{R}^{2n+k}$ . Como  $\omega(\mathbf{0})$  tem rank  $2n$ , podemos supôr que  $\omega|_{\mathbb{R}^{2n} \times \mathbf{0}_k}$  tem rank constante e igual a  $2n$ , isto é, essa restrição é não degenerada em  $\mathbb{R}^{2n} \times \mathbf{0}_k \cong \mathbb{R}^{2n}$  e portanto aí define uma forma simpléctica.

Resta então mostrar o teorema quando  $k = 0$ , isto é, quando  $\omega$  é uma forma simpléctica em  $M$ , cuja dimensão é  $2n$ .

♣.

♣ **Proposição 1.9** ... Seja  $M$  uma variedade de dimensão  $2n + k$  e  $\omega \in \Omega^2(M)$  uma 2-forma fechada de rank  $2n$ . Então a dimensão máxima de uma subvariedade integral de  $\omega$ , é igual a  $n + k$ .

**Dem.:** A distribuição Êcaracterística  $\mathfrak{C}_\omega$  sendo integrável, pelo teorema de Frobenius, podemos escolher coordenadas locais  $(x^i)_{i=1, \dots, 2n, 2n+1, \dots, 2n+k}$ , tais que os campos  $\partial/\partial x^\ell$ ,  $\ell = 2n + 1, \dots, 2n + k$  formam uma base para  $\mathfrak{C}_\omega$ .

Se  $X$  é um campo característico então  $X \lrcorner \omega = 0$  e também  $L_X \omega = d(X \lrcorner \omega) + X \lrcorner d\omega = 0$ . Portanto, se localmente:

$$\omega = \omega_{ij} dx^i \wedge dx^j$$

então:

$$\begin{aligned}\omega_{\ell m} &= 0 & \ell, m &= 2n+1, \dots, 2n+k \\ \frac{\partial \omega_{ij}}{\partial x^\ell} &= 0 & \ell &= 2n+1, \dots, 2n+k \quad \forall i, j\end{aligned}\quad (1.13.10)$$

o que significa que  $\omega$  é uma 2-forma apenas nas variáveis  $(x^i)_{i=1, \dots, 2n}$ :

$$\omega = \sum_{1 \leq i < j \leq 2n} \omega_{ij}(x^1, \dots, x^{2n}) dx^i \wedge dx^j \quad (1.13.11)$$

Esta forma já não tem vectores característicos. Mas, como sabemos, qualquer subvariedade integral maximal de  $\omega$  é obtida a partir de uma subvariedade integral maximal da forma (1.13.11), varrendo-a com os fluxos dos campos característicos  $\partial/\partial x^\ell$ ,  $\ell = 2n+1, \dots, 2n+k$ , isto é, “ampliando-a” nas direcções características  $(x^\ell)_{\ell=2n+1, \dots, 2n+k}$ .

Basta então provar a proposição quando  $\omega$  é simpléctica, mostrando que a dimensão de uma subvariedade integral maximal de uma forma simpléctica  $\omega$ , numa variedade de dimensão  $2n$ , é igual a  $n$ . Estas subvariedades integrais maximais chamam-se **subvariedades de Lagrange** de  $\omega$ . Este facto resulta por sua vez do seguinte lema de álgebra linear:

♣ **Lema 1.2** ... Seja  $(V, \omega)$  um espaço vectorial simpléctico de dimensão  $2n$  e  $S$  um subespaço totalmente isotrópico, isto é,  $\omega(u, v) = 0$ ,  $\forall u, v \in S$ . Então  $\dim S \leq n$ .

**Demonstração do Lema** ... Seja  $(u, v) \mapsto \langle u|v \rangle$  um produto interno definido positivo em  $V$ , e representemos  $\omega$  através de um operador  $J : V \rightarrow V$ , definido por:

$$\omega(u, v) = \langle u|J(v) \rangle, \quad u, v \in V$$

Como  $\omega$  é não degenerada  $J$  é um isomorfismo linear. Suponhamos que  $\dim S > n+1$ . Então, como  $\dim(S + J(S)) = \dim S + \dim J(S) - \dim(S \cap J(S))$ , viria que  $\dim(S \cap J(S)) \geq 2$  e portanto  $S \cap J(S) \neq \{0\}$ . Suponhamos então que  $v \in S \cap J(S)$ , com  $v \neq 0$ . Então  $v = J(u)$ , para algum  $u \in S$  e:

$$0 \neq \langle v|v \rangle = \langle v|J(u) \rangle = \omega(v, u) = 0$$

o que é absurdo.

♣.

Em particular:

- numa variedade simpléctica  $(M, \omega)$  de dimensão  $m = 2n$ , a dimensão máxima de uma subvariedade integral de  $\omega$ , é igual a  $2n - n = n$ . Uma tal subvariedade diz-se uma **subvariedade de Lagrange** de  $M$ .
- numa variedade de contacto  $(M, \omega)$  de dimensão  $m = 2n+1$ , a dimensão máxima de uma subvariedade integral de  $\omega$ , é igual a  $2n+1 - n = n+1$ . Uma tal subvariedade diz-se uma **subvariedade de Legendre** de  $M$ .

### 1.13.3 Invariantes integrais

Suponhamos agora que  $M$  é uma variedade de dimensão ímpar e que  $\theta \in \Omega^1 M$  é uma 1-forma diferencial tal que  $d\theta$  é não singular em cada ponto de  $\overline{M}$ . Então, em cada ponto  $p \in M$ , existe uma recta  $\ell_p$ , em  $T_p M$ , unívocamente determinada pela forma  $\theta$ , a que chamamos a **recta característica** da forma  $\theta$ . Portanto, se  $\mathbf{u}_p \in \ell_p$  é um vector não nulo em  $T_p M$ , que gera  $\ell_p$ , tem-se que :

$$d\theta(\mathbf{u}_p, \mathbf{v}_p) = 0, \quad \forall \mathbf{v}_p \in T_p M \quad (1.13.12)$$

Fica então definido um campo de rectas em  $M$ ,  $p \mapsto \ell_p$ , cujas curvas integrais se chamam as **linhas** ou **curvas características** da forma  $\theta$ .

Consideremos uma curva fechada  $\gamma_1$  em  $M$ , transversal ao campo  $\ell$  de rectas características da forma  $\theta$ . As linhas características de  $\theta$ , que partem dos pontos de  $\gamma_1$ , formam um **tubo de características**.

Figure 1.11: .

Temos então o seguinte:

♣ **Lema 1.3 (Lema de Stokes)** ... *Suponhamos que  $\sigma$  é um tubo de características de  $\theta$ , limitado por duas curvas fechadas  $\gamma_1$  e  $\gamma_2$ , isto é:*

$$\partial\sigma = \gamma_1 - \gamma_2$$

Então:

$$\int_{\gamma_1} \theta = \int_{\gamma_2} \theta$$

**Dem.:** Pelo teorema de Stokes:

$$\int_{\gamma_1} \theta - \int_{\gamma_2} \theta = \int_{\partial\sigma} \theta = \int_{\sigma} d\theta = 0$$

a última igualdade resulta de (1.13.12), uma vez que  $\sigma$  é um tubo de características.

♣.

Vamos aplicar o lema de Stokes à situação em que:

$$M = \mathbb{R} \times T^*\mathbb{R}^n \cong \mathbb{R}^{2n+1} \quad (1.13.13)$$

é o chamado **espaço de fases estendido**, munido de coordenadas  $(t, \mathbf{x}, \mathbf{p}) = (t, x^i, p_i)$ , e a 1-forma  $\theta$  é a **forma de Poincaré-Cartan**, dada por:

$$\theta = \mathbf{p} \, d\mathbf{x} - H dt = \sum_i p_i dx^i - H(t, \mathbf{x}, \mathbf{p}) dt \quad (1.13.14)$$

onde  $H \in C^\infty(\mathbb{R} \times T^*\mathbb{R}^n)$  é uma função, chamada o **Hamiltoniano** (dependente do tempo). Notemos que:

$$d\theta = \sum_i \left( dp_i \wedge dx^i - \frac{\partial H}{\partial x^i} dx^i \wedge dt - \frac{\partial H}{\partial p_i} dp_i \wedge dt \right) \quad (1.13.15)$$

e portanto a matriz de Gram de  $d\theta$ , na base  $\{\partial_{\mathbf{x}}, \partial_{\mathbf{p}}, \partial_t\}$ , é a matriz:

$$\begin{bmatrix} \mathbf{0} & -\mathbf{I}_n & -H_{\mathbf{x}} \\ \mathbf{I}_n & \mathbf{0} & -H_{\mathbf{p}} \\ H_{\mathbf{x}}^T & H_{\mathbf{p}}^T & 0 \end{bmatrix} \quad (1.13.16)$$

onde  $H_{\mathbf{p}} = \left[ \frac{\partial H}{\partial p_i} \right]$  e  $H_{\mathbf{x}} = \left[ \frac{\partial H}{\partial x^i} \right]$ . O rank desta matriz é obviamente  $2n$  e portanto  $d\theta$  é não singular. O seu núcleo,  $\ker d\theta$ , é gerado pelo vector:

$$H_{\mathbf{p}} \partial_{\mathbf{x}} - H_{\mathbf{x}} \partial_{\mathbf{p}} + \partial_t = \sum_i \left( \frac{\partial H}{\partial p_i} \frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i} \frac{\partial}{\partial p_i} \right) + \frac{\partial}{\partial t} \quad (1.13.17)$$

que gera portanto a recta característica da forma de Poincaré-Cartan  $\theta$ . As linhas características de  $\theta$  são as curvas integrais deste campo de vectores. Fica assim provado o seguinte teorema:

♣ **Teorema 1.6** ... *As linhas características da 1-forma de Poincaré-Cartan:*

$$\theta = \mathbf{p} \, d\mathbf{x} - H \, dt$$

no espaço das fases estendido  $\mathbb{R} \times T^*\mathbb{R}^n$ , projectam-se difeomòrficamente sobre o eixo dos  $t$ 's, e portanto podem ser parametrizadas na forma  $t \mapsto (\mathbf{x}(t), \mathbf{p}(t))$ . Estas funções satisfazem as **equações canónicas de Hamilton** seguintes:

$$\begin{cases} \dot{\mathbf{x}} = H_{\mathbf{p}} \\ \dot{\mathbf{p}} = -H_{\mathbf{x}} \end{cases} = \begin{cases} \frac{dx^i}{dt} = \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} = -\frac{\partial H}{\partial x^i} \end{cases}, \quad i = 1, \dots, n \quad (1.13.18)$$

♣.

Aplicando agora o Lema de Stokes 1.3 à forma  $\theta = \mathbf{p} \, d\mathbf{x} - H \, dt$ , obtemos o seguinte teorema fundamental:

Figure 1.12: .

♣ **Teorema 1.7** ... Suponhamos que  $\sigma$  é um tubo de características da forma de Poincaré-Cartan  $\theta = \mathbf{p} d\mathbf{x} - H dt$ , limitado por duas curvas fechadas  $\gamma_1$  e  $\gamma_2$ , isto é,  $\partial\sigma = \gamma_1 - \gamma_2$ . Então:

$$\int_{\gamma_1} \mathbf{p} d\mathbf{x} - H dt = \int_{\gamma_2} \mathbf{p} d\mathbf{x} - H dt \quad (1.13.19)$$

O integral  $\int \mathbf{p} d\mathbf{x} - H dt$  chama-se o **invariante integral de Hilbert**. ♣.

Vamos considerar, em particular, curvas fechadas  $\gamma \subset \{t\} \times T^*\mathbb{R}^n$  com  $t$  fixo, isto é, curvas fechadas constituídas por estados simultâneos. Ao longo de tais curvas  $dt = 0$  e  $\int \mathbf{p} d\mathbf{x} - H dt = \int \mathbf{p} d\mathbf{x}$ . Consideremos ainda a transformação:

$$\begin{aligned} \text{Fl}_{t_0}^{t_1} : \{t_0\} \times T^*\mathbb{R}^n &\longrightarrow \{t_1\} \times T^*\mathbb{R}^n \\ (\mathbf{x}_0, \mathbf{p}_0) &\longmapsto (\mathbf{x}_1, \mathbf{p}_1) = \text{Fl}_{t_0}^{t_1}(\mathbf{x}_0, \mathbf{p}_0) \end{aligned} \quad (1.13.20)$$

onde  $(\mathbf{x}_1, \mathbf{p}_1) = \text{Fl}_{t_0}^{t_1}(\mathbf{x}_0, \mathbf{p}_0) \in \{t_1\} \times T^*\mathbb{R}^n$  é o ponto obtido a partir de  $(\mathbf{x}_0, \mathbf{p}_0)$  seguindo o fluxo do **campo Hamiltoniano**  $X_H = H_{\mathbf{p}}\partial_{\mathbf{x}} - H_{\mathbf{x}}\partial_{\mathbf{p}}$ , isto é, resolvendo as equações canônicas de Hamilton (1.13.18), com condições iniciais  $\mathbf{x}(t_0) = \mathbf{x}_0$  e  $\mathbf{p}(t_0) = \mathbf{p}_0$ .

Figure 1.13: .

Se  $\gamma \subset \{t_0\} \times T^*\mathbb{R}^n$  é uma curva fechada, então  $\gamma_1 = \text{Fl}_{t_0}^{t_1}(\gamma)$  é também uma curva fechada em  $\{t_1\} \times T^*\mathbb{R}^n$ . Além disso elas limitam o mesmo tubo de características da

forma de Poincaré-Cartan  $\theta = \mathbf{p} d\mathbf{x} - H dt$ . Portanto, aplicando o teorema anterior, obtemos:

$$\int_{\gamma} \mathbf{p} d\mathbf{x} = \int_{\text{Fl}_{t_0}^{t_1}(\gamma)} \mathbf{p} d\mathbf{x} \quad (1.13.21)$$

isto é, “o fluxo do campo Hamiltoniano  $X_H$  preserva o integral da **forma de Liouville**  $\mathbf{p} d\mathbf{x} = \sum_i p_i dx^i$ , ao longo de curvas fechadas”.

# Capítulo 2

## Problemas variacionais paramétricos

### 2.1 Lagrangeanos paramétricos homogêneos

Nesta secção vamos discutir problemas variacionais descritos por funcionais do tipo:

$$\mathfrak{F}[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \quad (2.1.1)$$

em que o Lagrangeano  $L : T\mathbb{R}^n \rightarrow \mathbb{R}$  não depende do parâmetro  $t$ , e é homogêneo positivo de grau 1 em  $\dot{\mathbf{x}}$ , isto é:

$$L(\mathbf{x}, \lambda \dot{\mathbf{x}}) = \lambda L(\mathbf{x}, \dot{\mathbf{x}}), \quad \forall \lambda > 0 \quad (2.1.2)$$

O funcional  $\mathfrak{F}$  considera-se definido no conjunto de curvas geométricas de classe  $C^1$  regulares, em  $\mathbb{R}^n$ . Uma curva geométrica é uma classe de equivalência de curvas parametrizadas, em que duas dessas curvas são equivalentes sse diferem por reparametrizações de classe  $C^1$  que preservam orientação.

Para ver que de facto  $\mathfrak{F}$  está bem definido, consideremos duas curvas parametrizadas  $\mathbf{x} : [t_0, t_1] \rightarrow \mathbb{R}^n$  e  $\mathbf{y} : [\tau_0, \tau_1] \rightarrow \mathbb{R}^n$  equivalentes, de tal forma que existe um difeomorfismo  $\varphi : [\tau_0, \tau_1] \rightarrow [t_0, t_1]$ ,  $t = \varphi(\tau)$ , tal que  $\varphi'(\tau) > 0$  e  $\mathbf{y}(\tau) = \mathbf{x}(\varphi(\tau))$ . Vem então que:

$$\begin{aligned} \mathfrak{F}[\mathbf{y}(\cdot)] &= \int_{\tau_0}^{\tau_1} L(\mathbf{y}(\tau), \mathbf{y}'(\tau)) d\tau \\ &= \int_{\tau_0}^{\tau_1} L(\mathbf{x}(\varphi(\tau)), \dot{\mathbf{x}}(\varphi(\tau))\varphi'(\tau)) d\tau \\ &= \int_{\tau_0}^{\tau_1} L(\mathbf{x}(\varphi(\tau)), \dot{\mathbf{x}}(\varphi(\tau))) \varphi'(\tau) d\tau \\ &= \int_{t_0}^{t_1} L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \\ &= \mathfrak{F}[\mathbf{x}(\cdot)] \end{aligned} \quad (2.1.3)$$

Vejamos alguns exemplos de Lagrangeanos paramétricos homogêneos:

♣ Exemplos 2.1 ...

1.  $L(\mathbf{x}, \dot{\mathbf{x}}) = \|\dot{\mathbf{x}}\|$ . O funcional  $\mathfrak{F}[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} \|\dot{\mathbf{x}}\| dt$  representa o comprimento Euclidiano da curva  $\mathbf{x} : [t_0, t_1] \rightarrow \mathbb{R}^n$ .
2.  $L(\mathbf{x}, \dot{\mathbf{x}}) = n(\mathbf{x}) \|\dot{\mathbf{x}}\|$ . O funcional  $\mathfrak{F}[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} n(\mathbf{x}) \|\dot{\mathbf{x}}\| dt = \int_{t_0}^{t_1} n(\mathbf{x}) ds$  representa o comprimento óptico do raio  $\mathbf{x} : [t_0, t_1] \rightarrow \mathbb{R}^n$ , que se propaga num meio isotrópico não homogêneo de índice de refração  $n = n(\mathbf{x}) > 0$ .
3.  $L(\mathbf{x}, \dot{\mathbf{x}}) = \sqrt{g(\mathbf{x})(\dot{\mathbf{x}}, \dot{\mathbf{x}})} = \sqrt{g_{ij}(\mathbf{x})\dot{x}^i\dot{x}^j}$ , onde  $g$  é uma métrica Riemanniana em  $\mathbb{R}^n$ . O funcional  $\mathfrak{F}[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} \sqrt{g(\mathbf{x})(\dot{\mathbf{x}}, \dot{\mathbf{x}})} dt$  representa o comprimento da curva  $\mathbf{x} : [t_0, t_1] \rightarrow \mathbb{R}^n$ , relativamente à métrica Riemanniana  $g$ .

Começemos por discutir o seguinte problema clássico do cálculo de variações para Lagrangeanos paramétricos homogêneos:

- ♣ **Problema 2.1** ... *Entre as curvas geométricas de classe  $C^1$  regulares, em  $\mathbb{R}^n$ , que satisfazem as condições de fronteira:*

$$\mathbf{x}(t_0) = \mathbf{x}_0, \quad \mathbf{x}(t_1) = \mathbf{x}_1 \quad (2.1.4)$$

onde  $\mathbf{x}_0, \mathbf{x}_1$  são dois pontos fixos em  $\mathbb{R}^n$ , calcular a curva para a qual o valor do funcional (2.1.1) é estacionário <sup>1</sup>.

Figure 2.1:

Se  $\hat{\mathbf{x}} : [t_0, t_1] \rightarrow \mathbb{R}^n$  é uma solução do Problema 2.1, e se  $\mathbf{x}(\cdot; \lambda) : [t_0, t_1] \rightarrow \mathbb{R}^n$ ,  $\lambda \in \mathbb{R}$  é uma família a 1-parâmetro  $\lambda$  de curvas em  $\mathbb{R}^n$  (que depende diferenciavelmente do parâmetro  $\lambda$ ), tal que:

$$\begin{aligned} \mathbf{x}(\cdot; \lambda = 0) &= \hat{\mathbf{x}}(\cdot) \\ \mathbf{x}(t_0; \lambda) &= \mathbf{x}_0 \quad \text{e} \quad \mathbf{x}(t_1; \lambda) = \mathbf{x}_1, \quad \forall \lambda \\ \delta \hat{\mathbf{x}}(\cdot) = \boldsymbol{\eta}(\cdot) &\stackrel{\text{def}}{=} \left. \frac{d}{d\lambda} \right|_{\lambda=0} \mathbf{x}(\cdot; \lambda) \end{aligned} \quad (2.1.5)$$

<sup>1</sup>Mais geralmente,  $\mathbb{R}^n$  pode ser substituído por uma variedade suave  $M$  de dimensão  $n$ .

onde  $\boldsymbol{\eta}(\cdot) = \delta\widehat{\mathbf{x}}(\cdot)$  é uma variação com extremidades fixas, então:

$$\begin{aligned}
 0 &= \left. \frac{\partial}{\partial \lambda} \right|_{\lambda=0} \mathfrak{F}[\mathbf{x}(\cdot; \lambda)] \\
 &= \left. \frac{\partial}{\partial \lambda} \right|_{\lambda=0} \int_{t_0}^{t_1} L(\mathbf{x}(\cdot; \lambda), \dot{\mathbf{x}}(t; \lambda)) dt \\
 &= \int_{t_0}^{t_1} [L_{\mathbf{x}}(t) \boldsymbol{\eta}(t) + L_{\dot{\mathbf{x}}}(t) \dot{\boldsymbol{\eta}}(t)] dt \\
 &= \int_{t_0}^{t_1} \left[ L_{\mathbf{x}}(t) - \frac{d}{dt} L_{\dot{\mathbf{x}}}(t) \right] \boldsymbol{\eta}(t) dt + [L_{\dot{\mathbf{x}}}(t) \boldsymbol{\eta}(t)]_{t_0}^{t_1} \tag{2.1.6}
 \end{aligned}$$

onde  $L_{\mathbf{x}}(t) = L_{\mathbf{x}}(\widehat{\mathbf{x}}(t), \dot{\widehat{\mathbf{x}}}(t))$  e  $L_{\dot{\mathbf{x}}}(t) = L_{\dot{\mathbf{x}}}(\widehat{\mathbf{x}}(t), \dot{\widehat{\mathbf{x}}}(t))$ . A esta última fórmula é habitual chamar a **fórmula da primeira variação**, e notá-la por  $\delta\mathfrak{F}(\widehat{\mathbf{x}}; \boldsymbol{\eta})$ .

A este integral aplicamos o lema de Du Bois-Reymond e o facto de que  $\boldsymbol{\eta}(t_0) = \boldsymbol{\eta}(t_1) = 0$ , para concluir que  $\widehat{\mathbf{x}}(\cdot)$  deve satisfazer a **equação de Euler-Lagrange**:

$$-\frac{d}{dt} L_{\dot{\mathbf{x}}} + L_{\mathbf{x}} = \mathbf{0} \tag{2.1.7}$$

Qualquer solução das equações de Euler-Lagrange diz-se uma **extremal** do problema variacional 2.1.

Notas ...

1. Como, por hipótese, o Lagrangeano  $L$  é homogéneo de grau 1 nas variáveis  $\dot{\mathbf{x}}$ , isto é,  $L(\mathbf{x}, \lambda \dot{\mathbf{x}}) = \lambda L(\mathbf{x}, \dot{\mathbf{x}})$ ,  $\forall \lambda > 0$ , temos que  $L_{\mathbf{x}}$  é homogéneo de grau 1, e  $L_{\dot{\mathbf{x}}}$  é homogéneo de grau 0 nas variáveis  $\dot{\mathbf{x}}$ , isto é:

$$L_{\mathbf{x}}(\mathbf{x}, \lambda \dot{\mathbf{x}}) = \lambda L_{\mathbf{x}}(\mathbf{x}, \dot{\mathbf{x}}), \quad \text{e} \quad L_{\dot{\mathbf{x}}}(\mathbf{x}, \lambda \dot{\mathbf{x}}) = L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}), \quad \forall \lambda > 0 \tag{2.1.8}$$

É fácil verificar, usando estas relações, que, se  $\mathbf{x}(t)$  é solução da equação de Euler-Lagrange (2.1.7), então qualquer reparametrização de  $\mathbf{x}$ , digamos  $\mathbf{y}(\tau) = (\mathbf{x} \circ \varphi)(\tau)$ , onde  $\varphi'(\tau) > 0$ , é também solução dessa mesma equação.

2. Como o Lagrangeano  $L$  é homogéneo de grau 1 nas variáveis  $\dot{\mathbf{x}}$ , a identidade de Euler implica que:

$$L_{\dot{\mathbf{x}}}\dot{\mathbf{x}} = L, \quad \text{isto é} \quad \sum_{i=1}^n \dot{x}^i L_{\dot{x}^i}(\mathbf{x}, \dot{\mathbf{x}}) = L(\mathbf{x}, \dot{\mathbf{x}})$$

e portanto a energia de  $L$  é nula:

$$E_L(\mathbf{x}, \dot{\mathbf{x}}) \stackrel{\text{def}}{=} L_{\dot{\mathbf{x}}}\dot{\mathbf{x}} - L \equiv 0 \tag{2.1.9}$$

Esta energia é conservada, já que  $L$  não depende do parâmetro.

3. Por conservação de energia e pela igualdade (2.1.9), temos que  $L_{\dot{\mathbf{x}}}(\mathbf{x}(t), \dot{\mathbf{x}}(t)) \dot{\mathbf{x}}(t) - L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = 0$ . Portanto, derivando em ordem a  $t$ , obtemos:

$$\begin{aligned} 0 &= \left( \frac{d}{dt} L_{\dot{\mathbf{x}}} \right) \dot{\mathbf{x}} + L_{\dot{\mathbf{x}}} \ddot{\mathbf{x}} - L_{\mathbf{x}} \dot{\mathbf{x}} - L_{\dot{\mathbf{x}}} \ddot{\mathbf{x}} \\ &= \left( \frac{d}{dt} L_{\dot{\mathbf{x}}} - L_{\mathbf{x}} \right) \dot{\mathbf{x}} \end{aligned}$$

o que significa que as equações de Euler-Lagrange não são independentes e estão ligadas pela relação:

$$\sum_{i=1}^n \left( \frac{d}{dt} L_{\dot{x}^i} - L_{x^i} \right) \dot{x}^i = 0 \quad (2.1.10)$$

♣.

Suponhamos agora que temos um problema com extremidade móvel, mas condicionada a mover-se numa subvariedade  $\Sigma$ , de codimensão  $k$ , em  $\mathbb{R}^n$ , dada por uma equação do tipo:

$$\Phi(\mathbf{x}) = \mathbf{0} \quad (2.1.11)$$

onde  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^k$  é uma submersão. Mais precisamente, vamos discutir o problema seguinte:

- ♣ **Problema 2.2** ... *Entre as curvas geométricas que satisfazem as condições de fronteira:*

$$\mathbf{x}(t_0) = \mathbf{x}_0, \quad \Phi(\mathbf{x}(t_1)) = \mathbf{0} \quad (2.1.12)$$

*calcular a curva para a qual o valor do funcional:*

$$I[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \quad (2.1.13)$$

*é estacionário.*

Figure 2.2:

Se  $\mathbf{x} : [t_0, t_1] \rightarrow \mathbb{R}^n$  é uma solução deste problema, então também será solução do problema com extremidades fixas  $\mathbf{x}(t_0) = \mathbf{x}_0$  e  $\mathbf{x}(t_1) = \mathbf{x}_1$ . Portanto  $\mathbf{x}(\cdot)$  satisfaz a

equação de Euler-Lagrange. Por outro lado, se  $\lambda \mapsto \mathbf{x}(\cdot; \lambda)$ ,  $\lambda \in \mathbb{R}$  é uma família a 1-parâmetro  $\lambda$  de curvas em  $\mathbb{R}^n$  (que depende diferenciavelmente do parâmetro  $\lambda$ ), tal que:

$$\begin{aligned} \mathbf{x}(\cdot; \lambda = 0) &= \mathbf{x}(\cdot) \\ \mathbf{x}(t_0; \lambda) &= \mathbf{x}_0 \quad \text{e} \quad \Phi(\mathbf{x}(t_1; \lambda)) = \mathbf{0}, \quad \forall \lambda \\ \delta \mathbf{x}(\cdot) = \boldsymbol{\eta}(\cdot) &\stackrel{\text{def}}{=} \left. \frac{d}{d\lambda} \right|_{\lambda=0} \mathbf{x}(\cdot; \lambda) \end{aligned} \quad (2.1.14)$$

então, em particular, tem-se que  $\boldsymbol{\eta}(t_0) = \mathbf{0}$  e:

$$\boldsymbol{\eta}(t_1) = \left. \frac{d}{d\lambda} \right|_{\lambda=0} \mathbf{x}(t_1; \lambda) = \delta \mathbf{x}_1 \in T_{\mathbf{x}_1} \Sigma \quad \text{isto é} \quad d\Phi_{\mathbf{x}_1}(\delta \mathbf{x}_1) = \mathbf{0} \quad (2.1.15)$$

Pela fórmula da primeira variação (2.1.6), deduzimos que:

$$\begin{aligned} 0 &= \delta \mathfrak{F}(\mathbf{x}; \boldsymbol{\eta}) \\ &= \int_{t_0}^{t_1} \left[ L_{\mathbf{x}}(t) - \frac{d}{dt} L_{\dot{\mathbf{x}}}(t) \right] \boldsymbol{\eta}(t) dt + [L_{\dot{\mathbf{x}}}(t) \boldsymbol{\eta}(t)]_{t_0}^{t_1} \\ &= [L_{\dot{\mathbf{x}}}(t) \boldsymbol{\eta}(t)]_{t_0}^{t_1} \\ &= L_{\dot{\mathbf{x}}}(\mathbf{x}_1, \dot{\mathbf{x}}_1) \delta \mathbf{x}_1 \end{aligned} \quad (2.1.16)$$

Designando, como habitualmente, por:

$$\mathbf{p}_1 \stackrel{\text{def}}{=} L_{\dot{\mathbf{x}}}(\mathbf{x}_1, \dot{\mathbf{x}}_1) \quad (2.1.17)$$

o momento conjugado a  $\mathbf{x}$  (calculado em  $(\mathbf{x}_1, \dot{\mathbf{x}}_1)$ ), vemos que a extremal tem que verificar a **condição de transversalidade** ou **ortogonalidade** seguinte, na sua extremidade móvel:

$$\boxed{\mathbf{p}_1 \delta \mathbf{x}_1 = 0, \quad \forall \delta \mathbf{x}_1 \in T_{\mathbf{x}_1} \Sigma} \quad (2.1.18)$$

Por exemplo, se  $L(\mathbf{x}, \dot{\mathbf{x}}) = n(\mathbf{x}) \|\dot{\mathbf{x}}\|$  é o Lagrangeano óptico, então  $\mathbf{p} = \mathbf{p}(\mathbf{x}, \dot{\mathbf{x}}) = L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) = n(\mathbf{x}) \frac{\dot{\mathbf{x}}}{\|\dot{\mathbf{x}}\|}$ , e portanto a condição de transversalidade é:

$$n(\mathbf{x}_1) \frac{\dot{\mathbf{x}}_1}{\|\dot{\mathbf{x}}_1\|} \delta \mathbf{x}_1 = 0, \quad \forall \delta \mathbf{x}_1 \in T_{\mathbf{x}_1} \Sigma \quad (2.1.19)$$

Se  $n(\mathbf{x}_1) \neq 0$ , onde  $\mathbf{x}_1 \in \Sigma$ , isto traduz a ortogonalidade usual:

$$\dot{\mathbf{x}}_1 \perp T_{\mathbf{x}_1} \Sigma \quad (2.1.20)$$

## 2.2 Formalismo canónico

Como já vimos, sendo o Lagrangeano  $L$  homogéneo de grau 1 nas variáveis  $\dot{\mathbf{x}}$ , a identidade de Euler implica que:

$$L_{\dot{\mathbf{x}}} \dot{\mathbf{x}} = L, \quad \text{isto é} \quad \sum_{i=1}^n \dot{x}^i L_{\dot{x}^i}(\mathbf{x}, \dot{\mathbf{x}}) = L(\mathbf{x}, \dot{\mathbf{x}})$$

Derivando novamente em ordem a  $\dot{\mathbf{x}}$ , obtem-se:

$$L_{\dot{\mathbf{x}}\dot{\mathbf{x}}}\dot{\mathbf{x}} + L_{\dot{\mathbf{x}}} = L_{\dot{\mathbf{x}}}, \quad \Rightarrow \quad L_{\dot{\mathbf{x}}\dot{\mathbf{x}}}\dot{\mathbf{x}} = 0$$

isto é:

$$\sum_{i,j=1}^n L_{\dot{x}^i \dot{x}^k}(\mathbf{x}, \dot{\mathbf{x}}) \dot{x}^k = 0, \quad \forall \dot{\mathbf{x}} \neq \mathbf{0} \quad (2.2.1)$$

o que significa que a equação:

$$\mathbf{p} = L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}})$$

não pode ser resolvida em ordem a  $\dot{\mathbf{x}}$  e, portanto, não podemos definir o formalismo canónico via transformada de Legendre, como se fez para Lagrangeanos não paramétricos hiperregulares. Além disso, como também vimos antes, a energia de  $L$  é nula  $E_L = L_{\dot{\mathbf{x}}}\dot{\mathbf{x}} - L \equiv 0$ .

O formalismo canónico que vamos expôr deve-se a Rund, e baseia-se na introdução de um novo Lagrangeano  $Q$ , definido por  $Q = \frac{1}{2}L^2$ , isto é:

$$Q(\mathbf{x}, \dot{\mathbf{x}}) \stackrel{\text{def}}{=} \frac{1}{2}L^2(\mathbf{x}, \dot{\mathbf{x}}) \quad (2.2.2)$$

Eis algumas propriedades deste novo Lagrangeano  $Q$ :

- $Q(\mathbf{x}, \dot{\mathbf{x}}) \geq 0$  e  $Q(\mathbf{x}, \dot{\mathbf{x}}) = 0$  se e só se  $L(\mathbf{x}, \dot{\mathbf{x}}) = 0$ .
- $Q(\mathbf{x}, \lambda \dot{\mathbf{x}}) = \lambda^2 Q(\mathbf{x}, \dot{\mathbf{x}})$ ,  $\forall \lambda > 0$ , isto é,  $Q$  é homogéneo positivo de grau 2 nas variáveis  $\dot{\mathbf{x}}$ .
- Derivando esta última igualdade, duas vezes em ordem a  $\lambda$  (com  $(\mathbf{x}, \dot{\mathbf{x}})$  fixo), obtemos:

$$Q_{\dot{\mathbf{x}}}(\mathbf{x}, \lambda \dot{\mathbf{x}}) \dot{\mathbf{x}} = 2\lambda Q(\mathbf{x}, \dot{\mathbf{x}}), \quad \text{isto é} \quad \sum_i Q_{\dot{x}^i}(\mathbf{x}, \lambda \dot{\mathbf{x}}) \dot{x}^i = 2\lambda Q(\mathbf{x}, \dot{\mathbf{x}}) \quad (2.2.3)$$

e ainda:

$$Q_{\dot{\mathbf{x}}\dot{\mathbf{x}}}(\mathbf{x}, \lambda \dot{\mathbf{x}}) \dot{\mathbf{x}}\dot{\mathbf{x}} = 2Q(\mathbf{x}, \dot{\mathbf{x}}), \quad \text{isto é} \quad \sum_{ik} Q_{\dot{x}^i \dot{x}^k}(\mathbf{x}, \lambda \dot{\mathbf{x}}) \dot{x}^i \dot{x}^k = 2Q(\mathbf{x}, \dot{\mathbf{x}}) \quad (2.2.4)$$

♣.

Vejam agora qual a relação que existe entre as extremais associadas a cada um destes Lagrangeanos, para o caso mais frequente em que  $L(\mathbf{x}, \dot{\mathbf{x}}) > 0$ ,  $\forall \mathbf{x}, \dot{\mathbf{x}} \neq \mathbf{0}$ .

- **♣ Proposição 2.1** ... *Suponhamos que  $L(\mathbf{x}, \dot{\mathbf{x}}) > 0$ ,  $\forall \mathbf{x}, \dot{\mathbf{x}} \neq \mathbf{0}$  e seja  $Q = \frac{1}{2}L^2$ . Consideremos ainda os seguintes funcionais:*

$$\mathfrak{F}_L[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} L(\mathbf{x}, \dot{\mathbf{x}}) dt, \quad \mathfrak{F}_Q[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} Q(\mathbf{x}, \dot{\mathbf{x}}) dt \quad (2.2.5)$$

Então:

(i). Toda a  $Q$ -extremal  $\mathbf{x}(t)$ ,  $t_0 \leq t \leq t_1$ , satisfaz:

$$Q(\mathbf{x}(t), \dot{\mathbf{x}}(t)) \equiv \frac{1}{2}h^2 \quad (2.2.6)$$

para alguma constante  $h > 0$ , e é também uma  $L$ -extremal.

(ii). Reciprocamente, toda a  $L$ -extremal (geométrica)  $\mathbf{x}(t)$ ,  $t_0 \leq t \leq t_1$  possui uma única parametrização  $\mathbf{y}(\tau) = (\mathbf{x} \circ \varphi)(\tau)$ ,  $0 \leq \tau \leq T$ , para alguma constante  $T > 0$ , para a qual:

$$L(\mathbf{y}(\tau), \mathbf{y}'(\tau)) \equiv 1 \quad (2.2.7)$$

Esta curva parametrizada é então uma  $Q$ -extremal.

**Dem.:** (i). Como  $Q$  não depende de  $t$ , há conservação de energia - ao longo de uma  $Q$ -extremal,  $E_Q = Q_{\dot{\mathbf{x}}}\dot{\mathbf{x}} - Q$  é conservada. Mas, por (2.2.3):

$$E_Q = Q_{\dot{\mathbf{x}}}\dot{\mathbf{x}} - Q = 2Q - Q = Q$$

e portanto, ao longo de uma  $Q$ -extremal, tem-se:

$$Q(\mathbf{x}(t), \dot{\mathbf{x}}(t)) \equiv \frac{1}{2}h^2$$

para alguma constante  $h > 0$ . Mas isto acontece se e só se:

$$L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) \equiv h$$

Como  $Q_{\dot{\mathbf{x}}} = LL_{\dot{\mathbf{x}}}$  e  $Q_{\mathbf{x}} = LL_{\mathbf{x}}$ , obtemos:

$$-\frac{d}{dt}Q_{\dot{\mathbf{x}}} + Q_{\mathbf{x}} = h \left[ -\frac{d}{dt}L_{\dot{\mathbf{x}}} + L_{\mathbf{x}} \right] \quad (2.2.8)$$

e portanto, toda a  $Q$ -extremal que satisfaz (2.2.6) é também uma  $L$ -extremal.

(ii). Começemos agora com uma  $L$ -extremal  $\mathbf{x}(t)$ ,  $t_0 \leq t \leq t_1$ . Pretende-se calcular uma mudança do parâmetro  $t = \varphi(\tau)$  e  $\varphi' > 0$ , a que corresponda uma nova parametrização  $\mathbf{y}(\tau) = (\mathbf{x} \circ \varphi)(\tau)$ ,  $0 \leq \tau \leq T$ , para alguma constante  $T > 0$  a determinar pela condição de que:

$$L(\mathbf{y}(\tau), \mathbf{y}'(\tau)) = L(\mathbf{x}(\varphi(\tau)), \dot{\mathbf{x}}(\varphi(\tau))\varphi'(\tau)) = L(\mathbf{x}(t), \dot{\mathbf{x}}(t))\varphi'(\tau) \equiv 1$$

Mas a função  $f(t) = L(\mathbf{x}(t), \dot{\mathbf{x}}(t))$  é conhecida. É contínua e estritamente positiva. Pretende-se pois que:

$$f(t)\varphi'(\tau) \equiv 1, \quad \text{onde} \quad t = \varphi(\tau)$$

Por outras palavras,  $f(t)\frac{dt}{d\tau} \equiv 1$  donde se tira o valor de  $\tau$ :

$$\tau = \psi(t) = \int_{t_0}^t f(\xi)d\xi + c$$

onde  $c$  é uma constante. Os valores das constantes  $T$  e  $c$ , determinam-se pelas condições de que  $\psi(t_0) = 0$  e  $\psi(t_1) = T$ . Portanto  $c = 0$  e  $T = \int_{t_0}^{t_1} f(t) dt = \int_{t_0}^{t_1} L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt$ , isto é:

$$\tau = \psi(t) = \int_{t_0}^t L(\mathbf{x}(\xi), \dot{\mathbf{x}}(\xi)) d\xi$$

Como  $\mathbf{x}$  é uma  $L$ -extremal, também o é a nova reparametrização  $\mathbf{y} = \mathbf{x} \circ \varphi$ , como se viu antes. Mas, como  $L$  é constante sobre esta curva,  $\mathbf{y}$  é uma  $Q$ -extremal, como se deduz imediatamente de (2.2.8).



Concluimos pois que a classe de  $Q$ -extremais coincide com a classe de  $L$ -extremais normalizadas pela condição  $L(\mathbf{x}, \dot{\mathbf{x}}) \equiv 1$ . Enquanto que as  $L$ -extremais são invariantes por reparametrização, as  $Q$ -extremais vêm automaticamente parametrizadas com um parâmetro “natural”.

♣ **Exemplo 2.1 (Geodésicas)** ... Quando  $L(\mathbf{x}, \dot{\mathbf{x}}) = \sqrt{g(\mathbf{x})(\dot{\mathbf{x}}, \dot{\mathbf{x}})} = \sqrt{g_{ij}(\mathbf{x})\dot{x}^i\dot{x}^j}$ , onde  $g$  é uma métrica Riemanniana em  $\mathbb{R}^n$ , o funcional:

$$\mathfrak{F}[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} \sqrt{g(\mathbf{x})(\dot{\mathbf{x}}, \dot{\mathbf{x}})} dt$$

representa o comprimento da curva  $\mathbf{x} : [t_0, t_1] \rightarrow \mathbb{R}^n$ , relativamente à métrica Riemanniana  $g$ .

Consideremos agora o novo Lagrangeano:

$$Q(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}L^2(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}g(\mathbf{x})(\dot{\mathbf{x}}, \dot{\mathbf{x}}) = \frac{1}{2}g_{ij}(\mathbf{x})\dot{x}^i\dot{x}^j$$

a que chamamos a **energia (cinética) da métrica**  $g$ . As equações de Euler-Lagrange correspondentes ao Lagrangeano  $Q$  são:

$$-\frac{d}{dt}(g(\mathbf{x})\dot{\mathbf{x}}) + g_{\mathbf{x}}(\mathbf{x})\dot{\mathbf{x}}\dot{\mathbf{x}} = \mathbf{0} \quad (2.2.9)$$

O parâmetro “natural” é, neste caso, o comprimento de arco  $\tau = s$ :

$$s = \psi(t) = \int_{t_0}^t \sqrt{g(\mathbf{x}(\xi))(\dot{\mathbf{x}}(\xi), \dot{\mathbf{x}}(\xi))} d\xi$$

e a proposição diz que as geodésicas da métrica  $g$ , isto é, as  $L$ -extremais, quando parametrizadas por arco, têm energia cinética constante (igual a 1/2), e são também  $Q$ -extremais.

♣ **Exemplo 2.2 (O Lagrangeano óptico)** ... Como vimos o Lagrangeano óptico é:

$$L(\mathbf{x}, \dot{\mathbf{x}}) = n(\mathbf{x}) \|\dot{\mathbf{x}}\|$$

O funcional:

$$\mathfrak{F}_L[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} n(\mathbf{x}) \|\dot{\mathbf{x}}\| dt = \int_{t_0}^{t_1} n(\mathbf{x}) ds$$

representa o comprimento óptico do raio  $\mathbf{x} : [t_0, t_1] \rightarrow \mathbb{R}^3$ , que se propaga num meio isotrópico não homogéneo de índice de refração  $n = n(\mathbf{x}) > 0$ . Como já se viu,  $\mathfrak{F}_L[\mathbf{x}(\cdot)]$  representa também o tempo de percurso da luz ao longo desse mesmo raio.

As equações de Euler-Lagrange para os raios ( $L$ -extremais) são:

$$\frac{d}{dt} \left( n(\mathbf{x}) \frac{\dot{\mathbf{x}}}{\|\dot{\mathbf{x}}\|} \right) = \|\dot{\mathbf{x}}\| \nabla n(\mathbf{x}) \quad (2.2.10)$$

ou mais explicitamente:

$$\begin{cases} \frac{d}{dt} \frac{n \dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = n_x \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} \\ \frac{d}{dt} \frac{n \dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = n_y \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} \\ \frac{d}{dt} \frac{n \dot{z}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = n_z \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} \end{cases} \quad (2.2.11)$$

Consideremos agora o novo Lagrangeano:

$$Q(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} L^2(\mathbf{x}, \dot{\mathbf{x}}) = \frac{n^2 \dot{\mathbf{x}}^2}{2}$$

As equações de Euler-Lagrange correspondentes ao Lagrangeano  $Q$ , ou as equações dos  $Q$ -raios, são:

$$\mathbf{0} = -\frac{d}{dt} Q_{\dot{\mathbf{x}}} + Q_{\mathbf{x}} = -\frac{d}{dt} (n^2 \dot{\mathbf{x}}) + \dot{\mathbf{x}}^2 \frac{\nabla(n^2)}{2}$$

isto é:

$$\frac{d}{dt} (n^2 \dot{\mathbf{x}}) = \dot{\mathbf{x}}^2 \frac{\nabla(n^2)}{2} \quad (2.2.12)$$

Consideremos um  $L$ -raio,  $\mathbf{x} : [t_0, t_1] \rightarrow \mathbb{R}^3$ , isto é, uma solução das equações de Euler-Lagrange (2.2.10). A proposição anterior mostra que se escolhermos o novo parâmetro  $\tau \in [0, T]$ , definido através de:

$$\tau = \psi(t) = \int_{t_0}^t L(\mathbf{x}(\xi), \dot{\mathbf{x}}(\xi)) d\xi = \frac{1}{h} \int_{t_0}^t n(\mathbf{x}(\xi)) \|\dot{\mathbf{x}}(\xi)\| d\xi$$

onde  $T = \int_{t_0}^{t_1} n(\mathbf{x}(t)) \|\dot{\mathbf{x}}(t)\| dt$  é o comprimento óptico total de  $\mathbf{x}(\cdot)$ , então esse  $L$ -raio, com a nova parametrização, satisfaz  $L \equiv 1$  e é também um  $Q$ -raio, com energia total  $E_Q = Q = \frac{1}{2}$ .



Retomemos agora a exposição do formalismo canónico de Rund. Para isso, definamos o chamado **tensor fundamental**  $\mathbf{g}$ , através de:

$$g_{ik}(\mathbf{x}, \dot{\mathbf{x}}) \stackrel{\text{def}}{=} Q_{\dot{x}^i \dot{x}^k}(\mathbf{x}, \dot{\mathbf{x}}) \quad (2.2.13)$$

As funções  $g_{ik}$  são positivamente homogéneas de grau 0, relativamente às variáveis  $\dot{\mathbf{x}}$ , e verificam  $g_{ik} = g_{ki}$ . Além disso, de (2.2.4), deduzimos que:

$$Q(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} g_{ik}(\mathbf{x}, \dot{\mathbf{x}}) \dot{x}^i \dot{x}^k, \quad \forall \mathbf{x}, \dot{\mathbf{x}} \neq \mathbf{0} \quad (2.2.14)$$

Vamos agora impôr a **hipótese essencial** de que:

$$\det(g_{ik}) \neq 0 \quad (2.2.15)$$

e definamos um novo **momento conjugado** ( $Q$ -conjugado) à variável  $x^i$  através de:

$$q_i \stackrel{\text{def}}{=} g_{ik}(\mathbf{x}, \dot{\mathbf{x}}) \dot{x}^k \quad (2.2.16)$$

ou sucintamente:

$$\mathbf{q} = \mathbf{g}(\mathbf{x}, \dot{\mathbf{x}}) \dot{\mathbf{x}} \quad (2.2.17)$$

Notemos que, como  $Q_{\dot{\mathbf{x}}}$  é homogéneo de grau 1 nas variáveis  $\dot{\mathbf{x}}$ , a identidade Euler implica que:

$$Q_{\dot{\mathbf{x}} \dot{\mathbf{x}}} \dot{\mathbf{x}} = Q_{\dot{\mathbf{x}}}, \quad \text{isto é} \quad \sum_k Q_{\dot{x}^i \dot{x}^k}(\mathbf{x}, \dot{\mathbf{x}}) \dot{x}^k = Q_{\dot{x}^i}(\mathbf{x}, \dot{\mathbf{x}}) \quad (2.2.18)$$

e portanto podemos ainda escrever que  $q_i = g_{ik} \dot{x}^k = 2Q_{\dot{x}^i \dot{x}^k} \dot{x}^k = Q_{\dot{x}^i}$ . Como, além disso, por definição,  $Q = \frac{1}{2} L^2$  e  $p_i = L_{\dot{x}^i}$ , vem que  $q_i = Q_{\dot{x}^i} = L L_{\dot{x}^i} = L p_i$ , ou mais sucintamente:

$$\mathbf{q} = Q_{\dot{\mathbf{x}}} = L L_{\dot{\mathbf{x}}} = L \mathbf{p} \quad (2.2.19)$$

Portanto o novo momento conjugado  $\mathbf{q} = Q_{\dot{\mathbf{x}}}$  difere do momento usual  $\mathbf{p} = L_{\dot{\mathbf{x}}}$ , apenas pelo factor multiplicativo  $L$ :

$$\mathbf{q}(\mathbf{x}, \dot{\mathbf{x}}) = Q_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) = L(\mathbf{x}, \dot{\mathbf{x}}) L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) = L(\mathbf{x}, \dot{\mathbf{x}}) \mathbf{p}(\mathbf{x}, \dot{\mathbf{x}}) \quad (2.2.20)$$

Consideremos agora uma  $L$ -extremal  $\mathbf{x}(\tau)$ ,  $0 \leq \tau \leq T$ , devidamente parametrizada para que satisfaça:

$$L(\mathbf{x}(\tau), \dot{\mathbf{x}}(\tau)) \equiv 1 \quad (2.2.21)$$

Como vimos na proposição 2.1, é sempre possível introduzir uma tal parametrização natural. A  $L$ -extremal  $\mathbf{x}(\tau)$ , assim obtida, torna-se automaticamente uma  $Q$ -extremal, isto é, verifica as equações de Euler-Lagrange para o Lagrangeano  $Q$ :

$$-\frac{d}{d\tau} Q_{\dot{\mathbf{x}}} + Q_{\mathbf{x}} = 0 \quad (2.2.22)$$

Introduzindo os  $Q$ -momentos:

$$\mathbf{q} = Q_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}), \quad \Rightarrow \quad \dot{\mathbf{x}} = \dot{\mathbf{x}}(\mathbf{x}, \mathbf{q}) \quad (2.2.23)$$

e atendendo à hipótese fundamental (2.2.15), sabemos que as equações (2.2.22) são equivalentes às equações canônicas de Hamiltoniano:

$$\Phi(\mathbf{x}, \mathbf{q}) = (Q_{\dot{\mathbf{x}}}\dot{\mathbf{x}} - Q)|_{\dot{\mathbf{x}}=\dot{\mathbf{x}}(\mathbf{x}, \mathbf{q})} \quad (2.2.24)$$

isto é, às equações:

$$\begin{cases} \dot{\mathbf{x}} = \Phi_{\mathbf{q}} \\ \dot{\mathbf{q}} = -\Phi_{\mathbf{x}} \end{cases} \quad (2.2.25)$$

Mais detalhadamente - uma  $L$ -extremal  $\mathbf{x}(\tau)$ , parametrizada naturalmente (isto é, devidamente parametrizada para que satisfaça (2.2.21)), pode ser obtida como uma solução  $(\mathbf{x}(\tau), \mathbf{q}(\tau))$  das equações canônicas associadas ao Hamiltoniano  $\Phi$ , definido por (2.2.24), onde  $\mathbf{q}(\tau) = Q_{\dot{\mathbf{x}}}(\mathbf{x}(\tau), \dot{\mathbf{x}}(\tau))$ .

Notemos agora que, como  $Q_{\dot{\mathbf{x}}}\dot{\mathbf{x}} - Q = Q$ , então:

$$\Phi(\mathbf{x}, \mathbf{q}) = Q(\mathbf{x}, \dot{\mathbf{x}}), \quad \text{onde} \quad \mathbf{q} = Q_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) \Leftrightarrow \dot{\mathbf{x}} = \Phi_{\mathbf{q}}(\mathbf{x}, \mathbf{q}) \quad (2.2.26)$$

Definamos finalmente o **Hamiltoniano**  $H$ , correspondente ao Lagrangeano  $L$ , através de:

$$H(\mathbf{x}, \mathbf{q}) = L(\mathbf{x}, \dot{\mathbf{x}}), \quad \text{onde} \quad \dot{\mathbf{x}} = \Phi_{\mathbf{q}}(\mathbf{x}, \mathbf{q}) \quad (2.2.27)$$

Como  $\mathbf{q} = LL_{\dot{\mathbf{x}}}$ , vem que:

$$L(\mathbf{x}, \dot{\mathbf{x}}) = H(\mathbf{x}, L(\mathbf{x}, \dot{\mathbf{x}}) L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}})) = L(\mathbf{x}, \dot{\mathbf{x}}) H(\mathbf{x}, L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}))$$

já que  $H$  é homogêneo positivo na variável  $\mathbf{q}$ . Concluimos portanto que:

$$H(\mathbf{x}, L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}})) \equiv 1 \quad \text{se } L > 0 \quad (2.2.28)$$

Como  $\Phi = \frac{1}{2}H^2$ , as equações canônicas (2.2.25) podem ser escritas na forma:

$$\begin{cases} \dot{\mathbf{x}} = H H_{\mathbf{q}} \\ \dot{\mathbf{q}} = -H H_{\mathbf{x}} \end{cases} \quad (2.2.29)$$

Podemos pois enunciar a seguinte proposição:

- **♣ Proposição 2.2** ... *Suponhamos que  $L$  é um Lagrangeano paramétrico definido positivo. Então:*

$$\begin{aligned} Q(\mathbf{x}, \dot{\mathbf{x}}) &= \frac{1}{2}L^2(\mathbf{x}, \dot{\mathbf{x}}) \\ \mathbf{q} &= Q_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) = L(\mathbf{x}, \dot{\mathbf{x}}) L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) = L(\mathbf{x}, \dot{\mathbf{x}})\mathbf{p} \\ \Phi(\mathbf{x}, \mathbf{q}) &= Q(\mathbf{x}, \dot{\mathbf{x}}), \quad \forall \mathbf{q} = Q_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}), \quad \forall \dot{\mathbf{x}} = \Phi_{\mathbf{q}}(\mathbf{x}, \mathbf{q}) \\ H(\mathbf{x}, \mathbf{q}) &= L(\mathbf{x}, \dot{\mathbf{x}}), \quad \forall \mathbf{q} = Q_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}), \quad \forall \dot{\mathbf{x}} = \Phi_{\mathbf{q}}(\mathbf{x}, \mathbf{q}) \\ H(\mathbf{x}, L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}})) &\equiv 1 \end{aligned} \quad (2.2.30)$$

Seja  $\mathbf{x}(\tau)$  uma  $L$ -extremal regular parametrizada naturalmente, de tal forma a satisfazer a condição:

$$L(\mathbf{x}(\tau), \dot{\mathbf{x}}(\tau)) \equiv 1$$

Então  $\mathbf{x}(\tau)$  é uma  $Q$ -extremal e  $(\mathbf{x}(\tau), \mathbf{q}(\tau))$ , onde  $\mathbf{q}(\tau) = Q_{\dot{\mathbf{x}}}(\mathbf{x}(\tau), \dot{\mathbf{x}}(\tau))$ , satisfaz as equações canónicas:

$$\begin{cases} \dot{\mathbf{x}} = \Phi_{\mathbf{q}} \\ \dot{\mathbf{q}} = -\Phi_{\mathbf{x}} \end{cases} \quad \text{ou} \quad \begin{cases} \dot{\mathbf{x}} = H H_{\mathbf{q}} \\ \dot{\mathbf{q}} = -H H_{\mathbf{x}} \end{cases} \quad (2.2.31)$$

♣.

♣ **Exemplo 2.3 (O Hamiltoniano óptico)** ... Como vimos o Lagrangeano óptico é:

$$L(\mathbf{x}, \dot{\mathbf{x}}) = n(\mathbf{x}) \|\dot{\mathbf{x}}\|$$

e o  $Q$ -Lagrangeano é:

$$Q(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} n^2(\mathbf{x}) \dot{\mathbf{x}}^2$$

O Hamiltoniano  $\Phi$ , correspondente, via transformação de Legendre, ao Lagrangeano  $Q$  é:

$$\Phi(\mathbf{x}, \mathbf{q}) = \frac{\mathbf{q}^2}{2n^2} = \frac{p^2 + q^2 + r^2}{2n^2(x, y, z)} \quad (2.2.32)$$

De facto  $\mathbf{q} = Q_{\dot{\mathbf{x}}} = n^2 \dot{\mathbf{x}} \Rightarrow \dot{\mathbf{x}} = \frac{\mathbf{q}}{n^2}$ , ou mais explicitamente:

$$\begin{aligned} p &= \frac{\partial Q}{\partial \dot{x}} = n^2 \dot{x} \Rightarrow \dot{x} = \frac{p}{n^2} \\ q &= \frac{\partial Q}{\partial \dot{y}} = n^2 \dot{y} \Rightarrow \dot{y} = \frac{q}{n^2} \\ r &= \frac{\partial Q}{\partial \dot{z}} = n^2 \dot{z} \Rightarrow \dot{z} = \frac{r}{n^2} \end{aligned}$$

(note que  $\mathbf{q} = LL_{\dot{\mathbf{x}}}$ ). Portanto:

$$\begin{aligned} \Phi(\mathbf{x}, \mathbf{q}) &= (\mathbf{q}\dot{\mathbf{x}} - Q(\mathbf{x}, \dot{\mathbf{x}}))|_{\dot{\mathbf{x}}=\frac{\mathbf{q}}{n^2}} \\ &= Q(\mathbf{x}, \dot{\mathbf{x}})|_{\dot{\mathbf{x}}=\frac{\mathbf{q}}{n^2}} \\ &= \frac{\mathbf{q}^2}{2n^2} \end{aligned} \quad (2.2.33)$$

O Hamiltoniano óptico  $H$ , correspondente ao Lagrangeano óptico  $L$  é, por definição:

$$H(\mathbf{x}, \mathbf{q}) = L(\mathbf{x}, \dot{\mathbf{x}})|_{\dot{\mathbf{x}}=\Phi_{\mathbf{q}}(\mathbf{x}, \mathbf{q})} = (n(\mathbf{x})\|\dot{\mathbf{x}}\|)|_{\dot{\mathbf{x}}=\mathbf{q}/n^2} = \frac{\mathbf{q}}{n(\mathbf{x})} \quad (2.2.34)$$

As equações canónicas (2.2.31) são:

$$\begin{cases} \dot{\mathbf{x}} = HH_{\mathbf{q}} \\ \dot{\mathbf{q}} = -HH_{\mathbf{x}} \end{cases} \quad \text{isto é} \quad \begin{cases} \dot{\mathbf{x}} = \frac{\mathbf{q}}{n^2(\mathbf{x})} \\ \dot{\mathbf{q}} = \frac{\mathbf{q}^2}{n^3(\mathbf{x})} \nabla n(\mathbf{x}) \end{cases} \quad (2.2.35)$$

ou ainda:

$$\begin{cases} \dot{x} = \frac{p}{n^2} \\ \dot{y} = \frac{q}{n^2} \\ \dot{z} = \frac{r}{n^2} \\ \dot{p} = \frac{n_x(p^2+q^2+r^2)}{n^3} \\ \dot{q} = \frac{n_y(p^2+q^2+r^2)}{n^3} \\ \dot{r} = \frac{n_z(p^2+q^2+r^2)}{n^3} \end{cases} \quad (2.2.36)$$

Por (2.2.30) tem-se que:

$$H(\mathbf{x}, \mathbf{p}) \equiv 1, \quad \forall \mathbf{p} = L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}})$$

que, neste caso, é óbvio já que  $H(\mathbf{x}, L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}})) = \dot{\mathbf{x}}/\|\dot{\mathbf{x}}\| = 1$ .

Por outro lado, o Lagrangeano  $Q$  representa a energia cinética associada à métrica Riemanniana:

$$d\sigma = n(x, y, z) \sqrt{dx^2 + dy^2 + dz^2} = n ds \quad (2.2.37)$$

em  $\mathbb{R}^3$ . Portanto, as curvas integrais de  $X_{\Phi}$ , quando parametrizadas naturalmente, projectam-se em  $\mathbb{R}^3$ , exactamente nas geodésicas da métrica  $d\sigma$ , que, como sabemos, são as curvas que minimizam (localmente) o comprimento de arco correspondente à métrica  $d\sigma$ . Este comprimento de arco é o comprimento óptico da curva, e portanto obtemos mais uma vez o princípio de Fermat: “*Os raios de luz são as curvas que minimizam (localmente) o comprimento óptico*  $\int d\sigma = \int n ds$ .”



## 2.3 Campos de Meyer

Consideremos de novo um funcional do tipo:

$$\mathfrak{F}[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \quad (2.3.1)$$

em que o Lagrangeano  $L : T\mathbb{R}^n \rightarrow \mathbb{R}$  não depende do parâmetro e é homogéneo positivo de grau 1 em  $\dot{\mathbf{x}}$ .

Consideremos um domínio simplesmente conexo  $\mathcal{U} \subseteq \mathbb{R}_{\mathbf{x}}^n$  e uma família a  $(n - 1)$ -parâmetros de curvas:

$$\mathbf{x} = \mathbf{x}(t, \boldsymbol{\alpha}), \quad t \in I(\boldsymbol{\alpha}), \quad \boldsymbol{\alpha} \in \mathcal{A} \subseteq \mathbb{R}_{\boldsymbol{\alpha}}^{n-1} \quad (2.3.2)$$

onde supomos que os parâmetros  $\boldsymbol{\alpha} = (\alpha^1, \dots, \alpha^{n-1})$  variam num aberto  $\mathcal{A}$  do espaço dos parâmetros  $\mathbb{R}_{\boldsymbol{\alpha}}^{n-1}$ , e, para cada  $\boldsymbol{\alpha}$ ,  $I(\boldsymbol{\alpha})$  é um intervalo em  $\mathbb{R}_t$ , de tal forma que:

$$\Lambda \stackrel{\text{def}}{=} \{(t, \boldsymbol{\alpha}) : \boldsymbol{\alpha} \in \mathcal{A}, t \in I(\boldsymbol{\alpha})\} \quad (2.3.3)$$

é um domínio simplesmente conexo em  $\mathbb{R} \times \mathbb{R}^{n-1} = \mathbb{R}^n$ .

Quando a aplicação  $\Psi : \Lambda \longrightarrow \mathcal{U} \subseteq \mathbb{R}_x^n$ , definida por:

$$\Psi(t, \boldsymbol{\alpha}) = \mathbf{x}(t, \boldsymbol{\alpha}) \quad (2.3.4)$$

é um difeomorfismo, diz-se que  $\Psi$  define um **feixe de curvas** em  $\mathcal{U}$ . O feixe  $\Psi$  diz-se um **feixe de extremais** se, para cada  $\boldsymbol{\alpha}$ , a curva  $\mathbf{x}(\cdot, \boldsymbol{\alpha})$  fôr uma extremal (solução das equações de Euler-Lagrange).

Figure 2.3: Feixe de curvas

Fixemos agora um qualquer ponto  $\mathbf{x} \in \mathcal{U}$ . Então, por  $\mathbf{x}$  passa uma única curva do feixe  $\Psi$ , a que corresponde um valor (único) do parâmetro  $\boldsymbol{\alpha} = \boldsymbol{\alpha}(\mathbf{x}) \in \mathcal{A}$ , e ainda um único instante  $t = t(\mathbf{x}) \in I(\boldsymbol{\alpha}(\mathbf{x}))$ , tal que:

$$\mathbf{x} = \mathbf{x}(t(\mathbf{x}), \boldsymbol{\alpha}(\mathbf{x})) \quad (2.3.5)$$

De facto, como  $\Psi$  é um difeomorfismo a inversa  $\Psi^{-1}$  é também um difeomorfismo, e  $\Psi^{-1}(\mathbf{x}) = (t(\mathbf{x}), \boldsymbol{\alpha}(\mathbf{x})) \in \Lambda$ . Definamos agora um campo de vectores  $\mathfrak{J} \in \mathfrak{X}(\mathcal{U})$ , pondo, para cada  $\mathbf{x} \in \mathcal{U}$ :

$$\mathfrak{J}(\mathbf{x}) \stackrel{\text{def}}{=} \left. \frac{d}{dt} \right|_{t=t(\mathbf{x})} \mathbf{x}(t, \boldsymbol{\alpha}(\mathbf{x})), \quad \mathbf{x} \in \mathcal{U} \subseteq \mathbb{R}_x^n \quad (2.3.6)$$

O campo de vectores  $\mathfrak{J}$  diz-se o **campo de direcções** do feixe de curvas  $\Psi$ . Note que este campo nunca se anula em  $\mathcal{U}$ :  $\mathfrak{J}(\mathbf{x}) \neq \mathbf{0}, \forall \mathbf{x} \in \mathcal{U}$ . Além disso, as curvas do feixe,  $\mathbf{x}(t, \boldsymbol{\alpha})$  são soluções da ODE em  $\mathcal{U}$ :

$$\dot{\mathbf{x}} = \mathfrak{J}(\mathbf{x}) \quad (2.3.7)$$

É claro que se multiplicarmos o campo de direcções  $\mathfrak{J}$ , por uma função  $f(\mathbf{x}) > 0$  obtemos as mesmas direcções, mas agora de um feixe de curvas obtido do anterior por reparametrização das respectivas curvas, e que, por isso, deve ser considerado como equivalente ao primeiro. É muitas vezes útil utilizar certas parametrizações especiais. Quando  $L$  é definido positivo -  $L(\mathbf{x}, \mathbf{v}) > 0, \forall \mathbf{x}, \forall \mathbf{v} \neq \mathbf{0}$  - a mais frequente, é:

$$L(\mathbf{x}, \mathfrak{J}(\mathbf{x})) = 1 \quad (2.3.8)$$

o que sempre se consegue multiplicando o campo de direcções por uma função  $f(\mathbf{x}) > 0$  conveniente. Um tal campo diz-se um **campo normal** (ou **normalizado**). Estes campos podem também ser caracterizados pela equação:

$$L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) \mathfrak{J}(\mathbf{x}) = 1 \quad (2.3.9)$$

atendendo à identidade de Euler  $L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}) \mathbf{v} = L(\mathbf{x}, \mathbf{v})$ .

Consideremos agora uma dada extremal  $\hat{\mathbf{x}} : [t_0, t_1] \rightarrow \mathbb{R}^n$  (solução das equações de Euler-Lagrange), e passemos a discutir condições que garantam que  $\hat{\mathbf{x}}$  é um mínimo (local) de  $\mathfrak{F}$ .

Assim, suponhamos que é possível mergulhar a extremal  $\hat{\mathbf{x}}$  num feixe de extremais, com campo de direcções  $\mathfrak{J}(\mathbf{x})$ , e que tem a propriedade adicional seguinte - para todo o ponto  $\mathbf{x}$ , o Lagrangeano  $L = L(\mathbf{x}, \mathbf{v})$  anula-se sempre que  $\mathbf{v} = \mathfrak{J}(\mathbf{x})$  e é estritamente positivo sempre que  $\mathbf{v} \neq \mathfrak{J}(\mathbf{x})$ :

$$\forall \mathbf{x} \in \mathcal{U} : \quad \begin{cases} L(\mathbf{x}, \mathfrak{J}(\mathbf{x})) = 0 \\ L(\mathbf{x}, \mathbf{v}) > 0, \text{ para } \mathbf{v} \neq \mathfrak{J}(\mathbf{x}), \end{cases} \quad \mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n \quad (2.3.10)$$

Se isto fôr possível então  $\hat{\mathbf{x}}$  é um mínimo forte (local) de  $\mathfrak{F}$ . De facto, se  $\gamma : [t_0, t_1] \rightarrow \mathbb{R}^n$  é uma curva qualquer, com as mesmas extremidades de  $\hat{\mathbf{x}}$ , então (2.3.10) implica que:

$$\mathfrak{F}[\hat{\mathbf{x}}(\cdot)] = 0, \quad \mathfrak{F}[\gamma(\cdot)] > 0$$

uma vez que  $\dot{\hat{\mathbf{x}}}(t) = \mathfrak{J}(\hat{\mathbf{x}}(t))$ ,  $\forall t$ , e portanto:

$$\mathfrak{F}[\hat{\mathbf{x}}(\cdot)] < \mathfrak{F}[\gamma(\cdot)]$$

para toda a curva  $\gamma \neq \hat{\mathbf{x}}$ .

No entanto, em geral não é possível conseguir exactamente as condições (2.3.10). Para superar esta dificuldade, substituímos  $L$  por um Lagrangeano (gauge) equivalente  $\tilde{L}$ , definido por:

$$\tilde{L}(\mathbf{x}, \dot{\mathbf{x}}) \stackrel{\text{def}}{=} L(\mathbf{x}, \dot{\mathbf{x}}) - S_{\mathbf{x}}(\mathbf{x})\dot{\mathbf{x}} \quad (2.3.11)$$

onde  $S : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função suave.

Notemos que:

$$\begin{aligned} \mathfrak{F}_{\tilde{L}}[\mathbf{x}(\cdot)] &= \int_{t_0}^{t_1} [L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) - S_{\mathbf{x}}(\mathbf{x}(t))\dot{\mathbf{x}}(t)] dt \\ &= \int_{t_0}^{t_1} [L(\mathbf{x}(t), \dot{\mathbf{x}}(t))] dt - \int_{t_0}^{t_1} \left[ \frac{dS(\mathbf{x}(t))}{dt} \right] dt \\ &= \mathfrak{F}_L[\mathbf{x}(\cdot)] + S(\mathbf{x}(t_0)) - S(\mathbf{x}(t_1)) \end{aligned} \quad (2.3.12)$$

e portanto  $L$  e  $\tilde{L}$  têm exactamente as mesmas extremais. Daí o termos chamado a dois Lagrangeanos  $L$  e  $\tilde{L}$ , relacionados através da condição (2.3.11), **Lagrangeanos equivalentes**. Aliás, tem-se que:

$$\begin{aligned} -\frac{d}{dt}\tilde{L}_{\dot{\mathbf{x}}} + \tilde{L}_{\mathbf{x}} &= -\frac{d}{dt}(L_{\dot{\mathbf{x}}} - S_{\mathbf{x}}) + L_{\mathbf{x}} - S_{\mathbf{xx}}\dot{\mathbf{x}} \\ &= -\frac{d}{dt}L_{\dot{\mathbf{x}}} + L_{\mathbf{x}} \end{aligned}$$

o que mostra mais uma vez que  $L$  e  $\tilde{L}$  têm exactamente as mesmas extremais.

Suponhamos agora que conseguimos encontrar um Lagrangeano  $\tilde{L}$ , (gauge) equivalente a  $L$ , que verifica a propriedade acima descrita, isto é:

$$\forall \mathbf{x} \in \mathcal{U} : \quad \begin{cases} \tilde{L}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) = 0 \\ \tilde{L}(\mathbf{x}, \mathbf{v}) > 0, \text{ para } \mathbf{v} \neq \mathfrak{J}(\mathbf{x}), \end{cases} \quad \mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n \quad (2.3.13)$$

Então, se  $\gamma : [t_0, t_1] \rightarrow \mathbb{R}^n$  é uma curva qualquer, tal que:

$$S(\gamma(t_0)) = S(\hat{\mathbf{x}}(t_0)), \quad S(\gamma(t_1)) = S(\hat{\mathbf{x}}(t_1)) \quad (2.3.14)$$

virá que:

$$0 = \mathfrak{F}_{\tilde{L}}[\hat{\mathbf{x}}(\cdot)] < \mathfrak{F}_{\tilde{L}}[\gamma(\cdot)]$$

isto é:

$$0 = \mathfrak{F}_L[\hat{\mathbf{x}}(\cdot)] - S(\hat{\mathbf{x}}(t_1)) + S(\hat{\mathbf{x}}(t_0)) < \mathfrak{F}_L[\gamma(\cdot)] - S(\gamma(t_1)) + S(\gamma(t_0))$$

e portanto:

$$0 = \mathfrak{F}_L[\hat{\mathbf{x}}(\cdot)] < \mathfrak{F}_L[\gamma(\cdot)]$$

atendendo às condições (2.3.14). Portanto  $\hat{\mathbf{x}}$  será um mínimo forte (local) de  $\mathfrak{F}_L$ .

Resumindo - a solução que acabamos de propôr, e que se deve a Carathéodory<sup>2</sup>, consiste no seguinte: *encontrar um feixe de extremais  $\mathbf{x}(t, \boldsymbol{\alpha})$ , tal que  $\hat{\mathbf{x}}(\cdot) = \mathbf{x}(\cdot, \hat{\boldsymbol{\alpha}})$ , para algum  $\hat{\boldsymbol{\alpha}} \in \mathcal{A}$ , e cujo campo de direcções  $\mathfrak{J} \in \mathfrak{X}(\mathcal{U})$  satisfaça as condições (2.3.13), que, em termos de  $L$ , se escrevem na forma:*

$$\forall \mathbf{x} \in \mathcal{U} : \quad \begin{cases} L(\mathbf{x}, \mathfrak{J}(\mathbf{x})) = S_{\mathbf{x}}(\mathbf{x})\mathfrak{J}(\mathbf{x}) \\ L(\mathbf{x}, \mathbf{v}) > S_{\mathbf{x}}(\mathbf{x})\mathbf{v}, \text{ para } \mathbf{v} \neq \mathfrak{J}(\mathbf{x}), \end{cases} \quad \mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n \quad (2.3.15)$$

para alguma função suave  $S : \mathcal{U} \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ .

Atendendo a que  $L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}) \mathbf{v} = L(\mathbf{x}, \mathbf{v})$ , a primeira equação em (2.3.15) pode ser escrita na forma:

$$L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x}))\mathfrak{J}(\mathbf{x}) = S_{\mathbf{x}}(\mathbf{x})\mathfrak{J}(\mathbf{x})$$

e, como  $\mathfrak{J}(\mathbf{x}) \neq \mathbf{0}$ , ainda na forma equivalente:

$$L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) = S_{\mathbf{x}}(\mathbf{x}) \quad (2.3.16)$$

<sup>2</sup>é a chamada “Carathéodory royal road for field theory”.

a que se chama-se **equação de Carathéodory** (paramétrica).

Na discussão seguinte vamos supôr que  $L$  é um Lagrangeano paramétrico homogéneo definido positivo, isto é, para todo o  $\mathbf{x} \in \mathbb{R}^n$ :

$$L(\mathbf{x}, \mathbf{v}) > 0, \quad \forall \mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n \text{ tal que } \mathbf{v} \neq \mathbf{0} \quad (2.3.17)$$

- **♣ Definição 2.1** ... Um **campo extremal normal** (ou **campo de Meyer**) para um problema variacional homogéneo, com Lagrangeano definido positivo, é definido por um par  $(S, \mathfrak{J})$ , onde  $S : \mathcal{U} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função real e  $\mathfrak{J} \in \mathfrak{X}(\mathcal{U})$  é um campo de vectores não nulos, ambos  $C^\infty$ , que verificam a equação de Carathéodory:

$$L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) = S_{\mathbf{x}}(\mathbf{x}) \quad (2.3.18)$$

e ainda a condição de normalização:

$$L(\mathbf{x}, \mathfrak{J}(\mathbf{x})) \equiv 1, \quad \forall \mathbf{x} \quad (2.3.19)$$

A função  $S$  diz-se a **iconal** ou **função distância** do campo de Meyer.

♣.

Como  $L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}) \mathbf{v} = L(\mathbf{x}, \mathbf{v})$ , pela homogeneidade de  $L$ , a condição de normalização (2.3.19) pode ainda ser escrita na forma:

$$L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) \mathfrak{J}(\mathbf{x}) \equiv 1 \quad (2.3.20)$$

o que, conjugado com a equação de Carathéodory (2.3.18), implica a condição:

$$S_{\mathbf{x}}(\mathbf{x}) \mathfrak{J}(\mathbf{x}) = dS(\mathbf{x})(\mathfrak{J}(\mathbf{x})) \equiv 1, \quad \forall \mathbf{x} \quad (2.3.21)$$

Notemos que estas condições implicam que, se  $t \mapsto \mathbf{x}(t)$  é uma curva integral de  $\mathfrak{J}$ , isto é,  $\dot{\mathbf{x}}(t) = \mathfrak{J}(\mathbf{x}(t))$ ,  $\forall t$ , então é válida a seguinte “normalização”:

$$\begin{aligned} \frac{d}{dt} S(\mathbf{x}(t)) &= dS_{\mathbf{x}(t)}(\dot{\mathbf{x}}(t)) \\ &= dS_{\mathbf{x}(t)}(\mathfrak{J}(\mathbf{x}(t))) \\ &= \mathfrak{J}S(\mathbf{x}(t)) \\ &= 1 \\ &= L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) \end{aligned} \quad (2.3.22)$$

- **♣ Teorema 2.1** ... Para cada  $\mathbf{x} \in \mathcal{U}$ , consideremos o hiperplano afim  $\mathcal{H}_{\mathbf{x}}$ , em  $T_{\mathbf{x}}\mathbb{R}^n$ , definido por:

$$\mathcal{H}_{\mathbf{x}} = \{ \mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n : dS(\mathbf{x})(\mathbf{v}) = 1 \} \quad (2.3.23)$$

e suponhamos que  $\mathfrak{J}(\mathbf{x})$  é um mínimo local para a restrição da função:

$$\mathbf{v} \longmapsto L(\mathbf{x}, \mathbf{v}), \quad \mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n \quad (2.3.24)$$

ao hiperplano afim  $\mathcal{H}_{\mathbf{x}}$ .

Seja  $t \mapsto \mathbf{x}(t)$ ,  $0 \leq t \leq a$ , uma curva integral de  $\mathfrak{J}$  e  $t \mapsto \gamma(t)$ ,  $0 \leq t \leq a$ , uma curva próxima<sup>3</sup>, tal que  $S(\mathbf{x}(0)) = S(\gamma(0))$  e  $S(\mathbf{x}(a)) = S(\gamma(a))$ . Então:

$$\mathfrak{F}[\mathbf{x}(\cdot)] \leq \mathfrak{F}[\gamma(\cdot)]$$

Se, além disso, para cada  $\mathbf{x} \in \mathcal{U}$ ,  $\mathfrak{J}(\mathbf{x})$  é o único mínimo da função (2.3.24), então  $\mathfrak{F}[\mathbf{x}(\cdot)] < \mathfrak{F}[\gamma(\cdot)]$ , a não ser que  $\gamma$  seja uma reparametrização de  $\mathbf{x}$ .

**Dem.:** Podemos supôr, sem perda de generalidade, que:

$$S(\mathbf{x}(0)) = 0 \quad \text{e} \quad S(\mathbf{x}(a)) = 1$$

Como  $\mathbf{x}(\cdot)$  é curva integral de  $\mathfrak{J}$ , sabemos que  $\dot{\mathbf{x}} = \mathfrak{J}(\mathbf{x})$  e pela normalização (2.3.21),  $S_{\mathbf{x}}(\mathbf{x})\dot{\mathbf{x}} \equiv 1$ , isto é,  $\frac{d}{dt}S(\mathbf{x}(t)) = 1$ . Portanto  $S(\mathbf{x}(t)) = t$ , isto é, a curva integral  $\mathbf{x}(\cdot)$  pode ser parametrizada pelos valores do nível das hipersuperfícies de nível de  $S$  que intersecta (as sucessivas hipersuperfícies de nível são as “frentes de onda” e as curvas integrais de  $\mathfrak{J}$  são os “raios”).

Se  $\frac{d}{dt}S(\gamma(t)) > 0$ , podemos também mudar a parametrização de  $\gamma$  de tal forma a que  $S(\gamma(t)) = t$ , isto é, a curva  $\gamma$  pode também ser parametrizada pelos valores do nível das hipersuperfícies de nível de  $S$  que intersecta. Sendo assim, diremos que a curva  $\gamma$  está próxima da curva  $\mathbf{x}$  se:

- $\frac{d}{dt}S(\gamma(t)) > 0$ , isto é,  $S$  é uma função estritamente crescente ao longo de  $\gamma$ .
- $\forall t \in [0, a]$ ,  $\dot{\gamma}(t)$ , que agora satisfaz a condição  $dS(\dot{\gamma}(t)) = 1$ , está suficientemente próximo de  $\mathfrak{J}(\gamma(t))$ , de tal forma que  $L(\gamma(t), \dot{\gamma}(t)) \geq L(\gamma(t), \mathfrak{J}(\gamma(t)))$ .

Com estas hipóteses, deduzimos então que:

$$\begin{aligned} L(\gamma(t), \dot{\gamma}(t)) &\geq L(\gamma(t), \mathfrak{J}(\gamma(t))) \\ &= 1 = L(\mathbf{x}(t), \mathfrak{J}(\mathbf{x}(t))) = L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) \end{aligned}$$

e portanto:

$$\begin{aligned} \mathfrak{F}[\gamma(\cdot)] &= \int_0^a L(\gamma(t), \dot{\gamma}(t)) dt \\ &\geq \int_0^a 1 dt = \int_0^a L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt = \mathfrak{F}[\mathbf{x}(\cdot)] \end{aligned}$$



Podemos interpretar o teorema anterior da seguinte forma - em cada ponto  $\mathbf{x} \in \mathcal{U}$  temos um vector  $\mathfrak{J}(\mathbf{x})$  que representa a **direcção óptima** que o raio deve seguir quando passa em  $\mathbf{x}$ . Como  $\mathbf{x}(\cdot)$  é um raio, i.e., é uma curva integral de  $\mathfrak{J}$ ,  $\mathbf{x}(\cdot)$  é, neste sentido, uma curva optimal - por onde passa, ela segue sempre a direcção óptima. Qualquer outra

<sup>3</sup>num sentido que será claro na demonstração...

curva que parta de  $\mathbf{x}(0)$  para a frente de onda  $\mathcal{W}_a = \{\mathbf{x} : S(\mathbf{x}) \equiv a\}$ , deve violar em algum ponto esta optimalidade e, por isso, dará um valor maior para  $\mathfrak{F}$ .

Notemos ainda o seguinte: a hipótese (2.3.24) do teorema anterior diz que, para cada  $\mathbf{x} \in \mathcal{U}$  fixo,  $\mathfrak{J}(\mathbf{x})$  é um mínimo local para a restrição da função  $v \mapsto L(\mathbf{x}, \mathbf{v})$ , ao hiperplano afim  $\mathcal{H}_{\mathbf{x}} = \{\mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n : S_{\mathbf{x}}(\mathbf{x})\mathbf{v} = 1\}$ . Portanto, pelo método dos multiplicadores de Lagrange, sabemos que existe um multiplicador  $\lambda = \lambda(\mathbf{x})$  tal que:

$$L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) = \lambda(\mathbf{x}) S_{\mathbf{x}}(\mathbf{x}) \quad (2.3.25)$$

Aplicando ambos os membros ao vector  $\mathfrak{J}(\mathbf{x})$ , e usando a identidade Euler para  $L$ , obtemos:

$$\begin{aligned} \lambda(\mathbf{x}) S_{\mathbf{x}}(\mathbf{x}) \mathfrak{J}(\mathbf{x}) &= \lambda(\mathbf{x}) dS(\mathbf{x})(\mathfrak{J}(\mathbf{x})) \\ &= \lambda(\mathbf{x}) \\ &= L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) \mathfrak{J}(\mathbf{x}) \\ &= L(\mathbf{x}, \mathfrak{J}(\mathbf{x})) \\ &= 1 \end{aligned} \quad (2.3.26)$$

donde se deduz que  $\lambda(\mathbf{x}) = 1$  e a equação (2.3.25) fica:

$$L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) = S_{\mathbf{x}}(\mathbf{x}) \quad (2.3.27)$$

que não é mais do que a já conhecida equação de Carathéodory. Obtemos pois uma nova interpretação para esta equação.

Podemos ainda escrever a equação de Carathéodory na forma mais geométrica que passamos a expôr. Em primeiro lugar, ao campo de direcções  $\mathfrak{J}$ , associamos o chamado **campo de momentos**  $\mathcal{P}$ , pondo, para cada  $\mathbf{x} \in \mathbb{R}^n$ :

$$\mathcal{P}(\mathbf{x}) \stackrel{\text{def}}{=} L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) \in T_{\mathbf{x}}^*\mathbb{R}^n \quad (2.3.28)$$

A correspondência  $\mathbf{x} \mapsto \mathcal{P}(\mathbf{x})$ , pode ser vista como uma secção  $\mathcal{P} : \mathcal{U} \subseteq \mathbb{R}^n \longrightarrow T^*\mathbb{R}^n$ :

$$\mathcal{P} : \mathbf{x} \longmapsto (\mathbf{x}, \mathcal{P}(\mathbf{x})), \quad \text{onde} \quad \mathcal{P}(\mathbf{x}) = L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) \quad (2.3.29)$$

ou alternativamente como a 1-forma associada ao campo de direcções  $\mathfrak{J} \in \mathfrak{X}(\mathcal{U})$ , via  $\mathbf{v} \mapsto L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v})$ .

Se representamos por  $\theta$  a **forma canónica de Liouville** em  $T^*\mathbb{R}^n$ , que, em coordenadas canónicas é dada por  $\theta = \mathbf{p}d\mathbf{x}$ , vemos que a equação de Carathéodory (2.3.16) pode ser escrita na forma:

$$\mathcal{P}^*\theta = dS \quad (2.3.30)$$

o que significa que a imagem da secção  $\mathcal{P} : \mathbb{R}^n \longrightarrow T^*\mathbb{R}^n$  é uma **subvariedade de Lagrange** em  $T^*\mathbb{R}^n$ , isto é:

$$\mathcal{P}^*(\omega) = 0 \quad (2.3.31)$$

onde  $\omega = d\theta = d\mathbf{p} \wedge d\mathbf{x}$  é a forma simpléctica canónica em  $T^*\mathbb{R}^n$ .

Se  $\gamma : [t_0, t_1] \rightarrow \mathcal{U} \subseteq \mathbb{R}^n$  é uma curva qualquer que une os pontos  $P_0 = \gamma(t_0)$  a  $P_1 = \gamma(t_1)$ , em  $\mathcal{U}$ , temos que:

$$\begin{aligned} S(P_1) - S(P_0) &= \int_{\gamma} dS \\ &= \int_{\gamma} \mathcal{P}^*(\theta) \\ &= \int_{\gamma} \mathcal{P}(\mathbf{x}) d\mathbf{x} \\ &= \int_{\gamma} L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) d\mathbf{x} \end{aligned} \quad (2.3.32)$$

que é o chamado **integral invariante de Hilbert** paramétrico <sup>4</sup>.

Finalmente, recordemos que o Hamiltoniano  $H$  correspondente ao Lagrangeano  $L$ , foi definido na secção anterior através de:

$$H(\mathbf{x}, \mathbf{q}) = L(\mathbf{x}, \dot{\mathbf{x}}), \quad \text{para todo} \quad \dot{\mathbf{x}} = \Phi_{\mathbf{q}}(\mathbf{x}, \mathbf{q}) \Leftrightarrow \mathbf{q} = Q_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) \quad (2.3.34)$$

Aí se viu que:

$$H(\mathbf{x}, L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}})) \equiv 1 \quad (2.3.35)$$

Daí que, em particular, se obtenha:

$$1 = H(\mathbf{x}, L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathcal{J}(\mathbf{x}))) = H(\mathbf{x}, \mathcal{P}(\mathbf{x})) = H(\mathbf{x}, S_{\mathbf{x}}(\mathbf{x}))$$

que é a **equação de Hamilton-Jacobi** paramétrica, ou **equação iconal**:

$$H(\mathbf{x}, S_{\mathbf{x}}) = 1 \quad (2.3.36)$$

♣ **Exemplo 2.4** ... O Lagrangeano óptico é:

$$L(\mathbf{x}, \dot{\mathbf{x}}) = n(\mathbf{x}) \|\dot{\mathbf{x}}\|$$

Portanto:

$$\mathbf{q} = L(\mathbf{x}, \dot{\mathbf{x}}) L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) = n(\mathbf{x}) \|\dot{\mathbf{x}}\| n(\mathbf{x}) \frac{\dot{\mathbf{x}}}{\|\dot{\mathbf{x}}\|} = n^2(\mathbf{x}) \dot{\mathbf{x}}$$

donde:

$$H(\mathbf{x}, \mathbf{q}) = L(\mathbf{x}, \dot{\mathbf{x}})|_{\dot{\mathbf{x}}=\mathbf{q}/n^2} = (n(\mathbf{x}) \|\dot{\mathbf{x}}\|)|_{\dot{\mathbf{x}}=\mathbf{q}/n^2} = \frac{\|\mathbf{q}\|}{n(\mathbf{x})}$$

<sup>4</sup>Aliás este resultado não é inesperado, uma vez que a forma de Poincaré-Cartan, para Lagrangeanos paramétricos homogêneos reduz-se a:

$$\begin{aligned} \theta_L &= L_{\dot{\mathbf{x}}} d\mathbf{x} - E_L dt, \quad \text{onde} \quad E_L = L_{\dot{\mathbf{x}}} \dot{\mathbf{x}} - L \\ &= L_{\dot{\mathbf{x}}}(\mathbf{x}, \dot{\mathbf{x}}) d\mathbf{x} \end{aligned} \quad (2.3.33)$$

já que a energia de  $L$  é  $E_L = L_{\dot{\mathbf{x}}} \dot{\mathbf{x}} - L = 0$ . A forma  $\theta_L$  pode por isso ser considerada como uma 1-forma (semi-básica) em  $T\mathbb{R}^n$ .

Como  $\mathbf{q} = L\mathbf{p}$ , obtemos ainda  $1 = H(\mathbf{x}, \mathbf{p}) = \frac{\|\mathbf{p}\|}{n(\mathbf{x})}$  e, fazendo  $\mathbf{p} = S_{\mathbf{x}}$ , obtemos a equação iconal da óptica geométrica:

$$\|\nabla S\| = n \quad (2.3.37)$$

ou  $(\nabla S)^2 = n^2$ , onde identificamos  $S_{\mathbf{x}}$  com  $\nabla S$ , como é habitual.



## 2.4 Indicatriz. Função excesso. Fórmula de Weierstrass

Na discussão seguinte vamos continuar supôr que  $L$  é um Lagrangeano paramétrico homogéneo definido positivo, isto é, para todo o  $\mathbf{x} \in \mathbb{R}^n$  -  $L(\mathbf{x}, \mathbf{v}) > 0$ ,  $\forall \mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n$ ,  $\mathbf{v} \neq \mathbf{0}$ .

Para cada  $\mathbf{x} \in \mathbb{R}^n$ , definamos a **indicatriz** de  $\mathbf{x}$ , através de:

$$I_{\mathbf{x}} \stackrel{\text{def}}{=} \{ \mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n : L(\mathbf{x}, \mathbf{v}) \equiv 1 \} \quad (2.4.1)$$

Dado um qualquer vector não nulo  $\mathbf{w} \in T_{\mathbf{x}}\mathbb{R}^n$ , existe sempre um escalar  $\lambda = \lambda(\mathbf{w}) > 0$  tal que  $\lambda\mathbf{w} \in I_{\mathbf{x}}$ . De facto, pela homogeneidade de  $L$  e por (2.3.17),  $1 = L(\mathbf{x}, \lambda\mathbf{w}) = \lambda L(\mathbf{x}, \mathbf{w})$ , e basta tomar  $\lambda = 1/L(\mathbf{x}, \mathbf{w}) > 0$ . Portanto, cada raio orientado, partindo da origem de  $T_{\mathbf{x}}\mathbb{R}^n$ , intersecta uma e uma só vez a indicatriz  $I_{\mathbf{x}}$ , e esta é, por isso, um conjunto estrelado relativamente à origem em  $T_{\mathbf{x}}\mathbb{R}^n$  (não necessariamente convexo).

Fixemos agora um ponto  $\mathbf{v}_0 \in I_{\mathbf{x}}$ , de tal forma que:

$$L(\mathbf{x}, \mathbf{v}_0) = 1 \quad (2.4.2)$$

O hiperplano afim, em  $T_{\mathbf{x}}\mathbb{R}^n$ , tangente à indicatriz  $I_{\mathbf{x}}$ , no ponto  $\mathbf{v}_0$ , que designamos por  $\Pi_{\mathbf{v}_0}$ , é dado pela equação (ver a figura 2.4):

$$\Pi_{\mathbf{v}_0} = \{ \mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n : (\mathbf{v} - \mathbf{v}_0)L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0) = 0 \} \quad (2.4.3)$$

Mais uma vez devido à identidade Euler  $L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v})\mathbf{v} = L(\mathbf{x}, \mathbf{v})$ , e ao facto de  $L$  ser definido positivo, a equação para  $\Pi_{\mathbf{v}_0}$  pode ser escrita na forma:

$$\begin{aligned} 0 &= L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0)\mathbf{v} - L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0)\mathbf{v}_0 \\ &= L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0)\mathbf{v} - L(\mathbf{x}, \mathbf{v}_0) \\ &= L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0)\mathbf{v} - 1 \end{aligned}$$

isto é:

$$L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0)\mathbf{v} = 1 \quad (2.4.4)$$

Figure 2.4: Indicatriz

Representando por  $\mathbf{p}_0 = L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0)$ , o covector em  $T_{\mathbf{x}}^*\mathbb{R}^n$  correspondente a  $\mathbf{v}_0 \in T_{\mathbf{x}}\mathbb{R}^n$ , podemos ainda escrever a equação do hiperplano afim  $\Pi_{\mathbf{v}_0}$  na forma:

$$\Pi_{\mathbf{v}_0} = \{\mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n : \mathbf{p}_0\mathbf{v} = 1\}, \quad \text{onde} \quad \mathbf{p}_0 = L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0) \quad (2.4.5)$$

Consideremos agora, para cada  $\mathbf{x}$  fixo, a função  $L(\mathbf{x}, \cdot) : T_{\mathbf{x}}\mathbb{R}^n \cong \mathbb{R}^n \rightarrow \mathbb{R}$ :

$$\dot{\mathbf{x}} \longmapsto L(\mathbf{x}, \dot{\mathbf{x}})$$

O desenvolvimento de Taylor em torno de um ponto  $\mathbf{v}_0 \in T_{\mathbf{x}}\mathbb{R}^n$  é dado por:

$$L(\mathbf{x}, \mathbf{v}_0 + \mathbf{h}) = L(\mathbf{x}, \mathbf{v}_0) + L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0)\mathbf{h} + \mathcal{E}(\mathbf{x}, \mathbf{v}_0, \mathbf{h})$$

Pondo  $\mathbf{v} = \mathbf{v}_0 + \mathbf{h}$ , vem que:

$$\mathcal{E}(\mathbf{x}, \mathbf{v}_0, \mathbf{v}) \stackrel{\text{def}}{=} L(\mathbf{x}, \mathbf{v}) - L(\mathbf{x}, \mathbf{v}_0) - (\mathbf{v} - \mathbf{v}_0)L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0) \quad (2.4.6)$$

que é a chamada **função excesso** de Weierstrass. Como  $L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v})\mathbf{v} = L(\mathbf{x}, \mathbf{v})$ ,  $\forall \mathbf{v}$ , vem que:

$$\begin{aligned} \mathcal{E}(\mathbf{x}, \mathbf{v}_0, \mathbf{v}) &= L(\mathbf{x}, \mathbf{v}) - L(\mathbf{x}, \mathbf{v}_0) - (\mathbf{v} - \mathbf{v}_0)L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0) \\ &= L(\mathbf{x}, \mathbf{v}) - L(\mathbf{x}, \mathbf{v}_0) - \mathbf{v}L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0) + \mathbf{v}_0L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0) \\ &= L(\mathbf{x}, \mathbf{v}) - L(\mathbf{x}, \mathbf{v}_0) - \mathbf{v}L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0) + L(\mathbf{x}, \mathbf{v}_0) \\ &= L(\mathbf{x}, \mathbf{v}) - \mathbf{v}L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0) \\ &= \mathbf{v}[L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}) - L_{\dot{\mathbf{x}}}(\mathbf{x}, \mathbf{v}_0)] \\ &= \mathbf{v}[\mathbf{p}_{\mathbf{v}} - \mathbf{p}_{\mathbf{v}_0}] \end{aligned}$$

Resumindo:

$$\begin{aligned} \mathcal{E}(\mathbf{x}, \mathbf{v}_0, \mathbf{v}) &= L(\mathbf{x}, \mathbf{v}) - \mathbf{p}_{\mathbf{v}_0}\mathbf{v} \\ &= [\mathbf{p}_{\mathbf{v}} - \mathbf{p}_{\mathbf{v}_0}]\mathbf{v} \end{aligned} \quad (2.4.7)$$

Recordando agora que o hiperplano tangente  $\Pi_{\mathbf{v}_0}$ , à indicatriz  $I_{\mathbf{x}}$ , num dos seus pontos  $\mathbf{v}_0$  (de tal forma que  $L(\mathbf{x}, \mathbf{v}_0) = 1$ ), tem por equação:

$$\mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n : \quad \mathbf{p}_{\mathbf{v}_0}\mathbf{v} = 1$$

podemos enunciar a seguinte proposição:

- ♣ Proposição 2.3 ... (i). A condição:

$$\mathcal{E}(\mathbf{x}, \mathbf{v}_0, \mathbf{v}) \geq 0, \quad \forall \mathbf{v} \in I_{\mathbf{x}} \quad (2.4.8)$$

significa que a origem  $\mathbf{0}$  e a indicatriz  $I_{\mathbf{x}} = \{\mathbf{v} : L(\mathbf{x}, \mathbf{v}) = 1\}$  estão contidos no mesmo semi-espaço  $E_{\mathbf{v}_0} = \{\mathbf{v} : \mathbf{p}_{\mathbf{v}_0} \mathbf{v} \leq 1\}$  limitado por  $\Pi_{\mathbf{v}_0}$ . Além disso, se:

$$\mathcal{E}(\mathbf{x}, \mathbf{v}_0, \mathbf{v}) > 0, \quad \forall \mathbf{v} \in I_{\mathbf{x}}, \quad \text{com } \mathbf{v} \neq \mathbf{v}_0 \quad (2.4.9)$$

então  $\Pi_{\mathbf{v}_0}$  intersecta  $I_{\mathbf{x}}$  apenas em  $\mathbf{v}_0$ .

- (ii). A indicatriz  $I_{\mathbf{x}}$  é convexa se e só se:

$$\mathcal{E}(\mathbf{x}, \mathbf{v}_0, \mathbf{v}) \geq 0, \quad \forall \mathbf{v}_0, \mathbf{v} \in I_{\mathbf{x}} \quad (2.4.10)$$

- (iii). A indicatriz  $I_{\mathbf{x}}$  é estritamente convexa se e só se:

$$\mathcal{E}(\mathbf{x}, \mathbf{v}_0, \mathbf{v}) > 0, \quad \forall \mathbf{v}_0, \mathbf{v} \in I_{\mathbf{x}}, \quad \text{com } \mathbf{v} \neq \mathbf{v}_0 \quad (2.4.11)$$

♣.

Figure 2.5: Indicatriz

Consideremos novamente o hiperplano afim  $\mathcal{H}_{\mathbf{x}} = \{\mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n : S_{\mathbf{x}}(\mathbf{x})\mathbf{v} = 1\}$  e a função  $v \mapsto L(\mathbf{x}, \mathbf{v})$ , definida em  $T_{\mathbf{x}}\mathbb{R}^n$ . Como vimos, uma condição necessária para que  $\mathfrak{J}(\mathbf{x})$  seja um mínimo local da restrição  $L(\mathbf{x}, \cdot)|_{\mathcal{H}_{\mathbf{x}}}$  é dada exactamente pela equação de Carathéodory:

$$L_{\hat{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) = S_{\mathbf{x}}(\mathbf{x}) \quad (2.4.12)$$

Aplicando ambos os membros a  $\mathfrak{J}(\mathbf{x})$  e usando a homogeneidade de  $L$ , obtemos:

$$L(\mathbf{x}, \mathfrak{J}(\mathbf{x})) = S_{\mathbf{x}}(\mathbf{x}) \mathfrak{J}(\mathbf{x}) \quad (2.4.13)$$

Quanto à função excesso de Weierstrass, dada por (2.4.6), pondo  $\mathbf{v}_0 = \mathfrak{J}(\mathbf{x})$  em (2.4.7), e usando (2.4.12), obtemos:

$$\begin{aligned} \mathcal{E}(\mathbf{x}, \mathfrak{J}(\mathbf{x}), \mathbf{v}) &= L(\mathbf{x}, \mathbf{v}) - \mathbf{v} L_{\hat{\mathbf{x}}}(\mathbf{x}, \mathfrak{J}(\mathbf{x})) \\ &= L(\mathbf{x}, \mathbf{v}) - \mathbf{v} S_{\mathbf{x}}(\mathbf{x}) \end{aligned} \quad (2.4.14)$$

$\forall \mathbf{x} \in \mathcal{U}, \forall \mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n, \mathbf{v} \neq \mathbf{0}$ .

- **♣ Proposição 2.4 (Fórmula de Weierstrass)** ... Consideremos um feixe de Meyer em  $\mathcal{U} \subseteq \mathbb{R}^n$ , com campo de direcções  $\mathfrak{J} \in \mathfrak{X}(\mathcal{U})$  e iconal  $S$ , e sejam  $\mathbf{x}(\cdot), \gamma(\cdot) : [t_0, t_1] \rightarrow \mathcal{U}$  duas curvas de classe  $C^1$  em  $\mathcal{U}$ , tais que  $\dot{\mathbf{x}}(t) = \mathfrak{J}(\mathbf{x}(t))$ ,  $\forall t$ ,  $S(\mathbf{x}(t_0)) = S(\gamma(t_0))$  e  $S(\mathbf{x}(t_1)) = S(\gamma(t_1))$ . Então:

$$\mathfrak{F}[\gamma(\cdot)] - \mathfrak{F}[\mathbf{x}(\cdot)] = \int_{t_0}^{t_1} \mathcal{E}(\gamma(t), \mathfrak{J}(\gamma(t)), \dot{\gamma}(t)) dt \quad (2.4.15)$$

**Dem.:** Como  $\dot{\mathbf{x}}(t) = \mathfrak{J}(\mathbf{x}(t))$  vem que:

$$\begin{aligned} L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) &= L(\mathbf{x}(t), \mathfrak{J}(\mathbf{x}(t))) = S_{\mathbf{x}}(\mathbf{x}(t)) \mathfrak{J}(\mathbf{x}(t)) && \text{por (2.4.13)} \\ &= S_{\mathbf{x}}(\mathbf{x}(t)) \dot{\mathbf{x}}(t) = \frac{d}{dt} S(\mathbf{x}(t)) \end{aligned}$$

Daí que:

$$\begin{aligned} \mathfrak{F}[\mathbf{x}(\cdot)] &= \int_{t_0}^{t_1} L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt = S(\mathbf{x}(t_1)) - S(\mathbf{x}(t_0)) = S(\gamma(t_1)) - S(\gamma(t_0)) \\ &= \int_{t_0}^{t_1} \frac{d}{dt} S(\gamma(t)) dt = \int_{t_0}^{t_1} S_{\mathbf{x}}(\gamma(t)) \dot{\gamma}(t) dt \end{aligned}$$

Portanto:

$$\begin{aligned} \mathfrak{F}[\gamma(\cdot)] - \mathfrak{F}[\mathbf{x}(\cdot)] &= \int_{t_0}^{t_1} L(\gamma(\cdot), \dot{\gamma}(t)) dt - \int_{t_0}^{t_1} L(\mathbf{x}(\cdot), \dot{\mathbf{x}}(t)) dt \\ &= \int_{t_0}^{t_1} [L(\gamma(\cdot), \dot{\gamma}(t)) - S_{\mathbf{x}}(\gamma(t)) \dot{\gamma}(t)] dt \\ &= \int_{t_0}^{t_1} \mathcal{E}(\gamma(t), \mathfrak{J}(\gamma(t)), \dot{\gamma}(t)) dt \end{aligned} \quad (2.4.16)$$

atendendo a (2.4.14), com  $\mathbf{x} = \gamma(t)$  e  $\mathbf{v} = \dot{\gamma}(t)$ .

♣.

- **♣ Corolário 2.1 (Weierstrass)** ... Seja  $\mathbf{x}(\cdot) : I \rightarrow \mathbb{R}^n$  uma  $L$ -extremal regular e seja  $\mathcal{U}$  uma vizinhança aberta de  $\mathbf{x}(I)$ . Então, se  $\mathbf{x}(\cdot)$  puder ser mergulhada num feixe de Meyer em  $\mathcal{U}$  e se a função excesso de  $L$  fôr não negativa, a extremal  $\mathbf{x}(\cdot)$  minimiza a acção  $\mathfrak{F}$ , entre todas as curvas de classe  $C^1$ , em  $\mathcal{U}$ , que têm as mesmas extremidades de  $\mathbf{x}(\cdot)$ .

## 2.5 Discussão geométrica da equação de Hamilton-Jacobi reduzida

Seja  $\theta$  a 1-forma canónica de Liouville em  $T^*\mathbb{R}^n$ , dada, em coordenadas canónicas  $(\mathbf{x}, \mathbf{p})$ , por<sup>5</sup>:

$$\theta = \mathbf{p}d\mathbf{x} = \sum_{i=1}^n p_i dx^i \quad (2.5.1)$$

<sup>5</sup>podemos substituir  $\mathbb{R}^n$  por uma variedade diferenciável  $M$

e  $\omega = d\theta$  a forma simpléctica associada. Localmente  $\omega = d\mathbf{p} \wedge d\mathbf{x} = \sum_i dp_i \wedge dx^i$ . Recordemos que  $\theta$  se define por:

$$\theta_{(\mathbf{x},\mathbf{p})}(\xi) = \mathbf{p}(d\pi(\xi)), \quad \mathbf{p} \in T_{\mathbf{x}}^*\mathbb{R}^n, \quad \xi \in T_{(\mathbf{x},\mathbf{p})}T^*\mathbb{R}^n \quad (2.5.2)$$

Figure 2.6:

Além disso  $\theta$  satisfaz a propriedade seguinte:

- **♣ Proposição 2.5 (Propriedade tautológica)** ... Se  $\alpha \in \Omega^1\mathbb{R}^n$  é uma 1-forma em  $\mathbb{R}^n$ , interpretada como uma secção  $\alpha : \mathbb{R}^n \rightarrow T^*\mathbb{R}^n$ , do fibrado  $\pi : T^*\mathbb{R}^n \rightarrow \mathbb{R}^n$ , então:

$$\alpha^*\theta = \alpha \quad (2.5.3)$$

(no membro esquerdo desta fórmula,  $\alpha$  é interpretada como uma secção, enquanto que no membro direito é interpretada como uma 1-forma).

**Dem.:** De facto,  $(\alpha^*\theta)_{\mathbf{x}}(\mathbf{v}) = \theta_{(\mathbf{x},\alpha(\mathbf{x}))}(d\alpha_{\mathbf{x}}\mathbf{v}) = \alpha(\mathbf{x})(d\pi d\alpha_{\mathbf{x}}\mathbf{v}) = \alpha_{\mathbf{x}}(\mathbf{v}), \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{v} \in T_{\mathbf{x}}\mathbb{R}^n$ , uma vez que  $\pi \circ \alpha = \text{Id}$ .

♣.

- **♣ Definição 2.2** ... (i). Uma subvariedade imersa  $\varphi : \Lambda \subseteq \mathbb{R}_{\alpha}^k \hookrightarrow T^*\mathbb{R}^n$ , de dimensão  $k \leq n$ , diz-se uma **variedade isotrópica** se  $\varphi^*\omega = 0$ .  
(ii). Uma subvariedade imersa  $\varphi : \Lambda \subseteq \mathbb{R}_{\alpha}^n \hookrightarrow T^*\mathbb{R}^n$ , de dimensão  $n$ , diz-se uma **variedade de Lagrange** ou uma **Lagrangeana** se  $\varphi^*\omega = 0$ .  
(iii). Uma subvariedade imersa  $\varphi : \Lambda \subseteq \mathbb{R}_{\alpha}^n \hookrightarrow T^*\mathbb{R}^n$ , de dimensão  $n$ , diz-se uma **variedade Lagrangeana cónica** se  $\varphi^*\theta = 0$ , onde  $\theta$  é a 1-forma canónica de Liouville em  $T^*\mathbb{R}^n$ .

♣.

É claro que toda a Lagrangeana cónica é uma Lagrangeana. Uma Lagrangeana é uma variedade isotrópica de dimensão maximal (igual a  $n$ ). Isto resulta do seguinte facto de álgebra linear.

- **♣ Lema 2.1** ... Seja  $(V, \omega)$  um espaço vectorial simplético de dimensão  $2n$  e  $S$  um subespaço totalmente isotrópico, isto é,  $\omega(u, v) = 0, \forall u, v \in S$ . Então  $\dim S \leq n$ .

**Dem.:** Seja  $(u, v) \mapsto \langle u|v \rangle$  um produto interno definido positivo em  $V$ , e representemos  $\omega$  através de um operador  $J : V \rightarrow V$ , definido por:

$$\omega(u, v) = \langle u|J(v) \rangle, \quad u, v \in V$$

Como  $\omega$  é não degenerada  $J$  é um isomorfismo linear. Suponhamos que  $\dim S > n + 1$ . Então, como  $\dim(S + J(S)) = \dim S + \dim J(S) - \dim(S \cap J(S))$ , viria que  $\dim(S \cap J(S)) \geq 2$  e portanto  $S \cap J(S) \neq \{0\}$ . Suponhamos então que  $v \in S \cap J(S)$ , com  $v \neq 0$ . Então  $v = J(u)$ , para algum  $u \in S$  e:

$$0 \neq \langle v|v \rangle = \langle v|J(u) \rangle = \omega(v, u) = 0$$

o que é absurdo.



Vejamos agora alguns exemplos.

**♣ Exemplos 2.2** ...

1.  $\Lambda_{\mathbf{x}_0} = \{\mathbf{x}_0\} \times T_{\mathbf{x}_0}^* \mathbb{R}^n = \{(\mathbf{x}_0, \mathbf{p}) : \mathbf{p} \in \mathbb{R}^n\}$ , onde  $\mathbf{x}_0 \in \mathbb{R}^n$  é um ponto fixo, é uma Lagrangeana cônica.
2.  $\Lambda_{\mathbf{p}_0} = \mathbb{R}^n \times \{\mathbf{p}_0\}$ , onde  $\mathbf{p}_0 \in T_{\mathbf{x}}^* \mathbb{R}^n \cong \mathbb{R}^n$  é um covector fixo, é uma Lagrangeana que não é cônica.
3. Seja  $\Gamma \subset \mathbb{R}^n_{\mathbf{x}}$  uma subvariedade de dimensão  $d \leq n - 1$ , e  $\Lambda_{\Gamma}$  o chamado **fibrado conormal** a  $\Gamma$ , definido por:

$$\Lambda_{\Gamma} = \{(\mathbf{x}, \mathbf{p}) : \mathbf{x} \in \Gamma, \quad \mathbf{p} \in T_{\mathbf{x}}^* \mathbb{R}^n, \quad \text{tal que} \quad \langle T_{\mathbf{x}} \Gamma, \mathbf{p} \rangle = \mathbf{p}|_{T_{\mathbf{x}} \Gamma} = 0\} \quad (2.5.4)$$

O exemplo 1. é o caso especial em que  $\Gamma$  se reduz ao ponto  $\mathbf{x}_0$ .  $\Lambda_{\Gamma}$  é uma Lagrangeana cônica.

4. Seja  $\Gamma \subset \mathbb{R}^n_{\mathbf{x}}$  uma subvariedade de codimensão 1, e  $f : \Gamma \rightarrow \mathbb{R}$  uma função dada, com diferencial  $df_{\mathbf{x}} \in T_{\mathbf{x}}^* \Gamma$ , para cada ponto  $\mathbf{x} \in \Gamma$ . Define-se então:

$$\Lambda_{\Gamma, f} = \{(\mathbf{x}, \mathbf{p}) : \mathbf{x} \in \Gamma, \quad \mathbf{p} \in T_{\mathbf{x}}^* \mathbb{R}^n, \quad \text{tal que} \quad \langle T_{\mathbf{x}} \Gamma, \mathbf{p} \rangle = \mathbf{p}|_{T_{\mathbf{x}} \Gamma} = df_{\mathbf{x}}\} \quad (2.5.5)$$

O exemplo 3. é o caso especial em que  $\Gamma$  tem codimensão 1 e  $f \equiv 0$ .  $\Lambda_{\Gamma, f}$  é uma Lagrangeana que não é cônica (a não ser que  $f$  seja constante).



O exemplo seguinte tem importância crucial para o que se segue - consideremos o gráfico de uma 1-forma  $\alpha \in \Omega^1(\mathbb{R}^n_x)$ :

$$\Lambda_\alpha = \{(\mathbf{x}, \mathbf{p}) \in T^*\mathbb{R}^n : \mathbf{p} = \alpha(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n\} \quad (2.5.6)$$

- **♣ Proposição 2.6** ...  $\Lambda_\alpha$  é uma subvariedade Lagrangeana em  $T^*\mathbb{R}^n$  se e só se  $\alpha$  é fechada:  $d\alpha = 0$ .

**Dem.:** De facto, pela propriedade tautológica,  $\alpha^*\theta = \alpha$ , vem que  $d\alpha = d\alpha^*\theta = \alpha^*d\theta = \alpha^*\omega$  e, portanto,  $d\alpha = 0$  sse  $\alpha^*\omega = 0$ , isto é, sse  $\omega|_\Lambda = 0$ .

♣.

Quando  $\Lambda_\alpha$  é uma subvariedade Lagrangiana em  $T^*\mathbb{R}^n$ ,  $\alpha$  é fechada e portanto localmente exacta, isto é, numa vizinhança  $U$ , de cada ponto de  $\mathbb{R}^n_x$ , podemos encontrar uma função  $S : U \rightarrow \mathbb{R}$  tal que  $\alpha = dS$ . Diz-se então que  $S : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  é uma **função geradora** ou uma **iconal** (local) para  $\Lambda_\alpha$ . Pômos então, neste caso,  $\Lambda_\alpha = \Lambda_{dS}$ , em  $\mathcal{U}$ . Em coordenadas:

$$\alpha = dS = S_x dx = \frac{\partial S}{\partial x^i} dx^i, \quad \text{e} \quad \Lambda_\alpha = \Lambda_{dS} = \left\{ (\mathbf{x}, \mathbf{p}) : \mathbf{p} = S_x(\mathbf{x}) = \left( \frac{\partial S}{\partial x^i}(\mathbf{x}) \right)_{i=1, \dots, n} \right\} \quad (2.5.7)$$

Qualquer Lagrangeana  $\Lambda \subset T^*\mathbb{R}^n$ , que (localmente) se projecte difeomòrficamente sobre o espaço de configuração  $\mathbb{R}^n_x$ , é da forma  $\Lambda = \Lambda_\alpha = \Lambda_{dS}$ , para alguma iconal (local)  $S$ .

Por outro lado, dada uma Lagrangeana  $\Lambda \subset T^*\mathbb{R}^n$ , como a forma  $\theta = \mathbf{p}dx$  é fechada em  $\Lambda$ , o **integral de acção reduzida**  $\int_{P_0}^P \mathbf{p} dx$ , definido para todas as curvas contidas em  $\Lambda$ , que partem de um ponto fixo  $P_0 \in \Lambda$ , define uma função localmente unívoca  $W : \Lambda \rightarrow \mathbb{R}$ , a que chamamos uma **fase local** em  $\Lambda$ :

$$W(P) = \int_{P_0}^P \mathbf{p} dx \quad (2.5.8)$$

O integral de linha é calculado ao longo de uma qualquer curva em  $\Lambda$ , que una  $P_0$  a  $P$ . Se  $\Lambda$  (localmente) se projecta difeomòrficamente sobre o espaço de configuração  $\mathbb{R}^n_x$ , então, como vimos, será da forma  $\Lambda = \Lambda_\alpha = \Lambda_{dS}$ , e  $S = W \circ \alpha$  é uma iconal (local) para  $\Lambda$ .

É claro que as transformações canónicas transformam Lagrangeanas em Lagrangeanas. Por exemplo, a transformação  $(\mathbf{x}, \mathbf{p}) \mapsto (\mathbf{p}, -\mathbf{x})$  permuta  $\Lambda_{\mathbf{x}_0}$  com  $\Lambda_{\mathbf{p}_0}$ .

Suponhamos agora dado um Hamiltoniano  $H \in C^\infty(T^*\mathbb{R}^n)$ , que não depende do parâmetro  $t$ . Como sabemos, existe conservação de energia - uma curva integral  $(\mathbf{x}(t), \mathbf{p}(t))$  do campo Hamiltoniano  $X_H$ , isto é, uma solução das equações canónicas:

$$\begin{cases} \dot{\mathbf{x}} = H_{\mathbf{p}} \\ \dot{\mathbf{p}} = -H_{\mathbf{x}} \end{cases} \quad (2.5.9)$$

Figure 2.7:

está sempre contida num mesmo nível de energia:

$$\Sigma_h \stackrel{\text{def}}{=} \{(\mathbf{x}, \mathbf{p}) \in T^*\mathbb{R}^n : H(\mathbf{x}, \mathbf{p}) \equiv h\}, \quad h \text{ constante} \quad (2.5.10)$$

Nestas condições, temos então a seguinte proposição:

- **♣ Proposição 2.7 ... (i).** *O campo Hamiltoniano  $X_H = H_{\mathbf{p}}\partial_{\mathbf{x}} - H_{\mathbf{x}}\partial_{\mathbf{p}}$  é tangente à hipersuperfície de nível  $\Sigma_h$ .*

(ii).

$$\omega(X_H, \xi) = 0, \quad \forall \xi \in T\Sigma_h \quad (2.5.11)$$

(iii). *Seja  $\Lambda$  uma Lagrangeana em  $T^*\mathbb{R}^n$ , contida na hipersuperfície de nível  $\Sigma_h$ . Então o campo Hamiltoniano  $X_H$  é tangente a  $\Lambda$  e, em particular,  $\Lambda$  contém toda a curva integral de  $X_H$ , sempre que esta intersecta  $\Lambda$  em algum ponto.*

**Dem.:** Para mostrar (i)., temos que provar que  $dH(X_H) = 0$ . Mas, por definição,  $i_{X_H}\omega = dH$ . Portanto  $0 = \omega(X_H, X_H) = (i_{X_H}\omega)(X_H) = dH(X_H)$ .

Por outro lado,  $dH(\xi) = 0, \forall \xi \in T\Sigma_h$ , uma vez que  $H$  é constante em  $\Sigma_h$ . Logo  $0 = dH(\xi) = i_{X_H}\omega(\xi) = \omega(X_H, \xi)$ , o que mostra (ii).

Mostremos agora (iii). Seja  $P \in \Lambda \subset \Sigma_h$ , e seja  $\{\xi_1, \dots, \xi_n\}$  uma base para o espaço tangente  $T_P\Lambda \subset T_P\Sigma_h$ . Por (ii)., sabemos que  $\omega(X_H, \xi_i) = 0, i = 1, \dots, n$ . Mas, como  $T_P\Lambda$  é maximalmente isotrópico (a dimensão máxima de um subespaço onde  $\omega$  se anula é  $n$ ), isto implica que  $X_H = \sum \lambda^i \xi_i$ , o que significa que o campo Hamiltoniano  $X_H$  é tangente a  $\Lambda$ .

♣.

Consideremos agora uma variedade isotrópica  $\Lambda_0$ , de dimensão  $n - 1$ , contida na hipersuperfície de nível  $\Sigma_h$ . Por cada ponto  $P = (\mathbf{x}_0, \mathbf{p}_0) \in \Lambda_0$ , façamos passar uma curva integral  $(\mathbf{x}(t), \mathbf{p}(t))$  do campo Hamiltoniano  $X_H$ , com condição inicial  $(\mathbf{x}_0, \mathbf{p}_0) \in \Lambda_0$ . Suponhamos que a variedade  $\Lambda$ , que assim se obtém, tem dimensão  $n$  (o que acontece, se  $X_H$  fôr transversal a  $\Lambda_0$  e se  $t$  fôr suficientemente pequeno).

Figure 2.8:

Então  $\Lambda$  será uma Lagrangeana, de dimensão  $n$ , contida na hipersuperfície de nível  $\Sigma_h$ . Se  $\Lambda$  se define localmente por uma equação do tipo:

$$\mathbf{p} = \alpha(\mathbf{x}) = S_{\mathbf{x}}(\mathbf{x}) = \left( \frac{\partial S}{\partial x^i}(\mathbf{x}) \right)_{i=1, \dots, n} \quad (2.5.12)$$

isto é, se  $\Lambda$  fôr localmente da forma  $\Lambda = \Lambda_{dS}$ , então  $S$  satisfaz a **equação reduzida de Hamilton-Jacobi**:

$$H(\mathbf{x}, S_{\mathbf{x}}) = h \quad (2.5.13)$$

Neste caso, a função:

$$\mathcal{S}(t, \mathbf{x}) \stackrel{\text{def}}{=} -ht + S(\mathbf{x}) \quad (2.5.14)$$

satisfará a equação de Hamilton-Jacobi:

$$\mathcal{S}_t + H(\mathbf{x}, \mathcal{S}_{\mathbf{x}}) = 0 \quad (2.5.15)$$

Suponhamos agora que o Hamiltoniano é homogéneo positivo de grau 1 nas variáveis  $\mathbf{p}$ :

$$H(\mathbf{x}, \lambda \mathbf{p}) = \lambda H(\mathbf{x}, \mathbf{p}), \quad \forall \lambda > 0 \quad (2.5.16)$$

Como exemplos típicos temos:

- o Hamiltoniano óptico  $H(\mathbf{x}, \mathbf{p}) = \|\mathbf{p}\|/n(\mathbf{x})$ , que rege a propagação dos raios num meio isotrópico com índice de refração  $n(\mathbf{x})$ .
- As geodésicas de uma métrica Riemanniana podem ser deduzidas do Lagrangeano  $L = g_{ij}(\mathbf{x})\dot{x}^i\dot{x}^j$  ao qual corresponde o Hamiltoniano  $H = g^{ij}p_i p_j$ . No entanto, podemos também usar o Hamiltoniano  $\sqrt{H} = \sqrt{g^{ij}p_i p_j}$ . Para um nível de energia  $H = h$  constante, as equações de Hamilton são proporcionais com um factor de proporcionalidade constante. De facto, se  $H = h$  constante, então em  $\Sigma_h$ :

$$X_H = 2\sqrt{h} X_{\sqrt{H}} \quad (2.5.17)$$

- **♣ Proposição 2.8** ... O fluxo  $\Phi_t$ , de um campo Hamiltoniano  $X_H$ , associado a uma Hamiltoniano  $H$ , homogéneo positivo de grau 1 nas variáveis  $\mathbf{p}$ , preserva a 1-forma canónica de Liouville  $\boldsymbol{\theta} = \mathbf{p}d\mathbf{x}$ .

**Dem.:** É suficiente provar que a derivada de Lie  $L_{X_H}\boldsymbol{\theta} = 0$ . Como  $X_H = H_{\mathbf{p}}\partial_{\mathbf{x}} - H_{\mathbf{x}}\partial_{\mathbf{p}}$ , vem que:

$$\begin{aligned}
 L_{X_H}\boldsymbol{\theta} &= L_{X_H}(\mathbf{p}d\mathbf{x}) \\
 &= (i_{X_H}d + di_{X_H})(\mathbf{p}d\mathbf{x}) \\
 &= i_{X_H}(d\mathbf{p} \wedge d\mathbf{x}) + d(\mathbf{p}d\mathbf{x}(X_H)) \\
 &= d\mathbf{p}(X_H)d\mathbf{x} - d\mathbf{x}(X_H)d\mathbf{p} + d(\mathbf{p}H_{\mathbf{p}}) \\
 &= -H_{\mathbf{x}}d\mathbf{x} - H_{\mathbf{p}}d\mathbf{p} + H_{\mathbf{p}}d\mathbf{p} + \mathbf{p}dH_{\mathbf{p}} \\
 &= -H_{\mathbf{x}}d\mathbf{x} - \mathbf{p}dH_{\mathbf{p}} \\
 &= 0
 \end{aligned} \tag{2.5.18}$$

A última igualdade deve-se à identidade de Euler  $\mathbf{p}H_{\mathbf{p}} - H = 0$ , que é válida atendendo à homogeneidade de  $H$ . Dessa identidade obtem-se:

$$\begin{aligned}
 0 &= d(\mathbf{p}H_{\mathbf{p}}) - dH = H_{\mathbf{p}}d\mathbf{p} + \mathbf{p}d(H_{\mathbf{p}}) - dH \\
 &= H_{\mathbf{p}}d\mathbf{p} + \mathbf{p}d(H_{\mathbf{p}}) - H_{\mathbf{x}}d\mathbf{x} - H_{\mathbf{p}}d\mathbf{p} = \mathbf{p}d(H_{\mathbf{p}}) - H_{\mathbf{x}}d\mathbf{x}
 \end{aligned}$$

que é (2.5.18).

♣.

Desta proposição concluímos que o fluxo  $\Phi_t$ , de um Hamiltoniano homogéneo positivo de grau 1, nas variáveis  $\mathbf{p}$ , preserva a classe das Lagrangeanas cónicas (aquelas onde a forma  $\boldsymbol{\theta} = \mathbf{p}d\mathbf{x}$  se anula idênticamente).

Como já vimos, dado um ponto  $\mathbf{x}_0 \in \mathbb{R}_{\mathbf{x}}^n$ , a Lagrangeana:

$$\Lambda_{\mathbf{x}_0} = \{\mathbf{x}_0\} \times T_{\mathbf{x}_0}^*\mathbb{R}^n = \{(\mathbf{x}_0, \mathbf{p}) : \mathbf{p} \in \mathbb{R}^n\}$$

é cónica - é a fibra de  $T^*\mathbb{R}^n$  por cima de  $\mathbf{x}_0$ . Dado um Hamiltoniano homogéneo  $H$ , podemos construir um feixe central de trajectórias de centro  $\mathbf{x}_0$ , projectando no espaço de configuração  $\mathbb{R}_{\mathbf{x}}^n$ , as curvas integrais de  $X_H$ , com condição inicial  $(\mathbf{x}_0, \mathbf{p}_0)$ , onde  $\mathbf{p}_0$  varia numa variedade inicial dada  $V_0 \subset \Lambda_{\mathbf{x}_0}$ , de dimensão  $n - 1$ , tal que:

$$V_0 \subset \Lambda_{\mathbf{x}_0} \cap \Sigma_h \tag{2.5.19}$$

Para  $t > 0$  teremos uma certa variedade  $V_t$ , em  $T^*\mathbb{R}^n$ , também de dimensão  $n - 1$ , cuja projecção no espaço de configuração  $\mathbb{R}_{\mathbf{x}}^n$  é a chamada **frente de onda**, no instante  $t$ , correspondente ao feixe central de trajectórias de centro  $\mathbf{x}_0$ .

Consideremos agora uma Lagrangeana cónica qualquer  $\Lambda$ , em  $T^*\mathbb{R}^n$ , (portanto  $\dim \Lambda = n$ , e  $\mathbf{p}d\mathbf{x}|_{\Lambda} = 0$ ).

Figure 2.9:

A respectiva projecção, no espaço de configuração  $\mathbb{R}_x^n$ , tem dimensão  $\leq (n - 1)$ . De facto, se  $\xi = \mathbf{a}\partial_x + \mathbf{b}\partial_p$  é um vector tangente a  $\Lambda$ , num dos seus pontos  $(\mathbf{x}, \mathbf{p}) \in \Lambda$ , então  $d\pi(\xi) = \mathbf{a}\partial_x$  onde  $\mathbf{a}$ , tem que satisfazer a equação linear:

$$0 = (\mathbf{p}d\mathbf{x})(\bar{\xi}) = (\mathbf{p}d\mathbf{x})(\mathbf{a}\partial_x + \mathbf{b}\partial_p) = \mathbf{a}\mathbf{p}$$

Portanto, o espaço tangente à projecção  $\pi(\Lambda)$ , tem uma dimensão  $\leq (n - 1)$ . Em particular,  $\Lambda$  não pode ser definida como gráfico de uma 1-forma (mesmo localmente), o que impede de gerar  $\Lambda$  por uma iconal (local), contrariamente ao que eventualmente acontece com variedades não cónicas.

Por isso, de forma análoga ao caso anterior, em que  $\Lambda = \Lambda_{x_0}$ , consideramos uma variedade inicial  $V_0$ , de dimensão  $n - 1$ , contida em  $\Lambda$ , e definida por  $H \equiv h$  (constante):

$$V_0 \subset \Lambda \cap \Sigma_h$$

Dado um Hamiltoniano homogéneo  $H$ , construímos então o “tubo de trajectórias” de  $X_H$ , com “base” em  $V_0$ . Por outras palavras, varremos  $V_0$  pelo fluxo de  $X_H$ . Desta forma obtemos uma Lagrangeana  $\tilde{\Lambda}$ , contida na hipersuperfície de nível  $H \equiv h$  (constante):

$$\tilde{\Lambda} = \cup_t \Phi_t(V_0) \quad (2.5.20)$$

Sob certas condições, pode acontecer que  $\tilde{\Lambda}$  seja já gerada (localmente) por uma certa iconal  $S$ , de tal forma que  $\mathbf{p} = S_x(\mathbf{x})$ , onde  $S$  verifica a equação reduzida de Hamilton-Jacobi:

$$H(\mathbf{x}, S_x) \equiv h \quad (2.5.21)$$

Como  $\Lambda$  é cónica, a forma  $\theta = \mathbf{p}d\mathbf{x}$  anula-se em  $V_0 \subset \Lambda$ . Portanto ela também se anula em cada superfície  $V_t = \Phi_t(V_0)$ , já que, como vimos, o fluxo  $\Phi_t$  preserva a forma  $\theta$ . Mas isto implica que a fase local:

$$W(P) = \int_{P_0}^P \mathbf{p}d\mathbf{x}$$

definida em  $\tilde{\Lambda}$ , é constante em cada  $V_t = \Phi_t(V_0)$ ,  $\forall t$ . Vemos pois que as superfícies de nível  $S(\mathbf{x}) \equiv \text{constante}$ , onde  $S = W \circ \alpha$ , confundem-se exactamente com as projecções das superfícies  $V_t$  sobre o espaço de configuração  $\mathbb{R}_x^n$ , isto é, com as frentes de onda.

## 2.6 Aplicações à óptica geométrica

### 2.6.1 Construção das frentes de onda a partir dos raios

Consideremos a equação de Hamilton-Jacobi:

$$H(\mathbf{x}, d\psi(\mathbf{x})) = \frac{1}{2} \quad (2.6.1)$$

correspondente ao Hamiltoniano:

$$H(\mathbf{x}, \mathbf{p}) = \frac{\mathbf{p}^2}{n^2} \quad (2.6.2)$$

em  $T^*\mathbb{R}^3$ , munido das coordenadas canônicas  $\mathbf{x} = (x, y, z)$  e  $\mathbf{p} = (p, q, r)$ . As equações canônicas são dadas por (2.2.35):

$$\begin{cases} \dot{\mathbf{x}} = H_{\mathbf{p}} \\ \dot{\mathbf{p}} = -H_{\mathbf{x}} \end{cases} \quad (2.6.3)$$

Uma solução de (2.6.3) está contida em  $H^{-1}(1/2)$ :

$$H(\mathbf{x}(t), \mathbf{p}(t)) \equiv \frac{1}{2} \quad (2.6.4)$$

O problema que vamos agora discutir é o seguinte: supondo que sabemos resolver as equações canônicas (2.6.3), pretendemos calcular uma família de frentes de onda  $\psi = ct$ , onde  $\psi$  é uma solução da equação de Hamilton-Jacobi (2.6.1) que satisfaz a condição inicial:

$$\psi(\phi(\mathbf{u})) = F(\mathbf{u}), \quad \mathbf{u} = (\xi, \eta) \in U \subseteq \mathbb{R}^2 \quad (2.6.5)$$

Aqui  $\mathbf{x} = \phi(\mathbf{u})$  é a representação paramétrica de uma superfície  $\Gamma$  em  $\mathbb{R}_x^3$  (eventualmente degenerada), e  $F = F(\mathbf{u})$  é uma função dada.  $\Gamma$  pode representar, por exemplo, um obstáculo, a fronteira entre dois meios ópticos ou pode reduzir-se até a um ponto. A família de frentes de onda que pretendemos construir, e que satisfazem as condições iniciais (2.6.5), dadas em  $\Gamma$ , pode consistir de ondas incidentes, reflectidas ou refractadas.

Suponhamos que  $\psi = \psi(\mathbf{x})$  é uma solução da equação de Hamilton-Jacobi (2.6.1). Então as superfícies  $\psi = ct$  determinam uma família a um parâmetro de frentes de onda e uma família a dois parâmetros de raios ortogonais a essas frentes de onda. Seja  $\mathbf{x} = \mathbf{x}(\tau)$  um desses raios e  $P_0 = \mathbf{x}(\tau_0), P_1 = \mathbf{x}(\tau_1)$  dois pontos nesse raio. O comprimento óptico entre esses dois pontos:

$$\begin{aligned} V(P_0, P_1) &= \int \mathbf{p} d\mathbf{x} \\ &= \int_{\tau_0}^{\tau_1} \psi_x \dot{x} + \psi_y \dot{y} + \psi_z \dot{z} d\tau \\ &= \int_{\tau_0}^{\tau_1} (\nabla \psi)^2 d\tau \\ &= \int_{\tau_0}^{\tau_1} n^2 d\tau \end{aligned} \quad (2.6.6)$$

é dado pela diferença:

$$\psi(P_1) - \psi(P_0) = \int_{\tau_0}^{\tau_1} n^2 d\tau \quad (2.6.7)$$

É pois natural considerar o seguinte método para calcular a solução do problema anterior - por cada ponto  $\phi(\mathbf{u})$  da superfície inicial  $\Gamma$ , fazemos passar o único raio que por aí passa no instante  $\tau = 0$ , isto é, a projecção  $\mathbf{x}(\mathbf{u}; \tau)$  da solução das equações canónicas (2.6.3) com condição inicial:

$$\mathbf{x}(\mathbf{u}; \tau = 0) = \phi(\mathbf{u}), \quad \mathbf{p}(\mathbf{u}; \tau = 0) = \mathbf{p}_0(\mathbf{u}) \quad (2.6.8)$$

Aqui  $\phi(\mathbf{u})$  é dado, enquanto que  $\mathbf{p}_0(\mathbf{u})$  deve ser calculado de modo assegurar as condições iniciais. Como pretendemos que:

$$\psi(\phi(\mathbf{u})) = F(\mathbf{u})$$

devemos ter, pela regra da cadeia, que:

$$d\psi_{\phi(\mathbf{u})} \circ d\phi_{\mathbf{u}} = dF_{\mathbf{u}}, \quad \forall \mathbf{u} \quad (2.6.9)$$

Pondo:

$$\mathbf{p}_0(\mathbf{u}) = d\psi_{\phi(\mathbf{u})}$$

a equação (2.6.9) implica que  $\langle \phi_{\xi}(\mathbf{u}), \mathbf{p}_0(\mathbf{u}) \rangle = F_{\xi}(\mathbf{u})$ ,  $\langle \phi_{\eta}(\mathbf{u}), \mathbf{p}_0(\mathbf{u}) \rangle = F_{\eta}(\mathbf{u})$ . A estas equações devemos acrescentar ainda a equação  $H(\phi(\mathbf{u}), \mathbf{p}_0(\mathbf{u})) = 1/2$ . Portanto, para cada  $\mathbf{u} = (\xi, \eta)$ , as três equações:

$$\begin{cases} \langle \phi_{\xi}(\mathbf{u}), \mathbf{p}_0(\mathbf{u}) \rangle & = F_{\xi}(\mathbf{u}) \\ \langle \phi_{\eta}(\mathbf{u}), \mathbf{p}_0(\mathbf{u}) \rangle & = F_{\eta}(\mathbf{u}) \\ H(\phi(\mathbf{u}), \mathbf{p}_0(\mathbf{u})) & = 1/2 \end{cases}$$

que, em princípio, determinam  $\mathbf{p}_0 = \mathbf{p}(\mathbf{u}; \tau = 0)$ .

Mais formalmente,  $\Phi : \mathbf{u} \mapsto (\phi(\mathbf{u}), \mathbf{p}_0(\mathbf{u}))$ , onde  $\mathbf{p}_0(\mathbf{u})$  é determinado pelas equações (2.6.10), define uma subvariedade isotrópica de dimensão 2 em  $T^*\mathbb{R}^3$ , contida em  $\{H = 1/2\}$ . De facto as duas primeiras equações são equivalentes a:

$$\Phi^*(\mathbf{p}d\mathbf{x}) = dF \quad \Rightarrow \quad \Phi(d\mathbf{p} \wedge d\mathbf{x}) = 0 \quad (2.6.10)$$

Movendo esta subvariedade sob a acção do fluxo do campo Hamiltoniano  $X_H$  (que é tangente a  $\{H = 1/2\}$ ), supondo que  $X_H$  é transversal, obtemos uma subvariedade Lagrangeana (dimensão 3) contida em  $\{H = 1/2\}$ , que localmente é da forma gráfico de  $d\psi$ , e portanto este  $\psi$  é a solução pretendida da equação de Hamilton-Jacobi (2.6.1).

Pômos portanto:

$$\tilde{\psi}(\xi, \eta, \tau) = F(\xi, \eta) + \int_{r_{\tau}} \mathbf{p} d\mathbf{x} \quad (2.6.11)$$

onde  $r_{\tau}$  é a única solução das equações canónicas (2.6.3) que une  $(\phi(\mathbf{u}), \mathbf{p}_0)$ , com coordenadas  $(\xi, \eta, 0)$ , ao ponto  $(\xi, \eta, \tau)$ . Por (2.6.6), sabemos que o integral  $\int_{r_{\tau}} \mathbf{p} d\mathbf{x}$  dá exactamente o comprimento óptico  $\int n^2 d\tau$  desse raio.

É claro que estamos a supôr que, numa vizinhança da superfície  $\Gamma$ , em  $\mathbb{R}_{\mathbf{x}}^3$ , cuja representação paramétrica é  $\mathbf{x} = \boldsymbol{\phi}(\mathbf{u}) = \boldsymbol{\phi}(\xi, \eta)$ , podemos escolher  $(\xi, \eta, \tau)$  como coordenadas locais, o que acontece se o determinante Jacobiano  $\frac{\partial(x,y,z)}{\partial(\xi,\eta,\tau)} \neq 0$ . Sendo assim, pômos finalmente:

$$\psi(x, y, z) = \tilde{\psi}(\xi(x, y, z), \eta(x, y, z), \tau(x, y, z)) \quad (2.6.12)$$

onde  $\tilde{\psi}(\xi, \eta, \tau)$  é definida por (2.6.11).

Resta provar que (2.6.12) é a solução procurada. É óbvio que  $\psi(\xi, \eta, 0) = F(\xi, \eta)$ . Por outro lado, como  $H(\boldsymbol{\phi}(\mathbf{u}), \mathbf{p}_0(\mathbf{u})) = 1/2$  e  $H$  é um integral primeiro das equações canónicas, deduzimos que:

$$H(\mathbf{x}, \mathbf{p}) = 1/2 \quad (2.6.13)$$

Resta então provar que:

$$\mathbf{p} = d\psi(\mathbf{x}) \quad (2.6.14)$$

Derivando (2.6.11) em ordem a  $\tau$  vem que:

$$\tilde{\psi}_\tau = \mathbf{p} \mathbf{x}_\tau \quad (2.6.15)$$

Derivando agora (2.6.11) em ordem a  $\xi$  vem sucessivamente que:

$$\begin{aligned} \tilde{\psi}_\xi &= F_\xi + \int_0^\tau \mathbf{p}_\xi \mathbf{x}_\tau d\tau + \int_0^\tau \mathbf{p} \mathbf{x}_{\xi\tau} d\tau \\ &= F_\xi + \mathbf{p} \mathbf{x}_\xi - (\mathbf{p} \mathbf{x}_\xi|_{\tau=0}) + \int_0^\tau [\mathbf{p}_\xi \mathbf{x}_\tau - \mathbf{p}_\tau \mathbf{x}_\xi] d\tau \\ &= \mathbf{p} \mathbf{x}_\xi + \int_0^\tau [\mathbf{p}_\xi H_{\mathbf{p}} + H_{\mathbf{x}} \mathbf{x}_\xi] d\tau \\ &= \mathbf{p} \mathbf{x}_\xi + \int_0^\tau \frac{\partial H}{\partial \xi} d\tau \\ &= \mathbf{p} \mathbf{x}_\xi \end{aligned} \quad (2.6.16)$$

atendendo às equações canónicas, a que  $F_\xi = \mathbf{p} \mathbf{x}_\xi|_{\tau=0}$  e ainda a  $H \equiv 1/2$ . Anàlogamente se obtem que  $\tilde{\psi}_\eta = \mathbf{p} \mathbf{x}_\eta$ , e reunindo as três equações obtemos:

$$\begin{aligned} \tilde{\psi}_\tau &= \mathbf{p} \mathbf{x}_\tau \\ \tilde{\psi}_\xi &= \mathbf{p} \mathbf{x}_\xi \\ \tilde{\psi}_\eta &= \mathbf{p} \mathbf{x}_\eta \end{aligned} \quad (2.6.17)$$

Por outro lado, como  $\tilde{\psi}(\xi, \eta, \tau) = \psi(\mathbf{x}(\xi, \eta, \tau))$ , vem que:

$$\begin{aligned} \tilde{\psi}_\tau &= \psi_{\mathbf{x}} \mathbf{x}_\tau \\ \tilde{\psi}_\xi &= \psi_{\mathbf{x}} \mathbf{x}_\xi \\ \tilde{\psi}_\eta &= \psi_{\mathbf{x}} \mathbf{x}_\eta \end{aligned} \quad (2.6.18)$$

e finalmente, (2.6.17) e (2.6.18) implicam que:

$$\mathbf{p} = d\psi(\mathbf{x}) \quad (2.6.19)$$

como se pretendia.

Vamos aplicar esta teoria em duas situações concretas:

### 2.6.2 Propagação das frentes de onda através de uma descontinuidade do meio. Lei de Snell

Suponhamos que a superfície  $\Gamma$ , parametrizada como antes por  $\mathbf{x} = \phi(\mathbf{u})$ , separa dois meios notados por 1 e 2, respectivamente. O meio 1 supõe-se homogéneo e isotrópico com índice de refração  $n_1$ . Seja  $\psi_1(\mathbf{x}) = ct$  um sistema de frentes de onda que se propaga no meio 1, e suponhamos que:

$$\psi_1(\phi(\mathbf{u})) = F(\mathbf{u}) \quad (2.6.20)$$

Cada uma das frentes  $\psi_1 = ct$  dá origem a uma frente da família  $\psi_2 = ct$  que se propaga no meio 2. É natural supôr que, em cada ponto  $\phi(\mathbf{u}) \in \Gamma$ , a frente do meio 2 tem o mesmo valor da frente do meio 1, que intersecta  $\Gamma$  em  $\phi(\mathbf{u})$ , nomeadamente o valor  $F(\mathbf{u})$ . Queremos pois determinar a família  $\psi_2 = ct$ , que se propaga no meio 2, e que satisfaz:

$$\psi_2(\phi(\mathbf{u})) = F(\mathbf{u}) \quad (2.6.21)$$

De (2.6.20) deduzimos, pela regra da cadeia, que:

$$(d\psi_1)_{\phi(\mathbf{u})}(d\phi_{\mathbf{u}}(V)) = dF_{\mathbf{u}}(V), \quad \forall V \in T_{\mathbf{u}}\mathbb{R}^2$$

ou, pondo  $(d\psi_1)_{\phi(\mathbf{u})} = \mathbf{p}_1$ :

$$\begin{cases} \langle \phi_{\xi}(\mathbf{u}), \mathbf{p}_1 \rangle = F_{\xi}(\mathbf{u}) \\ \langle \phi_{\eta}(\mathbf{u}), \mathbf{p}_1 \rangle = F_{\eta}(\mathbf{u}) \end{cases}$$

o que permite calcular as derivadas de  $F$ . Para calcular agora  $\mathbf{p}_2$ , que permitirá o cálculo de  $\psi_2$ , de acordo com a secção anterior, temos que resolver as equações (ver (2.6.10)):

$$\begin{cases} \langle \phi_{\xi}(\mathbf{u}), \mathbf{p}_2 \rangle = F_{\xi}(\mathbf{u}) \\ \langle \phi_{\eta}(\mathbf{u}), \mathbf{p}_2 \rangle = F_{\eta}(\mathbf{u}) \\ H_2(\phi(\mathbf{u}), \mathbf{p}_2) = 1/2 \end{cases}$$

em ordem a  $\mathbf{p}_2$ . Subtraindo a primeira e segunda equações (2.6.22) das primeira e segunda equações (2.6.22), respectivamente, obtemos:

$$\begin{cases} \langle \phi_{\xi}(\mathbf{u}), \mathbf{p}_1 - \mathbf{p}_2 \rangle = 0 \\ \langle \phi_{\eta}(\mathbf{u}), \mathbf{p}_1 - \mathbf{p}_2 \rangle = 0 \\ H_2(\phi(\mathbf{u}), \mathbf{p}_2) = 1/2 \end{cases}$$

Se este sistema não tiver solução, não haverá frentes de onda penetrando no meio 2. Se tiver solução haverá frentes penetrando no meio 2. Seja:

$$\mathbf{N}(\mathbf{u}) = \phi_{\xi}(\mathbf{u}) \times \phi_{\eta}(\mathbf{u})$$

o vector normal a  $\Gamma$ , no ponto  $\phi(\mathbf{u})$ . As duas primeiras equações em (2.6.22) dizem que  $\mathbf{p}_1 - \mathbf{p}_2$  é colinear com  $\mathbf{N}$ , digamos  $\mathbf{p}_1 - \mathbf{p}_2 = \lambda \mathbf{N}$ , e portanto:

$$\mathbf{p}_1 \times \mathbf{N} = \mathbf{p}_2 \times \mathbf{N} \quad (2.6.22)$$

Esta fórmula diz que a normal à onda incidente, a normal à onda refractada e a normal à superfície  $\Gamma$  são coplanares. Por outro lado, da definição do produto vectorial, deduzimos a **lei de refracção de Snell**:

$$\|\mathbf{p}_2\| \sin \phi_2 = \|\mathbf{p}_1\| \sin \phi_1 \quad (2.6.23)$$

onde  $\phi_1$  e  $\phi_2$  são, respectivamente, os ângulos que  $\mathbf{p}_1$  e  $\mathbf{p}_2$  fazem com  $\mathbf{N}$ . Como  $\|\mathbf{p}_1\| = n_1$  e  $\|\mathbf{p}_2\| = n_2$ , a lei de Snell pode ser escrita na forma:

$$n_2 \sin \phi_2 = n_1 \sin \phi_1 \quad (2.6.24)$$

Figure 2.10: Lei de refracção de Snell.

Além da família  $\psi_1 = ct$ , que se propaga no meio 1, existe uma outra família  $\psi_3 = ct$ , que se propaga também no meio 1, e que satisfaz as condições de fronteira indicadas. Para determinar esta família, introduzimos um (co)vector  $\mathbf{p}_3$ , e as equações:

$$\begin{cases} \langle \phi_\xi(\mathbf{u}), \mathbf{p}_3 - \mathbf{p}_2 \rangle = 0 \\ \langle \phi_\eta(\mathbf{u}), \mathbf{p}_3 - \mathbf{p}_2 \rangle = 0 \\ H_1(\phi(\mathbf{u}), \mathbf{p}_3) = 1/2 \end{cases}$$

Usando exactamente os mesmos argumentos, deduzimos que  $\mathbf{p}_3 - \mathbf{p}_1 = LN$ , e portanto:

$$\mathbf{p}_3 \times \mathbf{N} = \mathbf{p}_1 \times \mathbf{N} \quad (2.6.25)$$

e finalmente:

$$n_1 \sin \phi_3 = n_1 \sin \phi_1 \quad (2.6.26)$$

Da figura 2.10, deduzimos que  $\phi_3 = \pi - \phi_1$  e, como antes,  $\mathbf{p}_1$ ,  $\mathbf{p}_3$  e  $\mathbf{N}$  são coplanares. A equação (2.6.26) é a lei da reflexão para meios isotrópicos.

### 2.6.3 Princípio de Huygens

Vamos agora aplicar a teoria anterior à situação em que  $\Gamma$  é ela própria uma frente de onda, digamos  $\psi(\mathbf{x}) = 0$ . A função  $F(\mathbf{u})$  é portanto a função nula, e as equações (2.6.10) são agora:

$$\begin{cases} \langle \phi_\xi(\mathbf{u}), \mathbf{p}_0 \rangle = 0 \\ \langle \phi_\eta(\mathbf{u}), \mathbf{p}_0 \rangle = 0 \\ H(\phi(\mathbf{u}), \mathbf{p}_0) = 1/2 \end{cases}$$

já que  $F_\xi = 0 = F_\eta$ , enquanto que a equação (2.6.12) para as frentes de onda é:

$$\psi(\xi, \eta, \tau) = \int_{r_\tau} \mathbf{p} \, d\mathbf{x} \quad (2.6.27)$$

onde  $r_\tau$  é a única solução das equações canónicas (2.6.3) que une  $(\phi(\mathbf{u}), \mathbf{p}_0)$ , com coordenadas  $(\xi, \eta, 0)$ , ao ponto  $(\xi, \eta, \tau)$ . As duas primeiras equações em (2.6.27) determinam uma recta de (co)vectores ortogonais ao espaço tangente a  $\Gamma = \{\psi = 0\}$ , no ponto  $\phi(\mathbf{u})$ . A terceira equação determina então a intersecção dessa recta com a hipersuperfície  $H = 1/2$ . Por exemplo, num meio isotrópico, com  $H = \mathbf{p}^2/n^2$ , onde  $\mathbf{p}^2 = \mathbf{p} \cdot \mathbf{p}$ , temos duas soluções para (2.6.27) que determinam duas famílias de frentes - uma propaga-se para um lado de  $\Gamma$  e a outra para o outro lado.

Suponhamos agora que  $\Gamma$  degenera num único ponto  $\Gamma = \mathbf{x}_0$ , de tal forma que a função  $\phi$  é agora constante (e igual a  $\mathbf{x}_0$ ). Neste caso, as equações (2.6.27) reduzem-se a uma única:

$$H(\mathbf{x}_0, \mathbf{p}_0) = 1/2 \quad (2.6.28)$$

Com  $\mathbf{x}_0$  fixo, (2.6.28) representa uma superfície no espaço cotangente  $T_{\mathbf{x}_0}^* \mathbb{R}^3$ , e pode localmente ser parametrizada por dois dos parâmetros  $(p_0, q_0, r_0) = \mathbf{p}_0$ .

Seleccionemos agora uma família a dois parâmetros de soluções (**bicaracterísticas**) das equações canónicas (2.6.3), digamos:

$$\begin{cases} \mathbf{x} = \mathbf{x}(\mathbf{p}_0; \tau) \\ \mathbf{p} = \mathbf{p}(\mathbf{p}_0; \tau) \end{cases} \quad (2.6.29)$$

determinadas pelas condições iniciais:

$$\begin{cases} \mathbf{x}(\mathbf{p}_0; \tau = 0) = \mathbf{x}_0 \\ \mathbf{p}(\mathbf{p}_0; \tau = 0) = \mathbf{p}_0 \end{cases} \quad (2.6.30)$$

onde, para cada  $\mathbf{x}_0, \mathbf{p}_0$  é solução de (2.6.28). A frente de onda é então determinada por:

$$\psi(\mathbf{p}_0; \tau) = \int_{r_\tau} \mathbf{p} \, d\mathbf{x} \quad (2.6.31)$$

onde  $r_\tau$  é a bicaracterística acima referida. Estas soluções especiais da equação de Hamilton-Jacobi, dizem-se **ondas esféricas** ou **onduletas** (wavelets), e a função:

$$V(P_0, P) = V(\mathbf{x}_0, \mathbf{x}) \stackrel{\text{def}}{=} \int_{r_\tau} \mathbf{p} \, d\mathbf{x} \quad (2.6.32)$$

é a chamada **função característica pontual de Hamilton**. Recorde que o integral  $\int_{r_\tau} \mathbf{p} \, d\mathbf{x}$  é igual ao comprimento óptico  $\int_{r_\tau} n^2 \, d\tau$ , do raio  $r_\tau$  e, portanto,  $V(P_0, P)$  é igual à distância óptica entre os pontos  $P_0$  e  $P_1$ .

Com a ajuda das onduletas  $V(P_0, P)$  podemos dar um outro método, que se deve a Huygens, para a construção das frentes de onda  $\psi = ct$  tais que  $\psi = 0$  é a superfície inicial  $\Gamma$ .

Suponhamos que a superfície inicial  $\Gamma$  é dada paramètricamente por:

$$\mathbf{x}_0 = \phi(\mathbf{u}), \quad \mathbf{u} = (\xi, \eta) \quad (2.6.33)$$

Cada ponto  $\mathbf{x}_0 \in \Gamma$  determina a família de onduletas:

$$V(\mathbf{u}; \mathbf{x}) = V(\xi, \eta; \mathbf{x}) = ct \quad (2.6.34)$$

e, quando  $\mathbf{u} = (\xi, \eta)$  varia e para um certo  $t$  fixo, obtemos uma família a dois parâmetros de frentes de onda. Vamos mostrar que a envolvente desta família é uma frente de onda  $\psi = ct$  da família determinada por  $\Gamma$ , tal que  $\psi = 0$  é exactamente  $\Gamma$ .

De facto, essa envolvente é obtida eliminando  $\xi$  e  $\eta$  nas equações seguintes:

$$\begin{aligned} V(\xi, \eta; \mathbf{x}) &= ct \\ V_\xi(\xi, \eta; \mathbf{x}) &= 0 \\ V_\eta(\xi, \eta; \mathbf{x}) &= 0 \end{aligned} \quad (2.6.35)$$

Em geral, consegue-se esta eliminação resolvendo as duas últimas equações em ordem a  $\xi$  e  $\eta$ , como funções de  $\mathbf{x}$ , e substituindo na primeira, para obter uma função:

$$\psi(\mathbf{x}) = V(\xi(\mathbf{x}), \eta(\mathbf{x}); \mathbf{x}) = ct \quad (2.6.36)$$

Pondo  $\mathbf{x} = (x, y, z)$ , obtemos então as seguintes derivadas:

$$\begin{aligned} \psi_x &= V_\xi \xi_x + V_\eta \eta_x + V_x = V_x \\ \psi_y &= V_\xi \xi_y + V_\eta \eta_y + V_y = V_y \\ \psi_z &= V_\xi \xi_z + V_\eta \eta_z + V_z = V_z \end{aligned} \quad (2.6.37)$$

já que  $V_\xi = 0 = V_\eta$ . Mas isto significa que  $\nabla\psi = \nabla V$ , isto é, a frente de onda que passa em  $\mathbf{x}$  é tangente à onduleta que passa nesse mesmo ponto  $\mathbf{x}$ .

Como  $V(\xi, \eta; \mathbf{x})$ , como função de  $\mathbf{x}$ , satisfaz a equação de Hamilton-Jacobi  $H(\mathbf{x}, dV(\mathbf{x})) = 1/2$ , também  $\psi$  a satisfaz, já que  $d\psi(\mathbf{x}) = dV(\mathbf{x})$ . Portanto  $\psi$  é uma frente de onda. Resta mostrar que, para  $t = 0$ ,  $\psi = 0$  é exactamente  $\Gamma$ . Das equações (2.6.35) podemos determinar  $\mathbf{x}$  como função de  $\xi, \eta, t$ :

$$\mathbf{x} = \mathbf{x}(\xi, \eta, t)$$

Para cada  $t$  fixo, estas equações dão uma representação paramétrica da superfície  $\psi(\mathbf{x}) = ct$ . Em particular, para  $t = 0$  obtemos uma representação paramétrica da superfície  $\psi(\mathbf{x}) = 0$ .

## 2.7 Apêndice. Equações de Maxwell e óptica geométrica

### 2.7.1 Propagação da luz num meio isotrópico não homogêneo

O **campo electromagnético** é caracterizado por dois campos de vectores em  $\mathbb{R}^3$ , dependentes do tempo:

$$\begin{aligned} (t, \mathbf{x}) &\longmapsto \mathbf{E}(t, \mathbf{x}) = (E_1(t, \mathbf{x}), E_2(t, \mathbf{x}), E_3(t, \mathbf{x})) \in \mathbb{R}^3 && \text{campo eléctrico} \\ (t, \mathbf{x}) &\longmapsto \mathbf{H}(t, \mathbf{x}) = (H_1(t, \mathbf{x}), H_2(t, \mathbf{x}), H_3(t, \mathbf{x})) \in \mathbb{R}^3 && \text{campo magnético} \end{aligned}$$

enquanto que as propriedades do meio são caracterizadas por duas funções escalares (que não dependem do tempo):

$$\begin{aligned} (t, \mathbf{x}) &\longmapsto \epsilon(\mathbf{x}) \in \mathbb{R} && \text{constante dielétrica} \\ (t, \mathbf{x}) &\longmapsto \mu(\mathbf{x}) \in \mathbb{R} && \text{permeabilidade magnética} \end{aligned}$$

Os campos  $\mathbf{E}$  e  $\mathbf{H}$  satisfazem um sistema 6 PDE's (lineares de primeira ordem):

$$\begin{cases} \text{[M1].} & \frac{\epsilon}{c} \mathbf{E}_t = \nabla \times \mathbf{H} \\ \text{[M2].} & \frac{\mu}{c} \mathbf{H}_t = -\nabla \times \mathbf{E} \end{cases}$$

às quais é habitual adicionar mais duas equações:

$$\begin{cases} \text{[M3].} & \nabla \cdot (\epsilon \mathbf{E}) = 0 \\ \text{[M4].} & \nabla \cdot (\mu \mathbf{H}) = 0 \end{cases}$$

onde  $\nabla = (\partial_x, \partial_y, \partial_z)$ , que dizem que não existem fontes (cargas ou correntes) de electricidade ou magnetismo. Recorde que  $\nabla \cdot = \mathbf{div}$  e  $\nabla \times = \mathbf{rot}$ .

Tomando os produtos internos de [M1] com  $\mathbf{E}$  e de [M2] com  $\mathbf{H}$ , respectivamente, e somando os resultados, obtemos:

$$\frac{1}{c} (\epsilon \mathbf{E} \cdot \mathbf{E}_t + \mu \mathbf{H} \cdot \mathbf{H}_t) - \mathbf{E} \cdot (\nabla \times \mathbf{H}) + \mathbf{H} \cdot (\nabla \times \mathbf{E}) = 0$$

Como:

$$\mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{H}) = \nabla \cdot (\mathbf{E} \times \mathbf{H})$$

vem que:

$$c \nabla \cdot (\mathbf{E} \times \mathbf{H}) + \frac{1}{2} \frac{\partial}{\partial t} (\epsilon \mathbf{E}^2 + \mu \mathbf{H}^2) = 0$$

onde pusemos  $\mathbf{E}^2 = \mathbf{E} \cdot \mathbf{E}$  e  $\mathbf{H}^2 = \mathbf{H} \cdot \mathbf{H}$ , ou ainda:

$$\frac{\partial}{\partial t} \underbrace{\frac{1}{8\pi} (\epsilon \mathbf{E}^2 + \mu \mathbf{H}^2)}_{\mathcal{E}(t, \mathbf{x})} + \mathbf{div} \underbrace{\frac{c}{4\pi} (\mathbf{E} \times \mathbf{H})}_{\mathbf{P}(t, \mathbf{x})} = 0 \quad (2.7.1)$$

onde definimos:

$$\begin{aligned} \mathcal{E}(t, \mathbf{x}) &= \frac{1}{8\pi} (\epsilon \mathbf{E}^2 + \mu \mathbf{H}^2), && \text{energia do campo} \\ \mathbf{P}(t, \mathbf{x}) &= \frac{c}{4\pi} (\mathbf{E} \times \mathbf{H}), && \text{vector de radiação} \end{aligned}$$

Temos então a seguinte **equação de continuidade** (ou conservação):

$$\frac{\partial \mathcal{E}}{\partial t} + \mathbf{div} \mathbf{P} = 0 \quad (2.7.2)$$

Se  $D \subset \mathbb{R}^3$  é um domínio fechado cujo bordo  $\partial D$  é uma superfície fechada, temos que:

$$\begin{aligned} 0 &= \int_D \left( \frac{\partial \mathcal{E}}{\partial t} + \mathbf{div} \mathbf{P} \right) dV \\ &= \frac{\partial}{\partial t} \int_D \mathcal{E} dV + \int_D \mathbf{div} \mathbf{P} dV \\ &= \frac{\partial}{\partial t} \int_D \mathcal{E} dV + \int_{\partial D} \mathbf{P} \cdot d\mathbf{S} \end{aligned} \quad (2.7.3)$$

O primeiro integral representa a variação temporal da energia total do campo em  $D$  e, portanto, o segundo integral representa o fluxo de energia através da superfície  $\partial D$ . Portanto o campo  $\mathbf{P} = \frac{c}{4\pi}(\mathbf{E} \times \mathbf{H})$  é interpretado como o fluxo de energia.

### 2.7.2 Representação integral das equações de Maxwell

Consideremos uma subvariedade compacta orientável de dimensão 4,  $V \subset \mathbb{R}_{t,\mathbf{x}}^4$ , cujo bordo  $\Gamma = \partial V$  é uma hipersuperfície fechada orientada pela normal exterior  $N$ . Se  $\partial V$  for dada (localmente) por uma equação do tipo  $\varphi(t, \mathbf{x}) = 0$ , o campo de vectores  $N$  é dado por:

$$N(t, \mathbf{x}) = \pm \frac{\mathbf{grad} \varphi}{\|\mathbf{grad} \varphi\|} = \pm \frac{(\varphi_t, \varphi_{\mathbf{x}})}{\sqrt{\varphi_t^2 + \varphi_{\mathbf{x}}^2}}$$

O resultado mais importante de que faremos uso sistemático é a fórmula de Gauss, que nos permite transformar equações que envolvem derivadas de um campo de vectores  $F \in \mathfrak{X}(\mathbb{R}^4)$  em equações que envolvem apenas o campo.

- **♣ Proposição 2.9 (Teorema de Gauss)** ... Seja  $F$  um campo de vectores de classe  $C^1$  definido num aberto que contem uma subvariedade compacta orientável  $n$ -dimensional  $V \subset \mathbb{R}^n$ , com bordo  $\partial V$ . Então:

$$\begin{aligned} \int_V \mathbf{div} F dv &= \int_{\partial V} F \cdot N d\sigma \\ &= \sum_{i=1}^n \int_{\partial V} F_i N_i d\sigma \\ &= \sum_{i=1}^n \int_{\partial V} (-1)^{i-1} F_i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n \end{aligned} \quad (2.7.4)$$

onde  $F = (F_i)$ ,  $N = (N_i)$  e  $\mathbf{div} F = \sum_i \frac{\partial F_i}{\partial x^i}$ .

♣.

Na nossa situação  $F = (0, \epsilon \mathbf{E})$  ou  $F = (0, \mu \mathbf{H})$ , isto é, em ambos os casos a  $t$ -componente é nula. Portanto, no segundo membro dos integrais anteriores, apenas figura a projecção  $\mathbf{N}$  da normal  $N$  no  $\mathbf{x}$ -espaço  $\mathbb{R}^3$ . Obtemos então:

$$\int_V \mathbf{div} (\epsilon \mathbf{E}) dv = \int_{\partial V} \epsilon (\mathbf{E} \cdot \mathbf{N}) d\sigma \quad (2.7.5)$$

Figure 2.11: .

e:

$$\int_V \mathbf{div}(\mu \mathbf{H}) dv = \int_{\partial V} \mu(\mathbf{H} \cdot \mathbf{N}) d\sigma \quad (2.7.6)$$

Mas, como por [M3] e [M4],  $\mathbf{div}(\epsilon \mathbf{E}) = 0 = \mathbf{div}(\mu \mathbf{H})$ , concluímos que:

- **♣ Proposição 2.10** ... Qualquer que seja a hipersuperfície fechada orientada,  $\Gamma$  em  $\mathbb{R}^4$ , tem-se que:

$$\int_{\Gamma} \epsilon(\mathbf{E} \cdot \mathbf{N}) d\sigma = 0 \quad (2.7.7)$$

$$\int_{\Gamma} \mu(\mathbf{H} \cdot \mathbf{N}) d\sigma = 0 \quad (2.7.8)$$

onde  $\mathbf{N}$  é a projecção, em  $\mathbb{R}_x^3$ , da normal exterior  $N$  a  $\Gamma$ .

♣.

Notemos que as condições (2.7.7) e (2.7.8) foram deduzidas das equações de Maxwell [M3] e [M4], sob as hipóteses de que  $\epsilon, \mu, \mathbf{E}$  e  $\mathbf{H}$  são de classe  $C^1$ . Neste caso, atendendo a (2.7.5) e (2.7.6), essas condições são de facto equivalentes às equações  $\mathbf{div}(\epsilon \mathbf{E}) = 0 = \mathbf{div}(\mu \mathbf{H})$ . No entanto, estas condições (2.7.7) e (2.7.8) podem ser aplicadas directamente, mesmo quando qualquer das funções envolvidas são descontínuas (desde que sejam integráveis, é claro!). Aliás o nosso objectivo é ver quais as condições que devem ser verificadas por essas possíveis descontinuidades e que podem ser deduzidas de (2.7.7) e (2.7.8).

Se no teorema de Gauss (para  $n = 4$ ),  $F$  tiver apenas uma componente não nula, digamos  $F_i$ , para um certo  $i \in \{0, 1, 2, 3\}$  fixo, vem que  $\int_V \frac{\partial F_i}{\partial x^i} dv = \int_{\partial V} F_i N_i d\sigma$ , donde se deduzem, fazendo os cálculos componente a componente, as seguintes fórmulas:

$$\int_V (\nabla \times \mathbf{E}) dv = \int_{\partial V} (\mathbf{N} \times \mathbf{E}) d\sigma \quad (2.7.9)$$

e:

$$\int_V (\nabla \times \mathbf{H}) dv = \int_{\partial V} (\mathbf{N} \times \mathbf{H}) d\sigma \quad (2.7.10)$$

Aplicamos estas fórmulas às equações de Maxwell [M1]:  $\nabla \times \mathbf{H} - \frac{\epsilon}{c} \mathbf{E}_t = \mathbf{0}$  e [M2]:  $\nabla \times \mathbf{E} + \frac{\mu}{c} \mathbf{H}_t = \mathbf{0}$ . Recordando que  $\epsilon$  e  $\mu$  não dependem de  $t$ , de tal forma que  $\epsilon \mathbf{E}_t = (\epsilon \mathbf{E})_t$  e  $\mu \mathbf{H}_t = (\mu \mathbf{H})_t$ , vem que:

$$\mathbf{0} = \int_V \left( \nabla \times \mathbf{H} - \frac{\epsilon}{c} \mathbf{E}_t \right) dv = \int_{\partial V} \left( \mathbf{N} \times \mathbf{H} - \frac{\epsilon}{c} N_0 \mathbf{E} \right) d\sigma \quad (2.7.11)$$

e:

$$\mathbf{0} = \int_V \left( \nabla \times \mathbf{E} + \frac{\mu}{c} \mathbf{H}_t \right) dv = \int_{\partial V} \left( \mathbf{N} \times \mathbf{E} + \frac{\mu}{c} N_0 \mathbf{H} \right) d\sigma \quad (2.7.12)$$

onde pusemos  $N = (N_0, \mathbf{N})$ . Como  $V$  é arbitrário, concluímos que:

- ♣ **Proposição 2.11** ... Qualquer que seja a hipersuperfície fechada orientada,  $\Gamma$  em  $\mathbb{R}^4$ , tem-se que:

$$\int_{\Gamma} \left( \mathbf{N} \times \mathbf{H} - \frac{\epsilon}{c} N_0 \mathbf{E} \right) d\sigma = \mathbf{0} \quad (2.7.13)$$

$$\int_{\Gamma} \left( \mathbf{N} \times \mathbf{E} + \frac{\mu}{c} N_0 \mathbf{H} \right) d\sigma = \mathbf{0} \quad (2.7.14)$$

onde  $N = (N_0, \mathbf{N})$  é a normal exterior a  $\Gamma$ .

♣.

Novamente notamos que estas condições envolvem apenas os campos  $\mathbf{E}$ ,  $\mathbf{H}$  e as funções  $\epsilon$ ,  $\mu$ , e não as suas derivadas. São equivalentes às equações de Maxwell [M1] e [M2], se essas derivadas existirem e forem contínuas, mas são mais gerais, no sentido em que devem ser verificadas também por campos ou funções descontínuos.

### 2.7.3 Propagação das discontinuidades. Frentes de onda

Os campos da óptica geométrica são, por definição, os valores dos campos electromagnéticos  $\mathbf{E}$  e  $\mathbf{H}$ , na fronteira que separa a região onde esses campos são nulos da região onde são não nulos. No espaço  $\mathbb{R}_{t,\mathbf{x}}^4$  esta fronteira é uma certa hipersuperfície  $\varphi(t, \mathbf{x}) = 0$ . Quando seccionamos essa hipersuperfície por hiperplanos  $t \equiv \text{constante}$ , e projectamos as secções no espaço  $\mathbb{R}_{\mathbf{x}}^3$ , paralelamente ao eixo dos  $t$ 's, obtemos uma família de superfícies  $\psi(\mathbf{x}) = ct$ . Para cada  $t$  fixo, a superfície  $\psi(\mathbf{x}) = ct$  é a fronteira, no espaço  $\mathbb{R}_{\mathbf{x}}^3$ , atingida pelos campos electromagnéticos  $\mathbf{E}(t, \mathbf{x})$  e  $\mathbf{H}(t, \mathbf{x})$ , no instante  $t$ . Estas superfícies chamam-se as frentes de onda. Os valores de  $\mathbf{E}$  e  $\mathbf{H}$ , num ponto  $\mathbf{x}$  e no instante  $t = \psi(\mathbf{x})/c$ , constituem os campos da óptica geométrica no ponto  $\mathbf{x}$ :

$$\mathbf{E}^*(t, \mathbf{x}) = \mathbf{E}(\psi(\mathbf{x})/c, \mathbf{x}), \quad \mathbf{H}^*(t, \mathbf{x}) = \mathbf{H}(\psi(\mathbf{x})/c, \mathbf{x})$$

Os raios electromagnéticos são as curvas ao longo das quais a energia do campo da óptica geométrica flui. Em meios isotrópicos, esses raios são as trajectórias ortogonais, no espaço  $\mathbb{R}_{\mathbf{x}}^3$ , à família de frentes de onda  $\psi = ct$ .

Vamos agora aplicar as relações integrais (2.7.7), (2.7.8) e (2.7.13), (2.7.14) ao problema seguinte:

- ♣ **Problema 2.3** ... Seja  $\varphi(t, \mathbf{x}) = 0$  uma hipersuperfície, em  $\mathbb{R}^4$ , na qual  $\epsilon, \mu, \mathbf{E}$  ou  $\mathbf{H}$  são descontínuos. Deduzir as **condições de descontinuidade** a que devem obedecer os campos  $\mathbf{E}$  e  $\mathbf{H}$  em  $\varphi = 0$ .

♣.

Para isso consideremos uma hipersuperfície fechada  $\Gamma$ , que é dividida em duas partes  $\Gamma_1$  e  $\Gamma_2$ , por  $\varphi = 0$ . Seja  $\Gamma_0$  a parte de  $\varphi = 0$  que está dentro de  $\Gamma$  (ver a figura 2.12).

Figure 2.12: .

A normal a  $\varphi = 0$  é proporcional a **grad**  $\varphi = (\varphi_t, \varphi_{\mathbf{x}})$ . Se  $p \in \{\varphi = 0\}$ , representamos por  $[\mathbf{E}](p)$  a descontinuidade de  $\mathbf{E}$  em  $p$ :

$$[\mathbf{E}](p) \stackrel{\text{def}}{=} \underbrace{\lim_{\substack{q \rightarrow p \\ q \in V_1}} \mathbf{E}(q)}_{\mathbf{E}_1(p)} - \underbrace{\lim_{\substack{q \rightarrow p \\ q \in V_2}} \mathbf{E}(q)}_{\mathbf{E}_2(p)} \quad (2.7.15)$$

e análogamente para  $[\mathbf{H}]$ ,  $[\epsilon]$  ou  $[\mu]$ .

Vamos agora aplicar a relação (2.7.13) à hipersuperfície fechada  $\Gamma = \Gamma_1 + \Gamma_2$ :

$$\int_{\Gamma_1 + \Gamma_2} \left( \mathbf{N} \times \mathbf{H} - \frac{\epsilon}{c} N_0 \mathbf{E} \right) d\sigma = \mathbf{0} \quad (2.7.16)$$

Mas (2.7.13) deverá ser também válida em  $\Gamma_1 + \Gamma_0$ :

$$\begin{aligned} \int_{\Gamma_1 + \Gamma_0} \left( \mathbf{N} \times \mathbf{H} - \frac{\epsilon}{c} N_0 \mathbf{E} \right) d\sigma &= \int_{\Gamma_1} \left( \mathbf{N} \times \mathbf{H} - \frac{\epsilon}{c} N_0 \mathbf{E} \right) d\sigma + \int_{\Gamma_0} \left( \mathbf{N} \times \mathbf{H} - \frac{\epsilon}{c} N_0 \mathbf{E} \right) d\sigma \\ &= \int_{\Gamma_1} \left( \mathbf{N} \times \mathbf{H} - \frac{\epsilon}{c} N_0 \mathbf{E} \right) d\sigma \\ &\quad + \int_{\Gamma_0} \left( \varphi_{\mathbf{x}} \times \mathbf{H}_1 - \frac{\epsilon_1}{c} \varphi_t \mathbf{E}_1 \right) \frac{d\sigma}{\sqrt{\varphi_t^2 + \varphi_{\mathbf{x}}^2}} = \mathbf{0} \end{aligned} \quad (2.7.17)$$

já que, em  $\Gamma_0$ , se tem:

$$N = \left( N_0 = \frac{\varphi_t}{\sqrt{\varphi_t^2 + \varphi_{\mathbf{x}}^2}}, \mathbf{N} = \frac{\varphi_{\mathbf{x}}}{\sqrt{\varphi_t^2 + \varphi_{\mathbf{x}}^2}} \right)$$

Como, por outro lado, (2.7.13) deverá ser também válida em  $\Gamma_2 + \Gamma_0$ , e como agora, em  $\Gamma_0$  se tem:

$$N = \left( N_0 = -\frac{\varphi_t}{\sqrt{\varphi_t^2 + \varphi_x^2}}, \mathbf{N} = -\frac{\varphi_x}{\sqrt{\varphi_t^2 + \varphi_x^2}} \right)$$

vem que:

$$\begin{aligned} \int_{\Gamma_2 + \Gamma_0} \left( \mathbf{N} \times \mathbf{H} - \frac{\epsilon}{c} N_0 \mathbf{E} \right) d\sigma &= \int_{\Gamma_2} \left( \mathbf{N} \times \mathbf{H} - \frac{\epsilon}{c} N_0 \mathbf{E} \right) d\sigma + \int_{\Gamma_0} \left( \mathbf{N} \times \mathbf{H} - \frac{\epsilon}{c} N_0 \mathbf{E} \right) d\sigma \\ &= \int_{\Gamma_2} \left( \mathbf{N} \times \mathbf{H} - \frac{\epsilon}{c} N_0 \mathbf{E} \right) d\sigma \\ &\quad - \int_{\Gamma_0} \left( \varphi_x \times \mathbf{H}_2 - \frac{\epsilon_2}{c} \varphi_t \mathbf{E}_2 \right) \frac{d\sigma}{\sqrt{\varphi_t^2 + \varphi_x^2}} = \mathbf{0} \end{aligned} \quad (2.7.18)$$

Agora adicionamos as duas equações (2.7.17) e (2.7.18) e subtraímos a equação (2.7.16), para obter:

$$\int_{\Gamma_0} \left( \varphi_x \times (\mathbf{H}_2 - \mathbf{H}_1) - \frac{\varphi_t}{c} (\epsilon_2 \mathbf{E}_2 - \epsilon_1 \mathbf{E}_1) \right) \frac{d\sigma}{\sqrt{\varphi_t^2 + \varphi_x^2}} = \mathbf{0} \quad (2.7.19)$$

e como (2.7.19) deverá verificar-se para toda a parte  $\Gamma_0$  de  $\{\varphi = 0\}$ , deveremos ter:

$$\varphi_x \times [\mathbf{H}] - \frac{\varphi_t}{c} [\epsilon \mathbf{E}] = \mathbf{0} \quad (2.7.20)$$

Consideremos agora a equação integral (2.7.7):  $\int_{\Gamma} \epsilon (\mathbf{E} \cdot \mathbf{N}) d\sigma$ . Aplicando-a às superfícies  $\Gamma_1 + \Gamma_2$ ,  $\Gamma_1 + \Gamma_0$  e  $\Gamma_2 + \Gamma_0$ , obtemos sucessivamente:

$$\begin{aligned} \int_{\Gamma_1} \epsilon (\mathbf{E} \cdot \mathbf{N}) d\sigma + \int_{\Gamma_2} \epsilon (\mathbf{E} \cdot \mathbf{N}) d\sigma &= 0 \\ \int_{\Gamma_1} \epsilon (\mathbf{E} \cdot \mathbf{N}) d\sigma + \int_{\Gamma_0} (\epsilon_1 \mathbf{E}_1 \cdot \varphi_x) \frac{d\sigma}{\sqrt{\varphi_t^2 + \varphi_x^2}} &= 0 \\ \int_{\Gamma_2} \epsilon (\mathbf{E} \cdot \mathbf{N}) d\sigma - \int_{\Gamma_0} (\epsilon_2 \mathbf{E}_2 \cdot \varphi_x) \frac{d\sigma}{\sqrt{\varphi_t^2 + \varphi_x^2}} &= 0 \end{aligned} \quad (2.7.21)$$

donde deduzimos que:

$$\int_{\Gamma_0} (\epsilon_2 \mathbf{E}_2 - \epsilon_1 \mathbf{E}_1) \cdot \varphi_x \frac{d\sigma}{\sqrt{\varphi_t^2 + \varphi_x^2}} = 0$$

e portanto:

$$[\epsilon \mathbf{E}] \cdot \varphi_x = 0 \quad (2.7.22)$$

Da mesma forma, usando as outras duas equações integrais, obtemos mais duas equações (basta trocar  $\epsilon$  com  $\mu$  e  $\mathbf{E}$  com  $-\mathbf{H}$ ). Resumindo toda esta discussão, temos a seguinte:

- **♣ Proposição 2.12** ... Um campo electromagnético que seja descontínuo numa hipersuperfície  $\varphi(t, \mathbf{x}) = 0$ , onde  $\mathbf{x} = (x, y, z)$ , deve satisfazer as seguintes condições de descontinuidade:

$$[\text{CD}] \dots \left\{ \begin{array}{l} \varphi_{\mathbf{x}} \times [\mathbf{H}] - \frac{\varphi_t}{c} [\epsilon \mathbf{E}] = \mathbf{0} \\ \varphi_{\mathbf{x}} \times [\mathbf{E}] + \frac{\varphi_t}{c} [\mu \mathbf{H}] = \mathbf{0} \\ \varphi_{\mathbf{x}} \cdot [\epsilon \mathbf{E}] = 0 \\ \varphi_{\mathbf{x}} \cdot [\mu \mathbf{H}] = 0 \end{array} \right. \quad (2.7.23)$$

$$\forall (t, \mathbf{x}) \in \Gamma_0 = \{\varphi = 0\}.$$

♣.

Vejamos alguns exemplos:

- **♣ Exemplo 2.5** ... Suponhamos que a superfície de descontinuidade é estacionária (independente do tempo  $t$ ):

$$\varphi(\mathbf{x}, t) = \psi(\mathbf{x}) = 0$$

e que  $\epsilon$  e  $\mu$  são descontínuos em  $\Sigma = \{\mathbf{x} \in \mathbb{R}^3 : \psi(\mathbf{x}) = 0\}$ .  $\Sigma$  representa pois a superfície de separação entre dois meios ópticos. Como neste caso  $\varphi_t = 0$  e  $\varphi_{\mathbf{x}} = \psi_{\mathbf{x}} = \nabla\psi$ , as equações (2.7.23) têm a forma:

$$\left\{ \begin{array}{l} \nabla\psi \times [\mathbf{H}] = \mathbf{0} = \nabla\psi \times [\mathbf{E}] \\ \nabla\psi \cdot [\epsilon \mathbf{E}] = \mathbf{0} = \nabla\psi \cdot [\mu \mathbf{H}] \end{array} \right. \quad (2.7.24)$$

$\forall \mathbf{x} \in \Sigma, \forall t \in \mathbb{R}$ . O vector  $\psi_{\mathbf{x}} = \nabla\psi$  é perpendicular a  $\Sigma$ . Portanto  $\nabla\psi \times [\mathbf{H}]$  e  $\nabla\psi \times [\mathbf{E}]$  são determinados pelas componentes tangenciais de  $\mathbf{H}$  e  $\mathbf{E}$ , respectivamente. Por outro lado,  $\nabla\psi \cdot [\epsilon \mathbf{E}]$  e  $\nabla\psi \cdot [\mu \mathbf{H}]$  são determinados pelas componentes normais de  $[\epsilon \mathbf{E}]$  e  $[\mu \mathbf{H}]$ , respectivamente. Portanto as equações (2.7.24) dizem que:

- as componentes tangenciais de  $\mathbf{H}$  e  $\mathbf{E}$  e as componentes normais de  $\epsilon \mathbf{E}$  e  $\mu \mathbf{H}$  são contínuas em qualquer superfície de descontinuidade estacionária de  $\epsilon$  e  $\mu$ .

• **♣ Exemplo 2.6** ... Suponhamos que  $\epsilon = \mu \equiv 1$  e que, no instante  $t = 0$ , os vectores  $\mathbf{E}(\mathbf{x}, 0)$  e  $\mathbf{H}(\mathbf{x}, 0)$  são nulos apenas numa pequena bola de raio  $\delta > 0$ , centrada na origem de  $\mathbb{R}^3$ . O que se espera é que, com o evoluir do tempo, este campo electromagnético se expanda de tal forma que, num certo instante  $t > 0$ , os vectores  $\mathbf{E}$  e  $\mathbf{H}$  sejam não nulos numa esfera de raio  $\delta + ct$ . Por outras palavras, o que se espera é que a superfície que separa a parte do espaço que ainda está em repouso, ainda sem a influência do campo electromagnético, da parte já excitada por esse campo, evolua no espaço quando o tempo avança. Uma superfície deste tipo chama-se sugestivamente uma **frente de onda**. Assim,

por exemplo, na situação atrás descrita, as frentes de onda serão as esferas concêntricas de equação:

$$\varphi(\mathbf{x}, t) = \|\mathbf{x}\| - \delta - ct = 0$$

Se os valores iniciais dos campos  $\mathbf{E}(\mathbf{x}, 0)$  ou  $\mathbf{H}(\mathbf{x}, 0)$  são não nulos na esfera inicial de raio  $\delta$ , então esta esfera é uma superfície na qual o campo electromagnético é descontínuo. Quando  $t > 0$  os correspondentes valores do campo na frente de onda  $\varphi(\mathbf{x}, t) = \|\mathbf{x}\| - \delta - ct = 0$  são também não nulos, pelo que o campo electromagnético é ainda descontínuo nessa frente de onda. Um observador situado no ponto  $\mathbf{x}$  interpreta essa descontinuidade como um sinal que o atinge quando a frente de onda passa em  $\mathbf{x}$ .



### 2.7.4 A equação iconal da óptica geométrica

Suponhamos agora que  $\varphi(\mathbf{x}, t) = 0$  representa uma hipersuperfície onde  $\mathbf{E}$  ou  $\mathbf{H}$  são descontínuos. Admitamos ainda que, quer  $\epsilon$  quer  $\mu$ , são ambos contínuos num aberto que contem a hipersuperfície  $\varphi = 0$ , e vejamos a que condições deverá satisfazer  $\varphi$ .

As duas primeiras condições de descontinuidade (2.7.23) dão neste caso:

$$\begin{cases} \varphi_{\mathbf{x}} \times [\mathbf{H}] - \epsilon \frac{\varphi_t}{c} [\mathbf{E}] = \mathbf{0} \\ \varphi_{\mathbf{x}} \times [\mathbf{E}] + \mu \frac{\varphi_t}{c} [\mathbf{H}] = \mathbf{0} \end{cases} \quad (2.7.25)$$

que são seis equações escalares lineares homogêneas para as componentes de  $[\mathbf{E}]$  e  $[\mathbf{H}]$ . Se este sistema admite soluções não nulas, como estamos a supôr que sim, então os coeficientes desse sistema deverão satisfazer certas condições que passamos a deduzir. Podemos supôr que  $\varphi_t \neq 0$ , caso contrário estaríamos na situação descrita no exemplo 2.5. Notemos ainda que se  $\mathbf{E} = \mathbf{0}$  então  $\mathbf{H} = \mathbf{0}$  e vice-versa, e, por isso, podemos supôr que ambos  $\mathbf{E}$  e  $\mathbf{H}$  são não nulos em  $\varphi = 0$ .

Resolvendo a primeira equação em (2.7.25), em ordem a  $\mathbf{E}$ , e substituindo na segunda obtemos:

$$\varphi_{\mathbf{x}} \times (\varphi_{\mathbf{x}} \times [\mathbf{H}]) + \frac{\epsilon\mu}{c^2} \varphi_t^2 [\mathbf{H}] = \mathbf{0}$$

Finalmente, aplicando a igualdade de Lagrange, obtemos:

$$(\varphi_{\mathbf{x}} \cdot [\mathbf{H}]) \varphi_{\mathbf{x}} - \varphi_{\mathbf{x}}^2 [\mathbf{H}] + \frac{\epsilon\mu}{c^2} \varphi_t^2 [\mathbf{H}] = \mathbf{0} \quad (2.7.26)$$

Mas a última equação em (2.7.23) diz-nos que (atendendo a que  $\mu$  é contínua)  $\varphi_{\mathbf{x}} \cdot [\mathbf{H}] = 0$ , e portanto, uma vez que  $\mathbf{H} \neq \mathbf{0}$ , obtemos finalmente a equação:

$$\varphi_{\mathbf{x}}^2 - \frac{\epsilon\mu}{c^2} \varphi_t^2 = 0 \quad \text{em} \quad \varphi = 0 \quad (2.7.27)$$

onde  $\varphi_{\mathbf{x}}^2 = \varphi_{\mathbf{x}} \cdot \varphi_{\mathbf{x}}$ .

De facto esta equação não é uma verdadeira PDE para  $\varphi$ , já que apenas deverá verificar-se em  $\varphi = 0$ . Mas como estamos a supôr que  $\varphi_t \neq 0$ , podemos resolver  $\varphi(\mathbf{x}, t) = 0$  explicitamente em ordem a  $t$ , para obter:

$$\varphi(\mathbf{x}, t) = \psi(\mathbf{x}) - ct = 0$$

A equação (2.7.27) fica agora na forma:

$$(\nabla\psi)^2 - \epsilon\mu = 0 \quad (2.7.28)$$

ou mais explicitamente:

$$\psi_x^2 + \psi_y^2 + \psi_z^2 - \epsilon\mu = 0 \quad (2.7.29)$$

que é agora uma verdadeira PDE para  $\psi = \psi(\mathbf{x})$ , que se diz a **equação eikonal da óptica geométrica**. As respectivas soluções são as frentes de onda da óptica geométrica.

No que se segue, vamos concentrar a nossa atenção nos valores do campo electromagnético em  $\varphi = 0$ , ou, equivalentemente, nos valores que eles tomam sobre as frentes de onda  $\psi(x, y, z) = ct$ , à medida que elas avançam no espaço, quando  $t$  cresce. Estes valores particulares do campo chamar-se-ão **sinais**. Estes sinais constituem o chamado **campo da óptica geométrica**.

Os valores do campo electromagnético  $\mathbf{E}$  e  $\mathbf{H}$  na frente de onda  $\psi(\mathbf{x}) = ct$ , podem ser representados pelas funções vectoriais:

$$\begin{aligned} \mathbf{E}^*(\mathbf{x}) &= \mathbf{E}\left(\mathbf{x}, \frac{1}{c}\psi(\mathbf{x})\right) \\ \mathbf{H}^*(\mathbf{x}) &= \mathbf{H}\left(\mathbf{x}, \frac{1}{c}\psi(\mathbf{x})\right) \end{aligned} \quad (2.7.30)$$

e portanto  $\mathbf{E}^*(\mathbf{x})$  e  $\mathbf{H}^*(\mathbf{x})$  dão o valor do sinal que é observado no ponto  $\mathbf{x}$ , no instante  $t = \frac{1}{c}\psi(\mathbf{x})$ .

Recordemos agora que os valores dos campos  $\mathbf{E}$  e  $\mathbf{H}$ , em  $\varphi = 0$ , são exactamente os valores das descontinuidades destes campos. Portanto eles devem verificar as condições de descontinuidade (2.7.23). Em particular, as duas primeiras equações dizem que, atendendo a (2.7.30), e ao facto de estarmos a admitir que  $\epsilon$  e  $\mu$  são contínuas:

$$\begin{aligned} \varphi_{\mathbf{x}} \times \mathbf{H}^* - \frac{\varphi_t}{c} \epsilon \mathbf{E}^* &= \mathbf{0} \\ \varphi_{\mathbf{x}} \times \mathbf{E}^* + \frac{\varphi_t}{c} \mu \mathbf{H}^* &= \mathbf{0} \end{aligned} \quad (2.7.31)$$

e, como  $\varphi = \psi - ct$ , o que implica que  $\varphi_t = -c$ , vem que:

$$\begin{aligned} \nabla\psi \times \mathbf{H}^* + \epsilon \mathbf{E}^* &= \mathbf{0} \\ \nabla\psi \times \mathbf{E}^* - \mu \mathbf{H}^* &= \mathbf{0} \end{aligned} \quad (2.7.32)$$

Tomando o produto interno destas duas equações com  $\nabla\psi$ , e, em seguida, o produto interno da primeira com  $\mathbf{H}^*$  ou da segunda com  $\mathbf{E}^*$ , obtemos:

$$\begin{aligned} \nabla\psi \cdot \mathbf{E}^* &= 0 \\ \nabla\psi \cdot \mathbf{H}^* &= 0 \\ \mathbf{H}^* \cdot \mathbf{E}^* &= 0 \end{aligned} \quad (2.7.33)$$

isto é:

- *Os vectores  $\mathbf{E}^*$ ,  $\mathbf{H}^*$  e  $\nabla\psi$  são sempre ortogonais. Além disso, como  $\nabla\psi$  é perpendicular à frente de onda  $\psi = ct$ , os vectores  $\mathbf{E}^*$  e  $\mathbf{H}^*$  são tangentes à frente de onda  $\psi = ct$ .*

Notemos que as equações (2.7.32) são 6 equações lineares homogêneas nas 6 componentes de  $\mathbf{E}^*$  e  $\mathbf{H}^*$ . Como  $\psi - ct = 0$  é uma superfície de descontinuidade, esses campos  $\mathbf{E}^*$  e  $\mathbf{H}^*$  são não nulos (se um é nulo, o outro também o é). Portanto, o determinante desse sistema terá de ser nulo, o que implica uma condição para a função  $\psi$ . Podemos obter essa condição resolvendo a segunda equação em (2.7.32) em ordem a  $\mathbf{H}^*$  e substituindo na primeira. O resultado é:

$$\nabla\psi \times (\nabla\psi \times \mathbf{E}^*) + \epsilon\mu \mathbf{E}^* = \mathbf{0}$$

Expandindo isto pela regra de Laplace, e usando (2.7.33), obtemos:

$$[(\nabla\psi)^2 - \epsilon\mu] \mathbf{E}^* = \mathbf{0}$$

e, finalmente, como  $\mathbf{E}^* \neq \mathbf{0}$ :

$$\psi_x^2 + \psi_y^2 + \psi_z^2 = \epsilon\mu \quad (2.7.34)$$

Concluindo: as frentes de onda  $\psi = ct$  devem ser solução da PDE de primeira ordem (2.7.34), que é a chamada **equação eikonal da óptica geométrica**.

# Capítulo 3

## Geometria Simpléctica e Mecânica

De acordo com a segunda Lei de Newton, uma partícula de massa  $m > 0$  move-se, sob a acção de um potencial  $V : \mathbb{R}^3 \rightarrow \mathbb{R}$ , segundo uma trajectória  $\mathbf{x}(t)$  em  $\mathbb{R}^3$ , de tal forma que:

$$\boxed{m\ddot{\mathbf{x}} = -\nabla V(\mathbf{x}) \quad \mathbf{x} \in Q = \mathbb{R}^3} \quad (3.0.1)$$

Por exemplo, o oscilador harmónico o potencial é quadrático e...

Se introduzimos os “momentos”  $p_i = m\dot{x}^i$ , e a energia total:

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2m}\|\mathbf{p}\|^2 + V(\mathbf{x}) \quad (\mathbf{x}, \mathbf{p}) \in \mathbb{R}^3 \times \mathbb{R}^3$$

então a segunda lei de Newton é equivalente às equações de Hamilton:

$$\boxed{\begin{cases} \dot{x}^i = \frac{\partial H}{\partial p_i} \\ \dot{p}_i = \frac{\partial H}{\partial x^i} \end{cases}} \quad (3.0.2)$$

Estudemos este sistema de equações de primeira ordem, para uma função  $H$ . Qualquer  $H(\mathbf{x}, \mathbf{p})$ ,  $(\mathbf{x}, \mathbf{p}) \in \mathbb{R}^3 \times \mathbb{R}^3$ . Para isso introduzimos a matriz:

$$\mathbf{J} = \begin{bmatrix} \mathbf{0} & -\mathbf{1}_3 \\ \mathbf{1}_3 & \mathbf{0} \end{bmatrix}$$

onde  $\mathbf{1}_3$  é a matriz identidade ( $3 \times 3$ ), e notamos que as equações (3.0.2) se podem escrever na forma:

$$\dot{\mathbf{x}} = -\mathbf{J} \nabla H(\mathbf{x}) \quad \mathbf{x} = (\mathbf{x}, \mathbf{p})$$

Se definimos o campo de vectores:

$$X_H \stackrel{\text{def}}{=} -\mathbf{J} \nabla H$$

então  $\mathbf{x}(t)$  satisfaz as equações de Hamilton se e só se  $\mathbf{x}(t)$  é uma curva integral de  $X_H$ , isto é, se  $\dot{\mathbf{x}}(t) = X_H(\mathbf{x}(t))$ .

A relação entre  $X_H$  e  $H$  pode ser explicada como segue: consideramos a forma bilinear anti-simétrica  $\omega$  em  $\mathbb{R}^3 \times \mathbb{R}^3$ , definida por:

$$\omega((\mathbf{x}_1, \mathbf{p}_1), (\mathbf{x}_2, \mathbf{p}_2)) \stackrel{\text{def}}{=} [\mathbf{x}_1, \mathbf{p}_1] \mathbf{J} [\mathbf{x}_2, \mathbf{p}_2]^t = \mathbf{x}_1 \cdot \mathbf{p}_2 - \mathbf{x}_2 \cdot \mathbf{p}_1$$

onde  $(\mathbf{x}_1, \mathbf{p}_1), (\mathbf{x}_2, \mathbf{p}_2) \in \mathbb{R}^3 \times \mathbb{R}^3$ , e  $\cdot$  é o produto interno usual em  $\mathbb{R}^3$ . temos então que  $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^3 \times \mathbb{R}^3$ :

$$\omega(X_H(\mathbf{x}), \mathbf{y}) = dH_{\mathbf{x}}(\mathbf{y})$$

## 3.1 Variedades simplécticas

Começamos com alguns resultados sobre geometria simpléctica linear.

♣ **Definição 3.1** ... Um “espaço vectorial simpléctico real”  $(V, \omega)$  é constituído por um espaço vectorial real de dimensão finita, munido de uma forma simpléctica, i.e., uma forma bilinear  $\omega : V \times V \rightarrow \mathbb{R}$  anti-simétrica e não degenerada.

Note que a dimensão de  $V$  tem que ser par. Se  $\{\mathbf{e}_i\}$  é uma base de  $V$ , e se  $\{\mathbf{e}^i\}$  é a respectiva base dual, então  $\omega = \omega_{ij} \mathbf{e}^i \otimes \mathbf{e}^j$ , e a matriz anti-simétrica  $[\omega_{ij} = \omega(\mathbf{e}_i, \mathbf{e}_j)]$  é não singular. Como  $\omega$  é não degenerada, a aplicação “bemol”:

$$\flat : V \rightarrow V^*, \quad \mathbf{v} \mapsto (\mathbf{v}^\flat : \mathbf{u} \mapsto \omega(\mathbf{v}, \mathbf{u}))$$

é um isomorfismo linear.

### Exemplos...

(i).  $V = W \times W^*$ , onde  $W$  é um espaço vectorial real de dimensão finita, e  $\omega$  é a forma simpléctica em  $V$ , definida por:

$$\omega((\mathbf{w}_1, \alpha_1), (\mathbf{w}_2, \alpha_2)) \stackrel{\text{def}}{=} \alpha_2(\mathbf{w}_1) - \alpha_1(\mathbf{w}_2)$$

onde  $\mathbf{w}_1, \mathbf{w}_2 \in W$ ,  $\alpha_1, \alpha_2 \in W^*$ . Se  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  é uma base para  $W$  e se  $\{\mathbf{e}^1, \dots, \mathbf{e}^n\}$  é a respectiva base dual, então a matriz de  $\omega$  na base  $\{(\mathbf{e}_1, \mathbf{0}), \dots, (\mathbf{e}_n, \mathbf{0}), (\mathbf{0}, \mathbf{e}^1), \dots, (\mathbf{0}, \mathbf{e}^n)\}$  de  $V$ , é a matriz:

$$\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{1}_n \\ -\mathbf{1}_n & \mathbf{0} \end{bmatrix}$$

(ii).  $V = W \times W$ , onde  $W$  é um espaço vectorial real de dimensão finita munido de um produto interno  $\cdot$ , e  $\omega$  é a forma simpléctica em  $V$ , definida por:

$$\omega((\mathbf{w}_1, \mathbf{w}_2), (\mathbf{z}_1, \mathbf{z}_2)) \stackrel{\text{def}}{=} \mathbf{w}_1 \cdot \mathbf{z}_2 - \mathbf{w}_2 \cdot \mathbf{z}_1$$

onde  $\mathbf{w}_1, \mathbf{w}_2, \mathbf{z}_1, \mathbf{z}_2 \in W$ .

♣ **Teorema 3.1** ... Um espaço vectorial simpléctico  $(V, \omega)$  admite sempre uma base simpléctica, isto é, uma base  $\{\mathbf{e}_1, \dots, \mathbf{e}_n, \mathbf{f}_1, \dots, \mathbf{f}_n\}$  que satisfaz as condições seguintes:

$$\omega(\mathbf{e}_i, \mathbf{e}_j) = 0 = \omega(\mathbf{f}_i, \mathbf{f}_j) \quad e \quad \omega(\mathbf{e}_i, \mathbf{f}_j) = \delta_{ij}$$

Portanto a matriz de  $\omega$  nessa base é:

$$\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{1}_n \\ -\mathbf{1}_n & \mathbf{0} \end{bmatrix}$$

**Dem.:** Começemos com um  $\mathbf{e}_1 \neq \mathbf{0}$ . Como  $\omega$  é não degenerada, existe  $\mathbf{f}_1 \in V - \{\mathbf{0}\}$  tal que  $\omega(\mathbf{e}_1, \mathbf{f}_1) = 1$ . O par  $(\mathbf{e}_1, \mathbf{f}_1)$  gera um subespaço  $V_2$  de dimensão 2. Consideremos então:

$$V_2^\perp \stackrel{\text{def}}{=} \{\mathbf{v} \in V : \omega(\mathbf{v}, \mathbf{e}_1) = 0 = \omega(\mathbf{v}, \mathbf{f}_1)\}$$

$(V_2^\perp, \omega|_{V_2^\perp})$  é um espaço vectorial simpléctico, e podemos proceder por indução sobre a dimensão de  $V$ , CQD.

Vamos agora introduzir o conceito de variedade simpléctica:

♣ **Definição 3.2** ... Uma “**variedade simpléctica**”  $(M, \omega)$  é uma variedade diferenciável  $C^\infty$ , munida de uma forma simpléctica, isto é, de uma 2-forma diferencial  $\omega$  fechada e não degenerada.

Note que a dimensão de  $M$  tem que ser par, digamos  $2n$ . Além disso,  $\forall x \in M$ , o par  $(T_x M, \omega_x)$  é espaço vectorial simpléctico, e como  $\omega$  é não degenerada a  $2n$ -forma:

$$\mu_\omega \stackrel{\text{def}}{=} \underbrace{\omega \wedge \dots \wedge \omega}_{n \text{ factores}} \quad (3.1.1)$$

é uma forma volume em  $M$ . Se  $\dim M = 2$ , uma variedade simpléctica é o mesmo que uma superfície com uma forma de área.

**Exemplo ... Estrutura simpléctica canónica em  $M = T^*Q$**

Seja  $Q$  uma variedade diferenciável  $C^\infty$  (“**espaço de configuração**”), e considere o respectivo fibrado cotangente  $M = T^*Q$  (“**espaço das fases**”). Em  $T^*Q$  define-se uma 1-forma diferencial canónica  $\theta$ , dita a “**forma de Liouville**”, através de:

$$\theta_{\alpha_q}(V_{\alpha_q}) \stackrel{\text{def}}{=} \langle \alpha_q, T\pi(V_{\alpha_q}) \rangle \quad (3.1.2)$$

onde  $\alpha_q \in T^*Q$ ,  $V_{\alpha_q} \in T_{\alpha_q}(T^*Q)$ , e  $\pi : T^*Q \rightarrow Q$  é a projecção canónica.

♣ **Exercício 3.1** (i). Sejam  $(x^1, \dots, x^n)$  um sistema de coordenadas locais em  $Q$ , e  $(x^1, \dots, x^n, p_1, \dots, p_n)$  o correspondente sistema de coordenadas locais em  $T^*Q$ . Mostre que a expressão local de  $\theta$  é:

$$\theta = p_i dx^i = \mathbf{p} \, d\mathbf{x}$$

(ii). Mostre que  $\theta$  tem a propriedade seguinte: “ $\theta$  é a única 1-forma em  $T^*Q$  tal que  $\alpha^*(\theta) = \alpha$ , para toda a 1-forma diferencial  $\alpha : Q \rightarrow T^*Q$ ”.

Definamos agora uma 2-forma  $\omega$  em  $T^*Q$  pondo:

$$\omega \stackrel{\text{def}}{=} -d\theta \tag{3.1.3}$$

♣ **Exercício 3.2** Mostre  $\omega$  é uma forma simplética em  $T^*Q$ .

A estrutura simplética assim obtida, diz-se a estrutura simplética canónica em  $T^*Q$ . A expressão local de  $\omega$  nas coordenadas locais  $(x^1, \dots, x^n, p_1, \dots, p_n)$ , é:

$$\omega = dx^i \wedge dp_i = -d\mathbf{p} \wedge d\mathbf{x}$$

Note que todas as cartas de um atlas trivializador canónico de  $T^*Q$  são cartas simpléticas, isto é, a representação local de  $\omega$  numa qualquer dessas cartas, tem a forma diagonal  $dx^i \wedge dp_i$ . Toda a variedade simplética  $(M, \omega)$ , admite um atlas simplético. Com efeito é válido o teorema seguinte:

♣ **Teorema 3.2** (“Teorema de Darboux” ... Suponha que  $\omega \in \Omega^2(M)$  é uma 2-forma não degenerada numa variedade de dimensão  $2n$ . Então  $\omega$  é fechada,  $d\omega = 0$ , se e só se existe uma carta  $(U; x^1, \dots, x^n, y^1, \dots, y^n)$  em torno de cada ponto  $p \in M$ , tal que:

$$\omega|_U = \sum_i dx^i \wedge dy^i$$

Portanto as formas simpléticas são sempre localmente “planas”, em contraste com as métricas riemannianas, por exemplo...

**Exemplo ... Estrutura simplética em  $M = TQ$ ;  $(Q, g)$  variedade riemanniana**

Seja  $(Q, g)$  uma variedade (pseudo-) riemanniana de dimensão  $n$ . A métrica  $g$  induz um isomorfismo “bemol” de fibrados vectoriais:

$$g^b : TQ \longrightarrow T^*Q$$

dado por:

$$g^b : X_q \mapsto g^b(X_q) = X_q^b \left( : Y_q \mapsto g_q(X_q, Y_q) \right) \quad X_q, Y_q \in T_qM$$

Em coordenadas locais  $(x^i, \dot{x}^i)$  e  $(x^i, p_i)$  para  $TQ$  e  $T^*Q$ , respectivamente,  $g^b$  é dada por:

$$\boxed{g^b : (x^i, \dot{x}^i) \mapsto (x^i, p_i = g_{ij}(q) \dot{x}^j)} \quad (3.1.4)$$

Considere a forma de Liouville  $\theta$  em  $T^*Q$ , e o pull-back:

$$\theta_g \stackrel{\text{def}}{=} (g^b)^*(\theta)$$

Nas coordenadas atrás referidas, temos que:

$$\theta_g = (g^b)^*(p_i dx^i) = (p_i \circ g^b) d(x^i \circ g^b) = g_{ij}(q) \dot{x}^j dx^i$$

Se considerarmos agora  $\omega_g = -d\theta_g = -d(g^b)^*(\theta) = (g^b)^*(-d\theta)$ , então  $\omega_g$  é uma 2-forma fechada não degenerada, e portanto  $TQ, \omega_g$  é uma variedade simpléctica. Em coordenadas locais:

$$\boxed{\omega_g = g_{ij}(q) dx^i \wedge d\dot{x}^j + \frac{\partial g_{ij}}{\partial x^k}(q) \dot{x}^i dx^j \wedge dx^k}$$

♣ **Exercício 3.3** ... Provar que  $(\theta_g)_{X_q}(V_{X_q}) = g_q(X_q, T\pi(V_{X_q}))$ ,  $\forall X_q \in TQ$  e  $\forall V_{X_q} \in T_{X_q}(TQ)$ .

♣ **Definição 3.3** ... Sejam  $(M, \omega_M)$  e  $(N, \omega_N)$  duas variedades simplécticas. Uma aplicação diferenciável  $F : M \rightarrow N$ , diz-se **“canónica ou simpléctica”**, se  $F^*\omega_N = \omega_M$ .

♣ **Exercício 3.4** ... Seja  $Q$  uma variedade e  $F \in \text{Diff}(Q)$  um difeomorfismo de  $Q$ . nestas condições, define-se o **“levantamento de  $F$  a  $T^*Q$ ”**, como sendo a aplicação  $T^*F : T^*Q \rightarrow T^*Q$ , definida através do diagrama seguinte:

$$\begin{array}{ccc} T^*Q & \xrightarrow{T^*F} & T^*Q \\ \pi \downarrow & & \downarrow \pi \\ Q & \xleftarrow{F} & Q \end{array}$$

isto é:

$$\boxed{T^*F : \alpha_q \mapsto T^*F(\alpha_q) : (X \mapsto T^*F(\alpha_q)(X) \stackrel{\text{def}}{=} \langle \alpha_q, TF(X) \rangle)} \quad (3.1.5)$$

$\forall \alpha_q \in T^*Q$  e  $\forall X \in T_{F^{-1}(q)}Q$ .

Mostre que  $T^*F$  é uma transformação simpléctica e que  $(T^*F)^*\theta = \theta$ , onde  $\theta$  é a forma de Liouville em  $T^*Q$ .

♣ **Exercício 3.5** ... Seja  $(Q, g)$  uma variedade riemanniana e  $F : Q \rightarrow \mathbb{R}$  uma função. Provar que  $TF : TQ \rightarrow TQ$  é simpléctica relativamente à forma simpléctica  $\omega_g$  e que  $(TF)^*\theta_g = \theta_g$ .

Vamos agora introduzir o conceito de sistema Hamiltoniano:

♣ **Definição 3.4** ... Um “**Sistema Hamiltoniano**”  $(M, \omega, H)$  é constituído por uma variedade simpléctica  $(M, \omega)$  e por uma função  $H \in C^\infty(M)$ , que se diz o “**Hamiltoniano**” (ou a energia) do sistema.

O campo de vectores  $X_H \in \mathfrak{X}(M)$  definido pela condição:

$$\boxed{i_{X_H} = dH} \quad (3.1.6)$$

isto é,  $\omega(X_H, Y) = dH(Y)$ ,  $\forall Y \in \mathfrak{X}(M)$ , diz-se o campo de vectores Hamiltoniano do sistema. O seu fluxo  $F_t^H = F_t^{X_H}$  diz-se o fluxo Hamiltoniano do sistema.

Quando  $(M, g)$  é uma variedade riemanniana e  $F : M \rightarrow \mathbb{R}$  é uma função  $C^\infty$ , recorde que se define o campo gradiente de  $F$ ,  $\mathbf{grad} F \in \mathfrak{X}(M)$ , através de:

$$g(\mathbf{grad} F, Y) = dF(Y) \quad \forall Y \in \mathfrak{X}(M)$$

A definição anterior é portanto formalmente análoga. No entanto, a anti-simetria da forma simpléctica conduz a propriedades conservativas, enquanto a simetria da métrica conduz a propriedades dissipativas do campo gradiente.

♣ **Exercício 3.6** ... (i). Seja  $(M, g)$  uma variedade riemanniana e  $F : M \rightarrow \mathbb{R}$  uma função  $C^\infty$ . Mostrar que a longo das órbitas não singulares de  $X = \mathbf{grad} F$ ,  $F$  é estritamente crescente e que portanto  $X$  não possui órbitas fechadas.

(ii). Suponha que  $M$  é uma variedade riemanniana completa. Mostre que existe uma constante  $C > 0$ , tal que  $\|X(p)\| < C$ ,  $\forall p \in M$ . Mostrar que  $X$  é um campo completo.

Veamos agora a representação local de um campo Hamiltoniano numa carta simpléctica, com coordenadas canónicas  $(x^1, \dots, x^n, p_1, \dots, p_n)$ . Nesta carta  $\omega = \dot{x}^i \wedge dp_i$ . Suponhamos que:

$$X_H = a^i \frac{\partial}{\partial x^i} + b_i \frac{\partial}{\partial p_i}$$

Então  $i_{X_H} dx^i = dx^i(X_H) = a^i$ ,  $i_{X_H} dp_i = dp_i(X_H) = b_i$  e a identidade que define  $X_H$ ,  $i_{X_H} \omega = dH$  conduz aos cálculos seguintes:

$$\begin{aligned} i_{X_H} \omega &= i_{X_H}(\dot{x}^i \wedge dp_i) \\ &= \sum_i (i_{X_H} \dot{x}^i) \wedge dp_i - \sum_i \dot{x}^i \wedge (i_{X_H} dp_i) \\ &= \sum_i (a^i dp_i - b_i dx^i) \end{aligned}$$

Como por outro lado  $dH = \frac{\partial H}{\partial x^i} dx^i + \frac{\partial H}{\partial p_i} dp_i$ , deduzimos que:

$$a^i = \frac{\partial H}{\partial p_i} \quad e \quad b_i = -\frac{\partial H}{\partial x^i}$$

isto é, a expressão local de um campo Hamiltoniano  $X_H$  numa carta simpléctica, com coordenadas canónicas  $(x^1, \dots, x^n, p_1, \dots, p_n)$  é:

$$\boxed{X_H = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i} \frac{\partial}{\partial p_i}} \quad (3.1.7)$$

ou em forma vectorial:

$$X_H = \mathbf{J} \begin{bmatrix} \frac{\partial H}{\partial \mathbf{x}} \\ \frac{\partial H}{\partial \mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{1}_n \\ -\mathbf{1}_n & \mathbf{0} \end{bmatrix} \begin{bmatrix} \frac{\partial H}{\partial \mathbf{x}} \\ \frac{\partial H}{\partial \mathbf{p}} \end{bmatrix} \quad (3.1.8)$$

Assim se  $(q(t), p(t))$  é uma curva integral de  $X_H$ , então ela deverá verificar as chamadas **“equações de Hamilton”**:

$$\boxed{\begin{cases} \dot{q}^i = \frac{\partial H}{\partial p_i} \\ \dot{p}_i = -\frac{\partial H}{\partial x^i} \end{cases}} \quad (3.1.9)$$

que é um sistema de equações diferenciais de primeira ordem.

♣ **Teorema 3.3** ... Seja  $\text{Fl}_t^H = \text{Fl}_t^{X_H}$  o fluxo hamiltoniano do campo  $X_H$ . Então:

(i)...  $H(\text{Fl}_t^H(x)) \equiv$  constante em  $t$ . Isto é,  $H$  é constante ao longo das curvas integrais de  $X_H$  **“Lei da conservação de energia”**.

(ii)...  $(\text{Fl}_t^H)^* \omega = \omega$ , isto é, para cada  $t$  a aplicação de avanço no tempo  $t$ ,  $\text{Fl}_t^H$  é simpéctica.

**Dem.:** ... (i). Seja  $\alpha_x(t) = \text{Fl}_t^H(x)$  a curva integral que em  $t = 0$  passa em  $x \in M$ . Vem então que:

$$\frac{d}{dt} H(\alpha_x(t)) = dH_{\alpha_x(t)} \dot{\alpha}_x(t) = dH_{\alpha_x(t)}(X_H(\alpha_x(t))) = \omega(X_H(\alpha_x(t)), X_H(\alpha_x(t))) = 0$$

(ii).

$$\frac{d}{dt} (\text{Fl}_t^H)^* \omega = (\text{Fl}_t^H)^* \mathcal{L}_{X_H} \omega = (\text{Fl}_t^H)^* (i_{X_H} d\omega + di_{X_H} \omega) = (\text{Fl}_t^H)^* (0 + dd\omega) = 0$$

isto é,  $(\text{Fl}_t^H)^* \omega$  é constante em  $t$ , e como  $(\text{Fl}_0^H) = \text{Id}$ , vem que  $(\text{Fl}_t^H)^* \omega = \omega$ , CQD.

♣ **Definição 3.5** ... Seja  $(M, \omega)$  uma variedade simpléctica, e  $f, g \in C^\infty(M)$ , com campos  $H$  amiltonianos associados,  $X_f$  e  $X_g$  respectivamente. Define-se o “**parêntesis de Poisson**” de  $f$  e  $g$ , através de:

$$\boxed{\{f, g\} \stackrel{def}{=} \omega(X_f, X_g)} \quad (3.1.10)$$

Numa carta simpléctica, com coordenadas canónicas  $(x^i, p_i)$ , relativamente às quais a expressão local de  $\omega$  é  $\omega = dx^i \wedge dp_i$ , é fácil ver que:

$$\boxed{\{f, g\} = \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial x^i} = (\mathbf{grad} f)^t \mathbf{J} \mathbf{grad} g} \quad (3.1.11)$$

Como:

$$\mathcal{L}_{X_f} g = i_{X_f} dg = i_{X_f} i_{X_g} \omega = \omega(X_f, X_g) = -\omega(X_g, X_f) = -\mathcal{L}_{X_g} f$$

vemos que:

$$\{f, g\} = \mathcal{L}_{X_f} g = -\mathcal{L}_{X_g} f$$

e portanto,  $f$  é constante ao longo das órbitas de  $X_g$ , se e só se  $\{f, g\} = 0$ , se e só se  $g$  é constante ao longo das órbitas de  $X_f$ .

Consideremos agora um difeomorfismo  $\varphi : M \rightarrow N$  entre duas variedades simplécticas  $(M, \omega_M)$  e  $(N, \omega_N)$ . Então, como  $\varphi^*(\mathcal{L}_X \alpha) = \mathcal{L}_{\varphi^* X} \varphi^* \alpha$ ,  $\forall \alpha \in \Omega(N)$ ,  $\forall X \in \mathfrak{X}(N)$ , vemos que:

$$\varphi^* \{f, g\} = \varphi^*(\mathcal{L}_{X_f} g) = \mathcal{L}_{\varphi^* X_f} \varphi^* g$$

e por outro lado:

$$\{\varphi^* f, \varphi^* g\} = \mathcal{L}_{X_{\varphi^* f}} \varphi^* g$$

Portanto  $\varphi$  preserva o parêntesis de Poisson de duas funções  $f, g \in C^\infty(N)$ , se e só se  $\varphi^* X_f = X_{\varphi^* f}$ ,  $\forall f \in C^\infty(N)$ . Isto é,  $\varphi$  preserva o parêntesis de Poisson se e só se preserva as equações de Hamilton.

Por outro lado:

$$i_{X_{\varphi^* f}} \omega = d(\varphi^* f) = \varphi^*(df) = \varphi^*(i_{X_f} \omega) = i_{\varphi^* X_f} \varphi^* \omega$$

e como  $\omega$  é não degenerada, e  $\forall v \in T_x M$ ,  $v = X_h(x)$ , para alguma função  $h \in C^\infty(U)$ , definida numa vizinhança de  $x$ , concluímos que  $\varphi$  é simpléctica se e só se  $\varphi^* X_f = X_{\varphi^* f}$ ,  $\forall f \in C^\infty(N)$ . Fica assim demonstrado o seguinte teorema:

♣ **Teorema 3.4** ... Seja  $\varphi : M \rightarrow N$  um difeomorfismo entre duas variedades simplécticas  $(M, \omega_M)$  e  $(N, \omega_N)$ . Então as condições seguintes são equivalentes:

- $\varphi$  é simpléctica.
- $\varphi$  preserva o parêntesis de Poisson de duas quaisquer funções  $f, g \in C^\infty(N)$ , isto é,  $\varphi^* \{f, g\} = \{\varphi^* f, \varphi^* g\}$ .
- $\varphi^* X_f = X_{\varphi^* f} \forall f \in C^\infty(N)$ .

A Lei da conservação de energia pode ser generalizada do seguinte modo:

♣ **Teorema 3.5** ... Seja  $X_H$  um campo Hamiltoniano numa variedade simpléctica  $(M, \omega)$ , com fluxo  $\text{Fl}_t^H$ . Então  $\forall f \in C^\infty(M)$  tem-se que:

$$\boxed{\frac{d}{dt}(f \circ \text{Fl}_t^H) = \{f \circ \text{Fl}_t^H, H\}} \quad (3.1.12)$$

**Dem.:** ...

$$\frac{d}{dt}(f \circ \text{Fl}_t^H) = \frac{d}{dt}((\text{Fl}_t^H)^* f) = (\text{Fl}_t^H)^* \mathcal{L}_{X_H} f = \mathcal{L}_{X_H}(f \circ \text{Fl}_t^H) = \{f \circ \text{Fl}_t^H, H\}$$

CQD.

Como já vimos  $T^*Q$  tem uma estrutura simpléctica canónica. Portanto se  $Q$  representa o espaço de configuração de um sistema mecânico, é possível estudar campos hamiltonianos no espaço de fases  $T^*Q$ , e os respectivos fluxos. Vamos agora estudar um tipo especial de hamiltoniano, particularmente importantes em mecânica clássica - os chamados Hamiltonianos de tipo mecânico. Para isso consideremos uma variedade riemanniana  $(Q, g)$  - o espaço de configuração de um sistema mecânico. Como já vimos  $g$  induz um isomorfismo de fibrados vectoriais:

$$g^b : TQ \longrightarrow T^*Q$$

Este isomorfismo permite munir cada  $T_q^*Q$ , de um produto interno, notado por  $g_q^*$ , e definido por:

$$g_q^*(\alpha_q, \beta_q) \stackrel{\text{def}}{=} g_q((g^b)^{-1}\alpha_q, (g^b)^{-1}\beta_q)$$

$\forall q \in Q, \forall \alpha_q, \beta_q \in T_q^*Q$ . Se  $g_{ij}(q)$  são os coeficientes da métrica  $g$  num sistema de coordenadas locais  $x^i$ , em  $Q$ , de tal forma que:

$$g = g_{ij}(q) \dot{x}^i \dot{x}^j$$

então as componentes de  $g^*$  são  $g^{ij}$ , onde  $g^{ij} = (g_{ij})^{-1}$ . Isto é:

$$g^* = g^{ij}(q) p_i p_j$$

(recorde que  $p_i = g_{ij}(q) \dot{x}^j$ ). Com estas notações passemos à definição de sistema hamiltoniano de tipo mecânico.

♣ **Definição 3.6** ... Uma função  $H : T^*Q \rightarrow \mathbb{R}$  diz-se um hamiltoniano de tipo mecânico, se  $H$  é da forma:

$$\boxed{H = K + V \circ \pi} \quad (3.1.13)$$

onde  $K : T^*Q \rightarrow \mathbb{R}$ , dada por:

$$K(\alpha_q) = \frac{1}{2} g_q^*(\alpha_q, \alpha_q), \quad \alpha_q \in T_x^*Q$$

é a chamada “energia cinética” do sistema,  $V : Q \rightarrow \mathbb{R}$  é a energia potencial, e  $\pi : T^*Q \rightarrow Q$  é a projecção canónica.

O sistema  $(T^*Q, \omega, H = K + V \circ \pi)$  diz-se um “sistema hamiltoniano de tipo mecânico”, com “espaço de configuração”  $Q$ , “espaço de fases”  $T^*Q$ , “energia total”  $H$ , “energia cinética”  $K$  e “energia potencial”  $V$ .

Para comparar este tratamento da mecânica clássica com o tratamento usual, baseado em princípios variacionais (ponto de vista de Lagrange), vamos introduzir alguns conceitos. Para sistemas hamiltonianos de tipo mecânico, fomos conduzidos a um certo campo de vectores  $X$  em  $T^*Q$

## 3.2 Sistemas mecânicos com simetria. Aplicação momento. Redução

♣ **Definição 3.7** ... Um “sistema mecânico com simetria  $(Q, K, V, G)$ ”, é constituído por:

- Uma variedade diferenciável  $Q$  - o espaço de configuração do sistema.
- Uma métrica riemanniana  $g$  em  $Q$ , e  $K = \frac{1}{2}g$  é a energia cinética dessa métrica:  $K(v_q) = \frac{1}{2}g(v_q, v_q)$ ,  $v_q \in T_qQ$ .
- Uma energia potencial  $V \in C^\infty(Q)$ .
- Uma acção de simetria, isto é, uma acção  $C^\infty$  de um grupo de Lie  $G$ , que actua à esquerda de  $Q$ , como um grupo de isometrias da métrica  $g$ , e preservando também o potencial  $V$ . Portanto se  $\Phi : G \times Q \rightarrow Q$  é a referida acção:

$$V \circ \Phi_g = V \quad \forall g \in G$$

e:

$$g_q(T\Phi_g(v_q), T\Phi_g(v_q)) = g_q(v_q, w_q) \quad \forall v_q, w_q \in T_qQ, \forall q \in Q$$

Se  $\mathfrak{g}$  é a álgebra de Lie de  $G$ , para cada  $\xi \in \mathfrak{g}$ , define-se o gerador infinitesimal da acção  $\Phi$ , associado a  $\xi$ , como sendo o campo de vectores  $\xi_Q \in \mathfrak{X}(Q)$  definido por:

$$\xi_Q(q) \stackrel{\text{def}}{=} \left. \frac{d}{dt} \right|_{t=0} \Phi_{\exp t\xi}(q) \quad (3.2.1)$$

Temos assim uma aplicação natural:

$$\mathfrak{g} \rightarrow T_qQ \quad \xi \mapsto \xi_Q(q)$$

:

♣ **Definição 3.8** ... Seja  $(Q, K, V, G)$  um sistema mecânico com simetria. Define-se então a respectiva “aplicação momento”, como sendo a aplicação:

$$\mathbf{J} : TQ \longrightarrow \mathfrak{g}^*$$

definida através de:

$$\mathbf{J} : v_q \mapsto \left( \mathbf{J}(v_q) : \xi \mapsto \mathbf{J}(v_q)(\xi) \stackrel{\text{def}}{=} g_q(v_q, \xi_Q(q)) \right) \quad (3.2.2)$$

Para cada  $\xi \in \mathfrak{g}$ , define-se ainda uma aplicação  $\widehat{\mathbf{J}}_\xi : TQ \rightarrow \mathbb{R}$ , através de:

$$\boxed{\widehat{\mathbf{J}}_\xi(v_q) \stackrel{\text{def}}{=} \mathbf{J}(v_q)(\xi)} \quad (3.2.3)$$

Se  $\Phi : G \times Q \rightarrow Q$  é a acção de simetria referida na definição anterior, então  $\Phi$  induz uma acção  $\Phi^T$  em  $TQ$ , dita a “**acção tangente**”, definida por:

$$\Phi^T(g, v_q) \stackrel{\text{def}}{=} T\Phi_g(v_q)$$

de tal forma que  $\pi$  é  $G$ -equivariante:

$$\begin{array}{ccc} TQ & \xrightarrow{\Phi_g^T} & TQ \\ \pi \downarrow & & \downarrow \pi \\ Q & \xrightarrow{\Phi_g} & Q \end{array}$$

Se  $\xi_{TQ} \in \mathfrak{X}(TQ)$  representa o gerador infinitesimal desta acção tangente, associado a um elemento  $\xi \in \mathfrak{g}$ :

$$\xi_{TQ}(v_q) = \left. \frac{d}{dt} \right|_{t=0} \Phi_{\exp t\xi}^T(v_q) \in T_{v_q}TQ$$

então a  $G$ -equivariância de  $\pi$  implica que:

$$\boxed{T\pi \circ \xi_{TQ} = \xi_Q \circ \pi} \quad (3.2.4)$$

Recorde que em  $TQ$  temos uma estrutura simpléctica dada pela forma simpléctica  $\omega_g = (g^\flat)^*\omega$ , onde  $\omega$  é a forma simpléctica canónica em  $T^*Q$ . A acção tangente  $\Phi^T$  é simpléctica, i.e., para cada  $g \in G$ ,  $\Phi_g^T : TQ \rightarrow TQ$  é um difeomorfismo simpléctico de  $(TQ, \omega_g)$ . De facto,  $(\Phi_g^T)^*\theta_g = \theta_g$ , onde  $\theta_g = (g^\flat)^*\theta$ , é o pull-back da forma de Liouville  $\theta$  em  $T^*Q$ , por  $g^\flat$ .

Consideremos de novo, para cada  $\xi \in \mathfrak{g}$ , a função  $\widehat{\mathbf{J}}_\xi : TQ \rightarrow \mathbb{R}$ , e verifiquemos que o respectivo campo hamiltoniano  $X_{\widehat{\mathbf{J}}_\xi} \in \mathfrak{X}(TQ)$ , é exactamente o gerador infinitesimal  $\xi_{TQ}$  da acção tangente. De facto:

$$\begin{aligned} (i_{\xi_{TQ}}\theta_g)(v_q) &= \theta_g(v_q)(\xi_{TQ}(v_q)) \\ &= g_q(v_q, T\pi\xi_{TQ}(v_q)) \\ &= g_q(\xi_Q(q)) \\ &= \widehat{\mathbf{J}}_\xi(v_q) \end{aligned}$$

isto é:

$$\widehat{\mathbf{J}}_\xi = i_{\xi_{TQ}}\theta_g \quad (3.2.5)$$

Por outro lado:

$$\begin{aligned} (\Phi_g^T)^*\theta_g = \theta_g &\Rightarrow \mathcal{L}_{\xi_{TQ}}\theta_g = 0 \\ &\Rightarrow di_{\xi_{TQ}}\theta_g + i_{\xi_{TQ}}d\theta_g = 0 \\ &\Rightarrow d\widehat{\mathbf{J}}_\xi = i_{\xi_{TQ}}\omega_g \quad \text{por (3.2.5)} \\ &\Rightarrow i_{X_{\widehat{\mathbf{J}}_\xi}}\omega_g = i_{\xi_{TQ}}\omega_g \\ &\Rightarrow \boxed{X_{\widehat{\mathbf{J}}_\xi} = \xi_{TQ}} \quad \text{porque } \omega_g \text{ é não degenerada} \end{aligned} \quad (3.2.6)$$

Após estas observações estamos aptos a enunciar e demonstrar o seguinte teorema fundamental:

♣ **Teorema 3.6** “**Teorema de E. Noether**” ... *Seja  $(Q, K, V, G)$  um sistema mecânico com simetria, e  $\mathbf{J} : TQ \rightarrow \mathfrak{g}$  a respectiva aplicação momento.*

*Então  $\mathbf{J}$  é integral primeiro do campo  $X_E$ , isto é,  $\mathbf{J}$  é constante ao longo das curvas integrais do campo hamiltoniano  $X_E$ , onde  $E = K + V \circ \pi : TQ \rightarrow \mathbb{R}$  é a energia total do sistema.*

**Dem.:** Sabemos que  $\Phi_g^T$  deixa  $E$  invariante:  $E \circ \Phi_g^T = E, \forall g \in G$ , por definição da acção de simetria. Em particular, para cada  $\xi \in \mathfrak{g}$ , temos que:

$$E(\Phi_{\exp t\xi}^T(v_q)) = E(v_q), \quad \forall v_q \in TQ$$

Derivando esta expressão em ordem a  $t$ , para  $t = 0$ , obtemos:

$$dE_{v_q} \xi_{TQ}(v_q) = 0 \quad \Rightarrow \quad \omega_g(X_E(v_q), \xi_{TQ}(v_q)) = 0 \quad \Rightarrow \quad \{E, \hat{\mathbf{J}}_\xi\} = 0$$

por definição de  $X_E$  e do parêntesis de Poisson, CQD.

### 3.2.1 O Problema de Kepler

Vamos considerar o seguinte um sistema mecânico com simetria:

$$\boxed{(Q, K, V, G) = (\mathbb{R}^2 - \{\mathbf{0}\}, K = \frac{1}{2}g, V(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|}, G = S = (2) \cong SS^1)} \quad (3.2.7)$$

onde  $g$  é a métrica euclideana usual em  $\mathbb{R}^2$ . Este sistema descreve o movimento de uma partícula de massa unitária, que se move no plano, sob a acção de um campo de forças central Newtoniano:

$$\mathbf{F}(\mathbf{x}) = -\text{grad } V(\mathbf{x}) = -\frac{1}{\|\mathbf{x}\|^2} \frac{\partial}{\partial \|\mathbf{x}\|} = -\frac{1}{r^2} \frac{\partial}{\partial r}$$

$\mathbf{x}$  é o vector de posição da partícula, e  $r = \|\mathbf{x}\|$ . É natural efectuar os cálculos em coordenadas polares  $r, \theta$  em  $Q = \mathbb{R}_+ \times SS^1$ . O grupo de simetria  $G = SO(2)$  actua em  $Q$  por rotações positivas, isto é, se  $R_\varphi$  é a rotação de ângulo  $\varphi$ , no sentido directo, a acção de simetria é  $\Phi : SO(2) \times Q \rightarrow Q$ , onde:

$$\Phi(R_\varphi, (r, \theta)) \stackrel{\text{def}}{=} (r, \theta + \varphi)$$

Claramente que  $\Phi$  é uma acção de simetria do sistema dado. Para calcular a respectiva aplicação momento, identificamos  $\mathfrak{so}(2) \cong T_1 SO(2)$  com  $i\mathbb{R} \cong \mathbb{R}$ , de tal forma que  $\exp(it\xi) = R_{t\xi} \in SO(2)$ . Portanto:

$$\begin{aligned} \xi_Q(q) &= \left. \frac{d}{dt} \right|_{t=0} \Phi(R_{t\xi}, (r, \theta)) \\ &= \left. \frac{d}{dt} \right|_{t=0} (r, \theta + t\xi) \\ &= \xi \frac{\partial}{\partial \theta} \Big|_q \quad \forall q = (r, \theta) \in Q \end{aligned} \quad (3.2.8)$$

A métrica  $g$  escreve-se em coordenadas polares na forma:

$$g = \dot{\mathbf{r}}^2 + r^2 \dot{\theta}^2$$

e se  $v_q = \dot{\mathbf{r}} \frac{\partial}{\partial \mathbf{r}} + \dot{\theta} \frac{\partial}{\partial \theta} \in \mathbf{T}_q \mathbf{Q}$ , então:

$$\begin{aligned} \mathbf{J}(v_q)\xi &= g_q(v_q, \xi_Q(q)) \\ &= g_q\left(\dot{\mathbf{r}} \frac{\partial}{\partial \mathbf{r}} + \dot{\theta} \frac{\partial}{\partial \theta}, \xi \frac{\partial}{\partial \theta}\right) \\ &= \xi r^2 \dot{\theta} \end{aligned} \tag{3.2.9}$$

Finalmente, identificando  $\mathbb{R} \cong \mathbb{R}^*$ , obtemos a aplicação momento  $\mathbf{J}$ , que não é mais do que o momento angular usual, expresso em coordenadas polares,  $\mathbf{J} : TQ \rightarrow \mathbb{R}$ :

$$\boxed{\mathbf{J}(r, \theta, \dot{\mathbf{r}}, \dot{\theta}) = r^2 \dot{\theta}} \tag{3.2.10}$$

### 3.2.2 Movimento livre de um sólido com um ponto fixo

Consideremos agora um sólido, constituído por pelo menos 3 pontos não colineares, que se move livremente em  $\mathbb{R}^3$  (ausência de forças externas), com um ponto fixo que supomos ser a origem  $\mathbf{0} \in \mathbb{R}^3$ .

O movimento deste sólido pode ser descrito da seguinte forma: consideramos um referencial fixo  $\mathcal{R}_f$  em  $\mathbb{R}^3$ , e um outro referencial  $\mathcal{R}_m$ , com a mesma origem  $\mathbf{0}$ , rigidamente ligado ao sólido, a que chamamos referencial móvel.

Se  $\mathbf{x}(t, \mathbf{a}) \in (\mathbb{R}^3, \mathcal{R}_f)$  representa a posição no instante  $t$ , relativamente ao referencial fixo  $\mathcal{R}_f$ , do ponto do sólido que no instante  $t = 0$  estava em  $\mathbf{a} \in (\mathbb{R}^3, \mathcal{R}_m)$ , então:

$$\boxed{\mathbf{x}(t, \mathbf{a}) = g(t) \mathbf{a}} \tag{3.2.11}$$

onde  $g(t) : (\mathbb{R}^3, \mathcal{R}_m) \rightarrow (\mathbb{R}^3, \mathcal{R}_f)$  é uma isometria linear positiva, isto é,  $g(t) \in SO(3)$ , com  $g(0) = \text{Id}$ .

As coordenadas de um ponto de  $\mathbb{R}^3$  relativamente ao referencial fixo  $\mathcal{R}_f$  dizem-se coordenadas espaciais, enquanto que as coordenadas de um ponto de  $\mathbb{R}^3$  relativamente ao referencial móvel  $\mathcal{R}_m$  dizem-se coordenadas do sólido.

# Bibliography

- [1] R. Abraham and J.E. Marsden, “*Foundations of Mechanics*”. 2nd. edition, Benjamin/Cummings Publishing Company 1978.
- [2] V.I. Arnold, “*Mathematical Methods of Classical Mechanics*”. 2nd. edition, GTM 60, Springer-Verlag 1981.
- [3] R. Courant and D. Hilbert, “*Methods of Mathematical Physics II*”. Interscience Publishers, 1962.
- [4] L. Elsgolts, “*Differential Equations and the Calculus of Variations*”. MIR Publishers, Moscow, 1977.
- [5] M. Giaquinta and S. Hidebrandt, “*Calculus of Variations I, II*”. Springer-Verlag 1996.
- [6] I.M. Gelfand and S.V. Fomin, “*Calculus of Variations*”. Dover Publications, Inc. 2000.
- [7] V. Guillemin and S. Sternberg, “*Geometric Asymptotics*”. AMS Mathematical Surveys 1977 (accessível em versão electrónica na página web da AMS).
- [8] R. Hermann “*Differential Geometry and the Calculus of Variations*”. Academic Press, 1968.
- [9] M. Kline and I.W. Kay, “*Electromagnetic Theory and Geometrical Optics*”. Interscience Publishers, 1964.
- [10] M.L. Krasnov, G.I. Makarenko and A.I. Kiseliöv, “*Calculo Variacional (exemplos y problemas)*”. Editorial MIR, Moscovo, 1976.
- [11] R.K. Luneburg, “*Mathematical theory of Optics*”. University of California Press, 1966.
- [12] H. Rund “*The Hamilton-Jacobi Theory in the Calculus of Variations, Its role in Mathematics and Physics*”. Robert Krieger Publishing Company, 1973.

- [13] M.I. Zelikin, “*Control Theory and Optimization I*”. Encyclopaedia of Mathematical Sciences, volume 86, Springer-Verlag, 2000.