Exercise: For all differential forms α and β, and any pair of real numbers a and b, show that

$$\mathrm{d}(a\alpha + b\beta) = a\mathrm{d}(\alpha) + b\mathrm{d}(\beta).$$

The property (ED3) extends to arbitrary differential forms

$$\mathrm{d}^2\alpha = \mathrm{d}(\mathrm{d}\alpha) = 0, \tag{16.5}$$

for, applying the operator d to (16.2) and using (ED3) gives

$$\mathrm{d}(\mathrm{d}\alpha) = \big(\mathrm{d}^2\alpha_{i_1\ldots i_r}\big) \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_r} = 0.$$

Example 16.1 Let $x = x^1$, $y = x^2$, $z = x^3$ be coordinates on the three-dimensional manifold $M = \mathbb{R}^3$. The exterior derivative of any 0-form $\alpha = f$ is

$$\mathrm{d}f = f_{,i}\,\mathrm{d}x^i = \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y + \frac{\partial f}{\partial z}\mathrm{d}z.$$

The three components are commonly known as the *gradient* of the scalar field f.

If $\omega = w_i\,\mathrm{d}x^i = A\,\mathrm{d}x + B\,\mathrm{d}y + C\mathrm{d}z$ is a differential 1-form then

$$\mathrm{d}\omega = \Big(\frac{\partial C}{\partial y} - \frac{\partial B}{\partial z}\Big)\mathrm{d}y \wedge \mathrm{d}z + \Big(\frac{\partial A}{\partial z} - \frac{\partial C}{\partial x}\Big)\mathrm{d}z \wedge \mathrm{d}x + \Big(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}\Big)\mathrm{d}x \wedge \mathrm{d}y.$$

The components of the exterior derivative are traditionally written as components of a vector field, known as the *curl* of the three-component vector field (A, B, C). Notice, however, that the *tensor* components of $\mathrm{d}\omega = -w_{[i,j]}\,\mathrm{d}x^i \wedge \mathrm{d}x^j$ are half the curl components,

$$(\mathrm{d}\omega)_{ij} = -w_{[i,j]} = \tfrac{1}{2}\big(w_{j,i} - w_{i,j}\big).$$

If $\alpha = \alpha_{ij}\,\mathrm{d}x^i \wedge \mathrm{d}x^j = P\,\mathrm{d}y \wedge \mathrm{d}z + Q\,\mathrm{d}z \wedge \mathrm{d}x + R\,\mathrm{d}x \wedge \mathrm{d}y$ is a 2-form then

$$\mathrm{d}\alpha = \Big(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\Big)\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z.$$

The single component of this 3-form is known as the *divergence* of the three-component vector field (P, Q, R). Equation (16.5) applied to the 0-form f and 1-form ω gives the following classical results:

$$\mathrm{d}^2 f = 0 \implies \text{curl grad} = 0,$$
$$\mathrm{d}^2\omega = 0 \implies \text{div curl} = 0.$$

Exercise: If $\alpha = \alpha_{ij}\,\mathrm{d}x^i \wedge \mathrm{d}x^j$ is a 2-form on a manifold M, show that

$$(\mathrm{d}\alpha)_{ijk} = \tfrac{1}{3}\big(\alpha_{ij,k} + \alpha_{jk,i} + \alpha_{ki,j}\big). \tag{16.6}$$

More generally, lumping together the permutations of the first r indices in Eq. (16.4) we obtain the following formula for the tensor components of the exterior derivative of an r-form α:

$$(\mathrm{d}\alpha)_{i_1\ldots i_{r+1}} = \frac{(-1)^r}{r+1} \sum_{\text{cyclic } \pi} (-1)^\pi \alpha_{i_{\pi(1)}\ldots i_{\pi(r)}, i_{\pi(r+1)}}. \tag{16.7}$$

Problems

Problem 16.1 Let $x^1 = x$, $x^2 = y$, $x^3 = z$ be coordinates on the manifold $\mathbb{R}^3$. Write out the components α_{ij} and $(\mathrm{d}\alpha)_{ijk}$, etc. for each of the following 2-forms:

$$\alpha = \mathrm{d}y \wedge \mathrm{d}z + \mathrm{d}x \wedge \mathrm{d}y,$$
$$\beta = x\,\mathrm{d}z \wedge \mathrm{d}y + y\,\mathrm{d}x \wedge \mathrm{d}z + z\,\mathrm{d}y \wedge \mathrm{d}x,$$
$$\gamma = \mathrm{d}(r^2(x\,\mathrm{d}x + y\,\mathrm{d}y + z\,\mathrm{d}z)) \quad \text{where} \quad r^2 = x^2 + y^2 + z^2.$$

Problem 16.2 On the manifold $\mathbb{R}^n$ compute the exterior derivative d of the differential form

$$\alpha = \sum_{i=1}^{n} (-1)^{i-1} x^i \,\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{i-1} \wedge \mathrm{d}x^{i+1} \wedge \cdots \wedge \mathrm{d}x^n.$$

Do the same for $\beta = r^{-n}\alpha$ where $r^2 = (x^1)^2 + \cdots + (x^n)^2$.

Problem 16.3 Show that the right-hand side of Eq. (16.6) transforms as a tensor field of type (0, 3). Generalize this result to the right-hand side of Eq. (16.7), to show that this equation could be used as a local definition of exterior derivative independent of the choice of coordinate system.

16.2 Properties of exterior derivative

If $\varphi : M \to N$ is a smooth map between two differentiable manifolds M and N, we define the induced map $\varphi^* : \Lambda_r(N) \to \Lambda_r(M)$ in a similar way to the pullback map, Eq. (15.18):

$$(\varphi^*\alpha)_p\big((X_1)_p,\ (X_2)_p, \ldots,\ (X_r)_p\big) = \alpha_{\varphi(p)}\big(\varphi_*(X_1)_p,\ \varphi_*(X_2)_p, \ldots,\ \varphi_*(X_r)_p\big).$$

As for covector fields, this map is well-defined on all differential r-forms, $\varphi^*\alpha$. The pullback of a 0-form $f \in \mathcal{F}(N) = \Lambda_0(N)$ is defined by $\varphi^* f = f \circ \varphi$, and it preserves wedge products

$$\varphi^*(\alpha \wedge \beta) = \varphi^*\alpha \wedge \varphi^*\beta,$$

which follows immediately from the definition $\alpha \wedge \beta = \mathcal{A}(\alpha \otimes \beta)$.

Exercise: Show that the composition of two maps φ and ψ results in a reverse composition of pullbacks, as in Eq. (15.19), $(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$.

Theorem 16.2 *For any differential form $\alpha \in \Lambda(N)$, the induced map φ^* commutes with the exterior derivative,*

$$\mathrm{d}\varphi^*\alpha = \varphi^*\,\mathrm{d}\alpha.$$

Proof: For a 0-form, $\alpha = f : N \to \mathbb{R}$, at any point $p \in M$ and any tangent vector X_p

$$\begin{aligned}\langle(\varphi^*\mathrm{d}f)_p, X_p\rangle &= \langle(\mathrm{d}f)_{\varphi(p)}, \varphi_* X_p\rangle \\ &= (\varphi_* X_p) f \\ &= X_p(f \circ \varphi) \\ &= \langle\big(\mathrm{d}(f \circ \varphi)\big)_p, X_p\rangle.\end{aligned}$$

As this equation holds for all tangent vectors X_p, we have

$$\varphi^*\mathrm{d}f = \mathrm{d}(f \circ \varphi) = \mathrm{d}(\varphi^* f).$$

For a general r-form, it is only necessary to prove the result in any local coordinate chart $(U; x^i)$. If $\alpha = \alpha_{i_1 \ldots i_r}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_r}$, then

$$\begin{aligned}\varphi^* \mathrm{d}\alpha &= \varphi^*(\mathrm{d}\alpha_{i_1 \ldots i_r} \wedge \mathrm{d}x^{i_1} \wedge \ldots \mathrm{d}x^{i_r}) \\ &= \mathrm{d}(\varphi^* \alpha_{i_1 \ldots i_r}) \wedge \mathrm{d}(\varphi^* x^{i_1}) \wedge \cdots \wedge \mathrm{d}(\varphi^* x^{i_r}) \\ &= \mathrm{d}(\varphi^* \alpha). \end{aligned}$$ ■

Applying the definition (15.33) of Lie derivative to the tensor field α and using $\widetilde{\sigma_t} = \left(\sigma_{-t}\right)^*$, where σ_t is a local one-parameter group generating a vector field X, it follows from Theorem 16.2 that the exterior derivative and Lie derivative commute,

$$\mathcal{L}_X\, \mathrm{d}\alpha = \mathrm{d}\mathcal{L}_X \alpha. \tag{16.8}$$

For any vector field X define the **interior product** $\mathrm{i}_X : \Lambda_r(M) \to \Lambda_{(r-1)}(M)$ as in Section 8.4,

$$\mathrm{i}_X \alpha = r C^1_1(X \otimes \alpha), \tag{16.9}$$

or equivalently, for arbitrary vector fields $X_1, X_2, \ldots, X_r$

$$(\mathrm{i}_{X_1} \alpha)(X_2, \ldots, X_r) = r\alpha(X_1, X_2, \ldots, X_r). \tag{16.10}$$

By Eq. (8.20) i_X is an antiderivation – for any differential r-form α and arbitrary differential form β

$$\mathrm{i}_X(\alpha \wedge \beta) = (\mathrm{i}_X \alpha) \wedge \beta + (-1)^r \alpha \wedge (\mathrm{i}_X \beta). \tag{16.11}$$

Exercise: Show that for any pair of vector fields X and Y, $\mathrm{i}_X \circ \mathrm{i}_Y = -\mathrm{i}_Y \circ \mathrm{i}_X$.

Theorem 16.3 (Cartan) *If X and Y are smooth vector fields on a differentiable manifold M and ω is a differential 1-form then*

$$\mathrm{i}_{[X,Y]} = \mathcal{L}_X \circ \mathrm{i}_Y - \mathrm{i}_Y \circ \mathcal{L}_X, \tag{16.12}$$

$$\mathcal{L}_X = \mathrm{i}_X \circ \mathrm{d} + \mathrm{d} \circ \mathrm{i}_X, \tag{16.13}$$

$$\mathrm{d}\omega(X, Y) = \tfrac{1}{2}\big(X(\langle Y, \omega\rangle) - Y(\langle X, \omega\rangle) - \langle [X, Y], \omega\rangle\big). \tag{16.14}$$

Proof: The first identity follows essentially from the fact that the Lie derivative $\mathcal{L}_X$ commutes with contraction operators, $\mathcal{L}_X C^i_j = C^i_j \mathcal{L}_X$ (see Problem 15.24). Thus for an arbitrary r-form α, using the Leibnitz rule (15.35) gives

$$\begin{aligned}\mathcal{L}_X(\mathrm{i}_Y \alpha) &= r C^1_1 \mathcal{L}_X(Y \otimes \alpha) \\ &= r C^1_1\big[(\mathcal{L}_X Y) \otimes \alpha + Y \otimes \mathcal{L}_X \alpha\big] \\ &= \mathrm{i}_{[X,Y]} \alpha + \mathrm{i}_Y(\mathcal{L}_X \alpha)\end{aligned}$$

as required.

To show (16.13) set K_X to be the operator $K_X = \mathrm{i}_X \circ \mathrm{d} + \mathrm{d} \circ \mathrm{i}_X : \Lambda_r(M) \to \Lambda_r(M)$. Using the fact that both i_X and d are antiderivations, Eqs. (16.11) and (16.3), it is straightforward to show that K_X is a derivation,

$$K_X(\alpha \wedge \beta) = K_X \alpha \wedge \beta + \alpha \wedge K_X \beta$$

for all differential forms α and β. From $\mathrm{d}^2 = 0$ the operator K_X commutes with d,

$$K_X \circ \mathrm{d} = \mathrm{i}_X \circ \mathrm{d}^2 + \mathrm{d} \circ \mathrm{i}_X \circ \mathrm{d} = \mathrm{d} \circ \mathrm{i}_X \circ \mathrm{d} = \mathrm{d} \circ K_X.$$

If α is a 0-form $\alpha = f$ then $\mathrm{i}_X f = 0$ by definition, and

$$K_X f = \mathrm{i}_X(\mathrm{d}f) + \mathrm{d}\mathrm{i}_X f = \langle \mathrm{d}f, X\rangle = Xf = \mathcal{L}_X f.$$

Hence, since K_X commutes both with d and $\mathcal{L}_X$,

$$K_X\, \mathrm{d}f = \mathrm{d}K_X f = \mathrm{d}(\mathcal{L}_X f) = \mathcal{L}_X\, \mathrm{d}f.$$

On applying the derivation property we obtain $K_X(g\, \mathrm{d}f) = \mathcal{L}_X(g\, \mathrm{d}f)$ and the required identity holds for any 1-form ω, as it can be expressed locally in a coordinate chart at any point as $\omega = w_i\, \mathrm{d}x^i$. The argument may be generalized to higher order r-forms to show that the operators $\mathcal{L}_X$ and K_X are identical on all of $\Lambda_r(M)$.

The final identity (16.14) is proved on applying (16.13) to a 1-form ω,

$$\langle Y, \mathcal{L}_X\omega\rangle = \langle Y, \mathrm{i}_X(\mathrm{d}\omega) + \mathrm{d}(\mathrm{i}_X\omega)\rangle$$

and using the Leibnitz rule for the Lie derivative,

$$\mathcal{L}_X\big(\langle Y, \omega\rangle\big) - \langle \mathcal{L}_X Y, \omega\rangle = \mathrm{i}_X\, \mathrm{d}\omega(Y) + Y(\mathrm{i}_X\omega).$$

Setting $r = 1$ and $\alpha = \omega$ in Eq. (16.10),

$$X\big(\langle Y, \omega\rangle\big) - \langle \mathcal{L}_X Y, \omega\rangle = 2\, \mathrm{d}\omega(X, Y) + Y\big(\langle X, \omega\rangle\big)$$

from which (16.14) is immediate. ■

If α is an r-form on M, a formula for $\alpha(X_1, X_2, \ldots, X_{r+1})$ that generalizes Eq. (16.14) is left to the reader (see Problem 16.5).

Problems

Problem 16.4 Let $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$ be the map

$$(x, y) \to (u, v, w) \quad \text{where} \quad u = \sin(xy),\ v = x + y,\ w = 2.$$

For the 1-form $\omega = w_1\mathrm{d}u + w_2\mathrm{d}v + w_3\mathrm{d}w$ on $\mathbb{R}^3$ evaluate $\varphi^*\omega$. For any function $f : \mathbb{R}^3 \to \mathbb{R}$ verify Theorem 16.2, that $\mathrm{d}(\varphi^* f) = \varphi^*\mathrm{d}f$.

Problem 16.5 If α is an r-form on a differentiable manifold M, show that for any vector fields $X_1, X_2, \ldots X_{r+1}$

$$\mathrm{d}\alpha(X_1, X_2, \ldots, X_{r+1}) = \frac{1}{r+1}\Big[\sum_{i=1}^{r+1}(-1)^{i+1} X_i\alpha(X_1, X_2, \ldots, \hat{X}_i, \ldots, X_{r+1}) + \sum_{i=1}^{r}\sum_{j=i+1}^{r+1}(-1)^{i+j}\alpha([X_i, X_j], \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_{r+1})\Big]$$

where $\hat{X}_i$ signifies that the argument X_i is to be omitted. The case $r = 0$ simply asserts that $\mathrm{d}f(X) = Xf$, while Eq. (16.14) is the case $r = 1$. Proceed by induction, assuming the identity is true for all $(r-1)$-forms, and use the fact that any r-form can be written locally as a sum of tensors of the type $\omega \wedge \beta$ where ω is a 1-form and β an r-form.

Problem 16.6 Show that the Laplacian operator on $\mathbb{R}^3$ may be defined by

$$\mathrm{d} * \mathrm{d}\phi = \nabla^2\phi \, \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = \Big(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}\Big)\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$$

where $*$ is the Hodge star operator of Section 8.6.

Use this to express the Laplacian operator in spherical polar coordinates (r, θ, ϕ).

16.3 Frobenius theorem: dual form

Let D^k be a k-dimensional distribution on a manifold M, assigning a k-dimensional subspace $D^k(p)$ of the tangent space at each point $p \in M$. Its **annihilator subspace** $(D^k)^\perp(p)$ (see Problem 3.16) consists of the set of covectors at p that vanish on $D^k(p)$,

$$(D^k)^\perp(p) = \{\omega_p \,|\, \langle \omega_p, X_p\rangle = 0 \text{ for all } X_p \in D^k(p)\}.$$

Since the distribution D^k is required to be C^∞, it follows from Theorem 3.7 that every point p has a neighbourhood U and a basis e_i of smooth vector fields on U, such that $e_{r+1}, \ldots, e_n$ span $D^k(q)$ at every point $q \in U$, where $r = n - k$. The dual basis of 1-forms ω^i defined by $\langle \omega^i, e_j\rangle = \delta^i_j$ has the property that the first r 1-forms $\omega^1, \omega^2, \ldots, \omega^r$ are linearly independent and span the annihilator subspace $(D^k)^\perp(q)$ at each $q \in U$.

The annihilator property is reciprocal: given r linearly independent 1-forms ω^a $(a = 1, \ldots, r)$ on an open subset U of M, they span the annihilator subspace $(D^k)^\perp$ of the $k = (n - r)$-dimensional distribution

$$D^k = \{X \in \mathcal{T}(U) \,|\, \langle \omega^a, X\rangle = 0\}.$$

As shown at the end of Section 8.3, the simple differential k-form

$$\Omega = \omega^1 \wedge \omega^2 \wedge \cdots \wedge \omega^k$$

is uniquely defined up to a scalar field factor by the subspace $(D^k)^\perp$, and has the property that a 1-form ω belongs to $(D^k)^\perp$ if and only if $\omega \wedge \Omega = 0$.

Suppose the distribution D^k is involutive, so that $X, Y \in D^k \Rightarrow [X, Y] \in D^k$. From Eq. (16.14)

$$\mathrm{d}\omega^a(X, Y) = \tfrac{1}{2}\big(X(\langle Y, \omega^a\rangle) - Y(\langle X, \omega^a\rangle) - \langle [X, Y], \omega^a\rangle\big) = 0$$

for any pair of vectors $X, Y \in D^k$. Conversely, if all ω^a and $\mathrm{d}\omega^a$ vanish when restricted to the distribution D^k, then $\langle [X, Y], \omega^a\rangle = 0$ for all $X, Y \in D^k$. Thus, a necessary and sufficient condition for a distribution D^k to be involutive is that for all $\omega^a \in (D^k)^\perp$ the exterior derivative $\mathrm{d}\omega^a$ vanishes on D^k.

Let $A^a_{ij} = -A^a_{ji}$ be scalar fields such that $\mathrm{d}\omega^a = A^a_{ij}\omega^i \wedge \omega^j$. If $\mathrm{d}\omega^a(e_\alpha, e_\beta) = 0$ for all $\alpha, \beta = r + 1, \ldots, n$, then $A^a_{\alpha\beta} = 0$ and

$$\mathrm{d}\omega^a = A^a_{bc}\omega^b \wedge \omega^c + A^a_{b\beta}\omega^b \wedge \omega^\beta + A^a_{\alpha c}\omega^\alpha \wedge \omega^c.$$

Thus, D^k is involutive if and only if for the 1-forms $\mathrm{d}\omega^a$ there exist 1-forms θ^a_b such that

$$\mathrm{d}\omega^a = \theta^a_b \wedge \omega^b.$$

On the other hand, the Frobenius theorem 15.4 asserts that D^k is involutive if and only if there exist local coordinates $(U; x^i)$ at any point p such that $e_\alpha = B^\beta_\alpha \partial_{x^\beta}$ for an invertible matrix of scalar fields $\left[B^\beta_\alpha\right]$ on U. In these coordinates, set $\omega^a = A^a_b \, \mathrm{d}x^b + W^a_\alpha \, \mathrm{d}x^\alpha$, and using $\langle \omega^a, e_\alpha \rangle = 0$, we have $W^a_\alpha = 0$. Hence an alternative necessary and sufficient condition for D^k to be involutive is the existence of coordinates $(U; x^i)$ such that

$$\omega^a = A^a_b \, \mathrm{d}x^b.$$

Theorem 16.4 *Let ω^a $(a = 1, \ldots, r)$ be a set of 1-forms on an open set U, linearly independent at every point $p \in U$. The following statements are all equivalent:*

(i) There exist local coordinates $(U; x^i)$ at every point $p \in U$ such that $\omega^a = A^a_b \, \mathrm{d}x^b$.
(ii) There exist 1-forms θ^a_b such that $\mathrm{d}\omega^a = \theta^a_b \wedge \omega^b$.
(iii) $\mathrm{d}\omega^a \wedge \Omega = 0$ where $\Omega = \omega^1 \wedge \omega^2 \wedge \cdots \wedge \omega^r$.
(iv) $\mathrm{d}\Omega \wedge \omega^a = 0$.
(v) There exists a 1-form θ such that $\mathrm{d}\Omega = \theta \wedge \Omega$.

Proof: We have seen by the above remarks that (i) $\Leftrightarrow$ (ii) as both statements are equivalent to the statement that the distribution D^k that annihilates all ω^a is involutive. Condition (ii) $\Leftrightarrow$ (iii) since $\omega^a \wedge \Omega = 0$, while the converse follows on setting $\mathrm{d}\omega^a = \theta^a_b \wedge \omega^b + A^a_{\alpha\beta} \omega^\alpha \wedge \omega^\beta$, where ω^i $(i = 1, \ldots, n)$ is any local basis of 1-forms completing the ω^a.

The implication (iii) $\Leftrightarrow$ (iv) follows at once from Eq. (16.3), and (v) $\Rightarrow$ (iv) since

$$\mathrm{d}\Omega = \theta \wedge \Omega \implies \mathrm{d}\Omega \wedge \omega^a = \theta \wedge \Omega \wedge \omega^a = 0.$$

Finally, (ii) $\Rightarrow$ (v), for if $\mathrm{d}\omega^a = \theta^a_b \wedge \omega^b$ then

$$\begin{aligned} \mathrm{d}\Omega &= \mathrm{d}\omega^1 \wedge \omega^2 \wedge \cdots \wedge \omega^r - \omega^1 \wedge \mathrm{d}\omega^2 \wedge \cdots \wedge \omega^r + \ldots \\ &= \theta^1_1 \wedge \omega^1 \wedge \omega^2 \wedge \cdots \wedge \omega^r - \omega^1 \wedge \theta^2_2 \wedge \omega^2 \wedge \cdots \wedge \omega^r + \ldots \\ &= (\theta^1_1 + \theta^2_2 + \cdots + \theta^r_r) \wedge \omega^1 \wedge \omega^2 \wedge \cdots \wedge \omega^r \\ &= \theta \wedge \Omega \end{aligned}$$

where $\theta = \theta^a_a$. Hence (iv) $\Rightarrow$ (iii)$\Rightarrow$ (ii) $\Rightarrow$ (v) and the proof is completed. ■

A system of linearly independent 1-forms $\omega^1, \ldots, \omega^r$ on an open set U, satisfying any of the conditions (i)–(v) of this theorem is said to be completely integrable . The equations defining the distribution D^k $(k = n - r)$ that annihilates these ω^a is given by the equations $\langle \omega^a, X \rangle = 0$, often written as a **Pfaffian system of equations**

$$\omega^a = 0 \quad (a = 1, \ldots, r).$$

Condition (i) says that locally there exist r functions $g^a(x^1, \ldots, x^n)$ on U such that

$$\omega^a = f^a_b \, \mathrm{d}g^b$$

where the functions f^a_b form a non-singular $r \times r$ matrix at every point of U. The functions g^a are known as a **first integral of the system**. The r-dimensional submanifolds $(N_\mathbf{c}, \psi_\mathbf{c})$ defined by $g^a(x^1, \ldots, x^n) = c^a = \text{const.}$ have the property

$$\psi^*_\mathbf{c} \omega^a = f^a_b \circ \psi_\mathbf{c} \, \mathrm{d}c^b = 0,$$

and are known as **integral submanifolds** of the system.

Example 16.2 Consider a single Pfaffian equation in three dimensions,

$$\omega = P(x, y, z)\,\mathrm{d}x + Q(x, y, z)\,\mathrm{d}y + R(x, y, z)\,\mathrm{d}z = 0.$$

If $\omega = f\,\mathrm{d}g$ where $f(0, 0, 0) \neq 0$, the function $f(x, y, z)$ is said to be an *integrating factor*. It is immediate then that

$$\mathrm{d}\omega = \mathrm{d}f \wedge \mathrm{d}g = \mathrm{d}f \wedge \frac{1}{f}\omega = \theta \wedge \omega$$

where $\theta = \mathrm{d}(\ln f)$. This is equivalent to conditions (ii) and (v) of Theorem 16.4. Conditions (iii) and (iv) are identical since $\Omega = \omega$, and follow at once from

$$\mathrm{d}\omega \wedge \omega = \theta \wedge \omega \wedge \omega = 0,$$

which reduces to Euler's famous integrability condition for the existence of an integrating factor,

$$P\Big(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\Big) + Q\Big(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\Big) + R\Big(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\Big) = 0.$$

For example, if $\omega = \mathrm{d}x + z\,\mathrm{d}y + \mathrm{d}z$ there is no integrating factor, $\omega = f\,\mathrm{d}g$, since

$$\mathrm{d}\omega \wedge \omega = \mathrm{d}z \wedge \mathrm{d}y \wedge \mathrm{d}x = -\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \neq 0.$$

On the other hand, if $\omega = 2xz\,\mathrm{d}x + 2yz\,\mathrm{d}y + \mathrm{d}z$, then

$$\mathrm{d}\omega \wedge \omega = (2x\,\mathrm{d}z \wedge \mathrm{d}x + 2y\,\mathrm{d}z \wedge \mathrm{d}y) \wedge \omega = 4xyz(\mathrm{d}z \wedge \mathrm{d}x \wedge \mathrm{d}y + \mathrm{d}z \wedge \mathrm{d}y \wedge \mathrm{d}x) = 0.$$

It should therefore be possible locally to express ω in the form $f\,\mathrm{d}g$. The functions f and g are not unique, for if $G(g)$ is an arbitrary function then $\omega = F\,\mathrm{d}G$ where $F = f/(\mathrm{d}G/\mathrm{d}g)$. To find an integrating factor f we solve a system of three differential equations

$$f\frac{\partial g}{\partial x} = 2xz, \tag{a}$$

$$f\frac{\partial g}{\partial y} = 2yz, \tag{b}$$

$$f\frac{\partial g}{\partial z} = 1. \tag{c}$$

Eliminating f from (a) and (b) we have

$$\frac{1}{x}\frac{\partial g}{\partial x} = \frac{1}{y}\frac{\partial g}{\partial y},$$

which can be expressed as

$$\frac{\partial g}{\partial x^2} = \frac{\partial g}{\partial y^2}.$$

This equation has a general solution $g(z, u)$ where $u = x^2 + y^2$, and eliminating f from (b) and (c) results in

$$\frac{\partial g}{\partial z} = 2yz\frac{\partial g}{\partial y} \implies \frac{\partial g}{\partial \ln z} = \frac{\partial g}{\partial y^2} = \frac{\partial g}{\partial u}.$$

Hence $g = G(\ln z + u)$, and since it is possible to pick an arbitrary function G we can set $g = ze^{x^2+y^2}$. From (c) it follows that $f = e^{-x^2-y^2}$, and it is easy to check that

$$\omega = e^{-x^2-y^2} d\left(ze^{x^2+y^2}\right) = 2xz\,dx + 2yz\,dy + dz.$$

Problems

Problem 16.7 Let $\omega = yz\,dx + xz + z^2\,dz$. Show that the Pfaffian system $\omega = 0$ has integral surfaces $g = z^3e^{xy} = \text{const.}$, and express ω in the form $f\,dg$.

Problem 16.8 Given an $r \times r$ matrix of 1-forms Ω, show that the equation

$$d\mathsf{A} = \Omega\mathsf{A} - \mathsf{A}\Omega$$

is soluble for an $r \times r$ matrix of functions A only if

$$\Theta\mathsf{A} = \mathsf{A}\Theta$$

where $\Theta = d\Omega - \Omega \wedge \Omega$.

If the equation has a solution for arbitrary initial values $\mathsf{A} = \mathsf{A}_0$ at any point $p \in M$, show that there exists a 2-form α such that $\Theta = \alpha\mathsf{I}$ and $d\alpha = 0$.

16.4 Thermodynamics

Thermodynamics deals with the overall properties of systems such as a vessel of gas or mixture of gases, a block of ice, a magnetized iron bar, etc. While such systems may be impossibly complex at the microscopic level, their thermodynamic behaviour is governed by a very few number of variables. For example, the state of a simple gas is determined by two variables, its volume V and pressure p, while a mixture of gases also requires specification of the molar concentrations $n_1, n_2, \ldots$ representing the relative number of particles of each species of gas. An iron bar may need information from among variables such as its length ℓ, cross-section A, tensile strength f and Young's modulus Y, the magnetic field $\mathbf{H}$, magnetization μ, electric field $\mathbf{E}$ and conductivity σ. In any case, the number of variables needed for a thermodynamic description of the system is tiny compared to the 10^{24} or so variables required for a complete description of the microscopic state of the system (see Section 14.4).

The following treatment is similar to that given in [10]. Every thermodynamic system will be assumed to have a special class of states known as **equilibrium states**, forming an n-dimensional manifold K, and given locally by a set of **thermodynamic variables** $\mathbf{x} = (x^1, x^2, \ldots, x^n)$. The dimension n is called the number of **degrees of freedom** of the thermodynamic system. Physically, we think of an equilibrium state as one in which the system remains when all external forces are removed. For a perfect gas there are two degrees of freedom, usually set to be $x^1 = p$ and $x^2 = V$. The variable p is called an **internal** or **thermal** variable, characterized physically by the fact that no work is done on or by the system if we change p alone, leaving V unaltered. Variables such as volume V, a change in which results in work being done on the system, are called **external** or **deformation** variables.

A **quasi-static** or **reversible** process, resulting in a transition from one equilibrium state $\mathbf{x}_1$ to another $\mathbf{x}_2$, is a parametrized curve $\gamma : [t_1, t_2] \to K$ such that $\gamma(t_1) = \mathbf{x}_1$ and $\gamma(t_2) = \mathbf{x}_2$. Since the curve passes through a continuous succession of equilibrium states, it should be thought of as occurring *infinitely slowly*, and its parameter t is not to be identified with real time. For example, a gas in a cylinder with a piston attached will undergo a quasi-static transition if the piston is withdrawn so slowly that the effect on the gas is reversible. If the piston is withdrawn rapidly the action is irreversible, as non-equilibrium intermediate states arise in which the gas swirls and eddies, creating regions of non-uniform pressure and density throughout the container. The same can be said of the action of a 'stirrer' on a gas or liquid in an adiabatic container – you can never 'unstir' the milk or sugar added to a cup of tea. Irreversible transitions from one state of the system cannot be represented by parametrized curves in the manifold of equilibrium states K. Whether the transition be reversible or irreversible, we assume that there is always associated with it a well-defined quantity ΔW, known as the **work done by the system**. The work done **on the system** is defined to be the negative of this quantity, $-\Delta W$.

We will also think of thermodynamic systems as being confined to certain 'enclosures', to be thought of as closed regions of three-dimensional space. Most importantly, a system K is said to be in an **adiabatic enclosure** if equilibrium states can only be disturbed by doing work on the system through mechanical means (reversible or irreversible), such as the movement of a piston or the rotation of a stirrer. In all cases, transitions between states of a system in an adiabatic enclosure are called **adiabatic processes**.

The boundary of an adiabatic enclosure can be considered as being an *insulating wall* through which no 'heat transfer' is allowed; a precise meaning to the concept of *heat* will be given directly. A **diathermic wall** within an adiabatic enclosure is one that permits heat to be transferred across it without any work being done. Two systems K_A and K_B are said to be in **thermal contact** if both are enclosed in a common adiabatic enclosure, but are separated by a diathermic wall. The states $\mathbf{x}_A$ and $\mathbf{x}_B$ of the two systems are then said to be in **thermal equilibrium** with each other.

Zeroth law of thermodynamics: temperature *For every thermodynamic system K there exists a function $\tau : K \to R$ called **empirical temperature** such that two systems K_A and K_B are in equilibrium with each other if and only if $\tau_A(\mathbf{x}_A) = \tau_B(\mathbf{x}_B)$.*

This law serves as little more than a definition of empirical temperature, but the fact that a single function of state achieves the definition of equilibrium is significant. Any set of states $\{\mathbf{x} \mid \tau(\mathbf{x}) = \text{const.}\}$ is called an **isotherm** of a system K.

Example 16.3 For an **ideal gas** we find $\tau = pV$ is an empirical temperature, and the isotherms are curves $pV = \text{const}$. Any monotone function $\tau' = \varphi \circ \tau$ will also do as empirical temperature. A system of ideal gases in equilibrium with each other, (p_1, V_1), $(p_2, V_2), \ldots, (p_n, V_n)$ have common empirical temperature $\tau = T_g$, called the **absolute gas temperature**, given by

$$T_g = \frac{p_1 V_1}{n_1 R} = \frac{p_2 V_2}{n_2 R} = \cdots = \frac{p_n V_n}{n_n R}$$

where n_i are the relative molar quantities of the gases involved and R is the universal gas constant. While an arbitrary function φ may still be applied to the absolute gas temperature, the same function must be applied equally to all component gases. In this example it is possible to eliminate all pressures except one, and the total system can be described by a single thermal variable, p_1 say, and n external deformation variables $V_1, V_2, \ldots, V_n$.

This example illustrates a common assumption made about thermodynamic systems of n degrees of freedom, that it is possible to pick coordinates $(x^1, x^2, \ldots, x^n)$ in a local neighbourhood of any point in K such that the first $n-1$ coordinates are external variables and x^n is an internal variable. We call this the **thermal variable assumption**.

First law of thermodynamics: energy. *For every thermodynamic system K there is a function $U : K \to \mathbb{R}$ known as **internal energy** and a 1-form $\omega \in \Lambda_1(K)$ known as the **work form** such that the work done by the system in any reversible process $\gamma : [t_1, t_2] \to K$ is given by*

$$\Delta W = \int_\gamma \omega$$

(see Example 15.9 for the definition of integral of ω along the curve γ). In every reversible adiabatic process

$$\gamma^*(\omega + \mathrm{d}U) = 0.$$

From Example 15.9 the integral of $\omega + \mathrm{d}U$ along the curve γ vanishes, since

$$\int_\gamma \omega + \mathrm{d}U = \int_{t_1}^{t_2} \langle \gamma^*(\omega + \mathrm{d}U), \frac{\mathrm{d}}{\mathrm{d}t} \rangle \mathrm{d}t = 0.$$

Furthermore, since

$$\int_\gamma \mathrm{d}U = \int_{t_1}^{t_2} \frac{\mathrm{d}U}{\mathrm{d}t} \mathrm{d}t = U(t_2) - U(t_1) = \Delta U$$

the conservation law of energy holds for any reversible adiabatic process,

$$\Delta W + \Delta U = 0.$$

Thus the change of internal energy is equal to the work done on the system, $-\Delta W$. The work done in any reversible adiabatic transition from one equilibrium state of a system K to another is independent of the path. In particular, no work is done by the system in any cyclic adiabatic process, returning a system to its original state – commonly known as the *impossibility of a perpetual motion machine of the first kind*.

The **heat 1-form** is defined to be $\theta = \omega + \mathrm{d}U$, and we refer to $\Delta Q = \int_\gamma \theta$ as the **heat added** to the system in any reversible process γ. The conservation of energy in the form

$$\Delta Q = \Delta W + \Delta U$$

is often referred to in the literature as the first law of thermodynamics. Adiabatic transitions are those with $\Delta Q = 0$.

If the thermal variable assumption holds, then it is generally assumed that the work form is a linear expansion of the external variables alone,

$$\omega = \sum_{k=1}^{n-1} P_k(x^1, \ldots, x^n)\, \mathrm{d}x^k,$$

where the component function P_k is known as the k**th generalized force**. Since U is a thermal variable, it is always possible to choose the nth coordinate as $x^n = U$, in which case

$$\theta = \sum_{k=1}^{n-1} P_k(x^1, \ldots, x^n)\, \mathrm{d}x^k + \mathrm{d}U = \sum_{k=1}^{n-1} P_k(x^1, \ldots, x^n)\, \mathrm{d}x^k + \mathrm{d}x^n.$$

Second law of thermodynamics

Not every transition between equilibrium states is possible, even if conservation of energy holds. The second law of thermodynamics limits the possible transitions consistent with energy conservation, and has a number of equivalent formulations. For example, the version due to Clausis asserts that no machine can perform work, or mechanical energy, while at the same time having no other effect than to lower the temperature of a thermodynamic system. Such a machine is sometimes referred to as a *perpetual motion machine of the second kind* – if it were possible one could draw on the essentially infinite heat reservoir of the oceans to perform an unlimited amount of mechanical work.

An equivalent version is Kelvin's principle: no cyclic quasi-static thermodynamic process permits the conversion of heat *entirely* into mechanical energy. By this is meant that no quasi-static thermodynamic cycle γ exists, the first half of which consists of a quasi-static process γ_1 purely of heat transfer in which no work is done, $\gamma_1^*\omega = 0$, while the second half γ_2 is adiabatic and consists purely of mechanical work, $\gamma_2^*\theta = 0$. Since U is a function of state it follows, on separating the cycle into its two parts, that

$$0 = \Delta U = \oint_\gamma \mathrm{d}U = \Delta_1 U + \Delta_2 U = \Delta_1 Q - \Delta_2 W.$$

Thus in any such cycle an amount of heat would be converted entirely into its mechanical equivalent of work.

Consider a quasi-static process γ_1 taking an equilibrium state $\mathbf{x}$ to another state $\mathbf{x}'$ along a curve of constant volume, $x^k = \text{const.}$ for $k = 1, \ldots, n-1$. Such a curve can be thought of as 'cooling at constant volume' and is achieved purely by heat transfer; no mechanical work is done,

$$\gamma_1^*\omega = \gamma_1^* \sum_{k=1}^{n-1} P_k\, \mathrm{d}x^k = \sum_{k=1}^{n-1} P_k\, \mathrm{d}(x^k \circ \gamma_1) = 0.$$

It then follows that no reversible adiabatic transition γ_2 such that $\gamma_2^*\theta = 0$ exists between these two states. Since processes such as γ_1 may always be assumed to be locally possible, it follows that every state $\mathbf{x}$ has equilibrium states in its neighbourhood that cannot be reached by quasi-static adiabatic paths. This leads to Carathéodory's more general version of the second law.

Second law of thermodynamics: entropy. *In a thermodynamic system K, every neighbourhood U of an arbitrary equilibrium state* $\mathbf{x}$ *contains a state* $\mathbf{x}'$ *that is inaccessible by a quasi-static adiabatic path from* $\mathbf{x}$.

Theorem 16.5 (Carathéodory) *The heat 1-form θ is integrable, $\theta \wedge \mathrm{d}\theta = 0$, if and only if every neighbourhood of any state $\mathbf{x} \in K$ contains a state $\mathbf{x}'$ adiabatically inaccessible from* $\mathbf{x}$.

Outline proof: If θ is integrable, then by Theorem 16.4 it is possible to find local coordinates $(U; y^i)$ of any state $\mathbf{x}$ such that $\theta\big|_U = Q_n \,\mathrm{d}y^n$. Adiabatics satisfy $\gamma^*\theta = 0$, or $y^n = \text{const}$. Hence, if U' is an open neighbourhood of $\mathbf{x}$ such that $U' \subseteq U$, any state $\mathbf{x}' \in U'$ such that $y'^n = y^n(\mathbf{x}') \neq y^n(\mathbf{x})$ is adiabatically inaccessible from $\mathbf{x}$.

Conversely, if $\theta \wedge \mathrm{d}\theta \neq 0$, then the 1-form θ is not integrable on an open subset U of every state $\mathbf{x} \in M$. Hence the distribution D^{n-1} such that $\theta \in (D^{n-1})^\perp$ is not involutive on an open neighbourhood U' of $\mathbf{x}$, so that $[D^{n-1},\ D^{n-1}] = \mathcal{T}(U')$. Let X and Y be vector fields in D^{n-1} such that $[X, Y]$ is not in the distribution. It may then be shown that every state $\mathbf{x}'$ is accessible by a curve of the form

$$t \mapsto \psi_{-\sqrt{t}} \circ \phi_{-\sqrt{t}} \circ \psi_{\sqrt{t}} \circ \phi_{\sqrt{t}} \mathbf{x}$$

where ψ_t and ϕ_t are local flows generated by the vector field X and Y (see Example 15.14). ■

For a reversible adiabatic process γ at constant volume we have $\gamma^*\theta = 0$ and

$$\gamma^*\omega = \sum_{k=1}^{n-1} P_k \frac{\mathrm{d}x^k(\gamma(t))}{\mathrm{d}t} \mathrm{d}t = 0.$$

Hence there is no change in internal energy for such processes, $\Delta U = 0$. On the other hand, for an irreversible adiabatic process at constant volume, such as stirring a gas in an adiabatic enclosure, there is always an increase in internal energy, $U' > U$. Hence all states with $U' < U$ are adiabatically inaccessible by adiabatic processes at constant volume, be they reversible or not. As remarked above, it is impossible to 'unstir' a gas. In general, for any two states $\mathbf{x}$ and $\mathbf{x}'$ either (i) $\mathbf{x}$ is adiabatically inaccessible to $\mathbf{x}'$, (ii) $\mathbf{x}'$ is adiabatically inaccessible to $\mathbf{x}$, or (iii) there exists a reversible quasi-static process from $\mathbf{x}$ and $\mathbf{x}'$.

From Theorem 16.5 and Carathéodory's statement of the second law, the heat form θ can be expressed as

$$\theta = f \,\mathrm{d}s$$

where f and s are real-valued functions on K. Any function $s(x^1, \ldots, x^{n-1}, U)$ for which this holds is known as an **empirical entropy**. A reversible adiabatic process $\gamma : [a, b] \to K$ is clearly **isentropic**, $s = \text{const.}$, since $\gamma^*\theta = 0$ along the process, and the hypersurface $s = \text{const.}$ through any state $\mathbf{x}$ represents the local boundary between adiabatically accessible and inaccessible states from $\mathbf{x}$.

For most thermodynamic systems the function s is globally defined by the identity $\theta = f\mathrm{d}s$. Since a path in K connecting adiabatically accessible states has $\mathrm{d}U/\mathrm{d}t \geq 0$, we can assume that s is a monotone increasing function of U for fixed volume coordinates

$x^1, \ldots, x^{n-1}$. For any path γ with $x^k = \text{const.}$ for $k = 1, \ldots, n-1$, such that $\gamma^*\omega = 0$, it follows that

$$\frac{\mathrm{d}U}{\mathrm{d}t} = \langle \dot{\gamma}, \theta \rangle = f \frac{\mathrm{d}s}{\mathrm{d}t}$$

and the function f must be everywhere positive.

Absolute entropy and temperature

Consider two systems A and B in an adiabatic enclosure and in equilibrium through mutual contact with a diathermic wall. In place of variables $x^1, \ldots, x^{n-1}, U_A$ for states of system A let us use variables $x^1, \ldots, x^{n-2}, s_A, \tau_A$ where τ_A is the empirical temperature, and similarly use variables $y^1, \ldots, y^{m-2}, s_B, \tau_B = \tau_A$ for states of system B. The combined system then has coordinates $x^1, \ldots, x^{n-2}, y^1, \ldots, y^{m-2}, s_A, s_B, \tau = \tau_A = \tau_B$. Since work done in any reversible process is an additive quantity, $\Delta W = \Delta W_A + \Delta W_B$, we may assume from the first law of thermodynamics that U is an additive function, $U = U_A + U_B$. Hence the work 1-form may be assumed to be additive, $\omega = \omega_A + \omega_B$, and so is the heat 1-form

$$\theta = \omega + \mathrm{d}U = \omega_A + \omega_B + \mathrm{d}U_A + \mathrm{d}U_B = \theta_A + \theta_B, \tag{16.15}$$

which can be written

$$f\,\mathrm{d}s = f_A\,\mathrm{d}s_A + f_B\,\mathrm{d}s_B, \tag{16.16}$$

where $f_A = f_A(x^1, \ldots, x^{n-2}, s_A, \tau)$ and $f_B = f_B(y^1, \ldots, y^{m-2}, s_B, \tau)$. Since s is a function of all variables $s = s(x^1, \ldots, y^{m-2}, s_A, s_B, \tau)$, it follows that $s = s(s_A, s_B)$ and

$$\frac{f_A}{f} = \frac{\partial s}{\partial s_A}, \qquad \frac{f_B}{f} = \frac{\partial s}{\partial s_B}. \tag{16.17}$$

Hence $f = f(s_A, s_B, \tau)$, $f_A = f_A(s_A, \tau)$, $f_B = f_B(s_B, \tau)$ and

$$\frac{\partial \ln f_A}{\partial \tau} = \frac{\partial \ln f_B}{\partial \tau} = \frac{\partial \ln f}{\partial \tau} = g(\tau)$$

for some function g. Setting $T(\tau) = \exp\left(\int g(\tau)\,\mathrm{d}\tau\right)$,

$$f_A = T(\tau)F_A(s_A), \qquad f_B = T(\tau)F_B(s_B), \qquad f = T(\tau)F(s_A, s_B),$$

and Eq. (16.16) results in

$$F\mathrm{d}s = F_A\,\mathrm{d}s_A + F_B\,\mathrm{d}s_B. \tag{16.18}$$

By setting $S_A = \int F_A(s_A)\,\mathrm{d}s_A$ and $S_B = \int F_B(s_B)\,\mathrm{d}s_B$, we have

$$F\,\mathrm{d}s = \mathrm{d}S_A + \mathrm{d}S_B = \mathrm{d}S$$

where $S = S_A + S_B$. Hence

$$\theta_A = f_A\,\mathrm{d}s_A = TF_A\,\mathrm{d}s_A = T\,\mathrm{d}S_A, \qquad \theta_B = T\,\mathrm{d}S_B$$

and

$$\theta = TF\,\mathrm{d}s = T\,\mathrm{d}S, \tag{16.19}$$

which is consistent with the earlier requirement of additivity of heat forms, Eq. (16.15). The particular choice of empirical temperature T and entropy S such that (16.19) holds, and which has the additivity property $S = S_A + S_B$, is called **absolute temperature** and **absolute entropy**. In the literature one often finds the formula $\mathrm{d}Q = T\,\mathrm{d}S$ in place of (16.19) but this notation is not good, for the right-hand side is not an exact differential as $\mathrm{d}\theta \neq 0$ in general.

When $\mathrm{d}\theta \neq 0$ the original variables τ and s are independent and only simple scaling freedoms are available for absolute temperature and entropy. For example, if

$$\theta = T\,\mathrm{d}S = T'\,\mathrm{d}S'$$

then

$$\frac{T'(\tau)}{T(\tau)} = \frac{\mathrm{d}S'(s)}{\mathrm{d}S(s)} = a = \text{const.},$$

where $a > 0$ if the rule $\Delta S > 0$ for adiabatically accessible states is to be preserved. Hence

$$T' = aT, \qquad S' = \frac{1}{a}S + b.$$

Only a positive scaling may be applied to absolute temperature and there is an **absolute zero** of temperature; absolute entropy permits an affine transformation, consisting of both a rescaling and change of origin.

Example 16.4 An ideal or perfect gas is determined by two variables, volume V and absolute temperature T. The heat 1-form is given by

$$\theta = \mathrm{d}U + p\,\mathrm{d}V = T\,\mathrm{d}S\,.$$

Using

$$\mathrm{d}\Big(\frac{\theta}{T}\Big) = \mathrm{d}^2 S = 0$$

we have

$$-\frac{1}{T^2}\mathrm{d}T \wedge \mathrm{d}U + \mathrm{d}\Big(\frac{p}{T}\Big) \wedge \mathrm{d}V = 0$$

and setting $U = U(V, T),\ p = p(V, T)$ results in

$$\Big[-\frac{1}{T^2}\Big(\frac{\partial U}{\partial V}\Big)_T - \frac{p}{T^2} + \frac{1}{T}\Big(\frac{\partial p}{\partial T}\Big)_V\Big]\mathrm{d}T \wedge \mathrm{d}V = 0.$$

Hence

$$T\Big(\frac{\partial p}{\partial T}\Big)_V = \Big(\frac{\partial U}{\partial V}\Big)_T + p. \tag{16.20}$$

For a gas in an adiabatic enclosure, classic experiments of Gay-Lussac and Joule have led to the conclusion that $U = U(T)$. Substituting into (16.20) results in

$$\frac{\partial \ln p}{\partial T} = \frac{1}{T},$$

which integrates to give a function $f(V)$ such that

$$f(V)p = T.$$

Comparing with the discussion in Example 16.3, we have for a single mole of gas

$$Vp = RT_g,$$

and since $T = T(T_g)$ it follows that after a suitable scaling of temperature we may set $f(V) = V$ and $T = T_g$. Thus for an ideal gas the absolute temperature is identical with absolute gas temperature.

From $\theta = T\,\mathrm{d}S = \mathrm{d}U + p\,\mathrm{d}V$ we have

$$\mathrm{d}S = \frac{1}{T}\mathrm{d}U + \frac{R}{V}\mathrm{d}V$$

and the formula for absolute entropy of an ideal gas is

$$S = \int \frac{1}{T}\frac{\mathrm{d}U}{\mathrm{d}T}\mathrm{d}T + R\ln V.$$

Problem

Problem 16.9 For a reversible process $\sigma : T \to K$, using absolute temperature T as the parameter, set

$$\sigma^*\theta = c\,\mathrm{d}T$$

where c is known as the **specific heat** for the process. For a perfect gas show that for a process at constant volume, $V = \text{const.}$, the specific heat is given by

$$c_V = \left(\frac{\partial U}{\partial T}\right)_V.$$

For a process at constant pressure show that

$$c_p = c_V + R,$$

while for an adiabatic process, $\sigma^*\theta = 0$,

$$pV^\gamma = \text{const.} \quad \text{where} \quad \gamma = \frac{c_p}{c_V}.$$

16.5 Classical mechanics

Classical analytic mechanics comes in two basic forms, Lagrangian or Hamiltonian. Both have natural formulations in the language of differential geometry, which we will outline in this section. More details may be found in [2, 10–14] and [4, chap. 13].

Calculus of variations

The reader should have at least a rudimentary acquaintance with the calculus of variations as found in standard texts on applied mathematics such as [15]. The following is a brief introduction to the subject, as it applies to parametrized curves on manifolds.

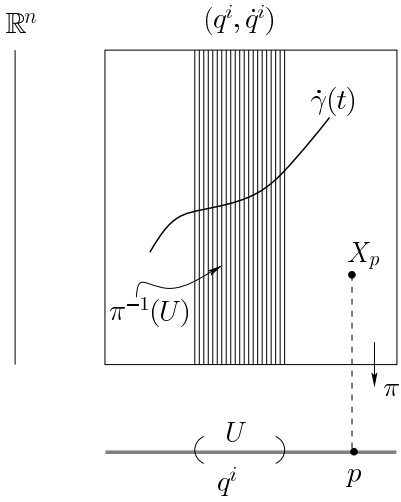

Figure 16.1 Tangent bundle

Let M be any differential manifold and TM its tangent bundle (refer to Section 15.3; see Fig. 16.1). If $\gamma : \mathbb{R} \to M$ is a smooth curve on M, define its **lift** to TM to be the curve $\dot{\gamma} : \mathbb{R} \to TM$ traced out by the tangent vector to the curve, so that $\dot{\gamma}(t)$ is the tangent vector to the curve at $\gamma(t)$ and $\pi(\dot{\gamma}(t)) = \gamma(t)$.

Exercise: Show that if $X(t)$ is the tangent to the curve $\dot{\gamma}(t)$ then $\pi_* X(t) = \dot{\gamma}(t) \in T_{\gamma(t)}(M)$.

A function $L : TM \to \mathbb{R}$ is called a **Lagrangian function**, and for any parametrized curve $\gamma : [t_0, t_1] \to M$ we define the corresponding **action** to be

$$S[\gamma] = \int_{t_0}^{t_1} L(\dot{\gamma}(t))\, \mathrm{d}t. \tag{16.21}$$

If $(q^1, \ldots, q^n)$ are local coordinates on M let the induced local coordinates on the tangent bundle TM be written $(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$. This notation may cause a little concern to the reader, but it is much loved by physicists – the quantities $\dot{q}^i$ are independent quantities, not to be thought of as 'derivatives' of q^i unless a specific curve γ having coordinate representation $q^i = q^i(t)$ is given. In that case, and only then, we find $\dot{q}^i(t) = \mathrm{d}q^i(t)/\mathrm{d}t$ *along the lift* $\dot{\gamma}(t)$ *of the curve*. Otherwise, the $\dot{q}^i$ refer to all possible components of tangent vectors at that point of M having coordinates q^j. A Lagrangian can be written as a function of $2n$ variables, $L(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$.

By a **variation** of a given curve $\gamma : [t_0, t_1] \to M$ (see Fig. 16.2) is meant a one-parameter family of curves $\gamma : [t_0, t_1] \times [-a, a] \to M$ such that for all $\lambda \in [-a, a]$

$$\gamma(t_0, \lambda) = \gamma(t_0) \quad \text{and} \quad \gamma(t_1, \lambda) = \gamma(t_1)$$

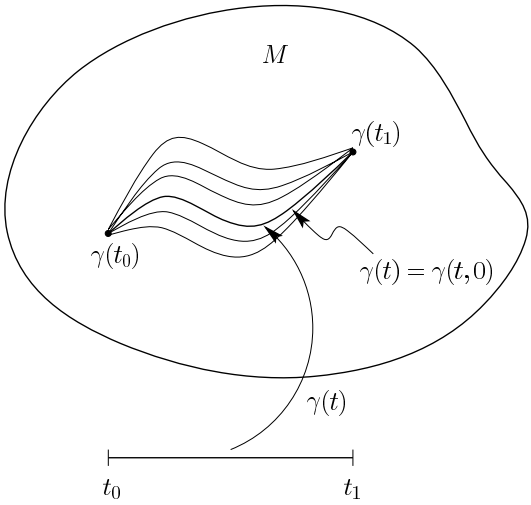

Figure 16.2 Variation of a curve

and the member of the family defined by $\lambda = 0$ is the given curve γ,

$$\gamma(t, 0) = \gamma(t) \qquad \text{for all } t_0 \leq t \leq t_1.$$

For each t in the range $[t_0, t_1]$ we define the *connection curve* $\gamma_t : [-a, a] \to M$ by $\gamma_t(\lambda) = \gamma(t, \lambda)$. Its tangent vector along the curve $\lambda = 0$ is written $\delta\gamma$, whose value at $t \in [t_0, t_1]$ is determined by the action on an arbitrary function $f : M \to \mathbb{R}$,

$$\delta\gamma_t(f) = \frac{\partial f(\gamma_t(\lambda))}{\partial\lambda}\Big|_{\lambda=0}. \tag{16.22}$$

This is referred to as the **variation field** along the curve. In traditional literature it is simply referred to as the 'variation of the curve'. Since all curves of the family meet at the end points $t = t_0, t_1$, the quantity on the right-hand side of Eq. (16.22) vanishes,

$$\delta\gamma_{t_0} = \delta\gamma_{t_1} = 0. \tag{16.23}$$

The lift of the variation field to the tangent bundle is a curve $\delta\dot{\gamma} : [t_0, t_1] \to TM$, which starts at the zero vector in the fibre above $\gamma(t_0)$ and ends at the zero vector in the fibre above $\gamma(t_1)$. In coordinates,

$$\delta\dot{\gamma}(t) = \big(\delta q^1(t), \ldots, \delta q^n(t), \delta\dot{q}^1(t), \ldots, \delta\dot{q}^n(t)\big)$$

where

$$\delta q^i(t) = \frac{\partial q^i(t,\lambda)}{\partial\lambda}\Big|_{\lambda=0}, \qquad \delta\dot{q}^i(t) = \frac{\partial\dot{q}^i(t,\lambda)}{\partial\lambda}\Big|_{\lambda=0} = \frac{\partial\delta q^i(t)}{\partial t}.$$

Exercise: Justify the final identity in this equation.

The action $S[\gamma]$ becomes a function of λ if γ is replaced by its variation γ_λ. We say that a curve $\gamma : [t_0, t_1] \to M$ is an **extremal** if for every variation of the curve

$$\delta S \equiv \frac{\mathrm{d}S}{\mathrm{d}\lambda}\Big|_{\lambda=0} = \int_{t_0}^{t_1} \delta L \, \mathrm{d}t = 0 \tag{16.24}$$

where

$$\begin{aligned} \delta L &\equiv \frac{\partial L\big(\dot{\gamma}_\lambda\big)}{\partial \lambda}\Big|_{\lambda=0} \\ &= \langle \mathrm{d}L \,|\, \delta\dot{\gamma} \rangle \\ &= \frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i. \end{aligned}$$

Substituting in (16.24) and performing an integration by parts results in

$$0 = \delta S = \int_{t_0}^{t_1} \left[\frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}^i}\Big) \right] \mathrm{d}t + \frac{\partial L}{\partial \dot{q}^i}\delta q^i \Big|_{t_0}^{t_1}.$$

The final term vanishes on account of $\delta q^i = 0$ at $t = t_0, t_1$, and since the $\delta q^i(t)$ are essentially arbitrary functions on the interval $[t_0, t_1]$ subject to the end-point constraints it may be shown that the term in the integrand must vanish,

$$\frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}^i}\Big) = 0. \tag{16.25}$$

These are known as the **Euler–Lagrange equations**.

Example 16.5 In the plane, the shortest curve between two fixed points is a straight line. To prove this, use the length as action

$$S = \int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2}\, \mathrm{d}t.$$

Setting $t = x$ and replacing $\dot{}$ with $'$ there is a single variable $q^1 = y$ and the Lagrangian is $L = \sqrt{1 + (y')^2}$. The Euler–Lagrange equation reads

$$\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\partial L}{\partial y'}\Big) = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{y'}{\sqrt{1 + (y')^2}}\Big) = 0$$

with solution

$$\frac{y'}{\sqrt{1 + (y')^2}} = \text{const.}$$

Hence $y' = a$ for some constant a, and the extremal curve is a straight line $y = ax + b$.

Lagrangian mechanics

In Newtonian mechanics a dynamical system of N particles is defined by positive real scalars $m_1, m_2, \ldots, m_N$ called the **masses** of the particles, and parametrized curves $t \mapsto \mathbf{r}_a = \mathbf{r}_a(t)$ $(a = 1, 2, \ldots, N)$ where each $\mathbf{r}_a \in \mathbb{R}^3$. The parameter t is interpreted as *time*.

The **kinetic energy** of the system is defined as

$$T = \frac{1}{2}\sum_{a=1}^{N} m_a \dot{\mathbf{r}}_a^2 \quad \text{where} \quad \dot{\mathbf{r}}_a = \frac{\mathrm{d}\mathbf{r}_a}{\mathrm{d}t} \quad \text{and} \quad \dot{\mathbf{r}}_a^2 = \dot{x}_a^2 + \dot{y}_a^2 + \dot{z}_a^2.$$

We will also assume **conservative systems** in which Newton's second law reads

$$m_a \ddot{\mathbf{r}}_a = -\nabla_a U \equiv -\frac{\partial U(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)}{\partial \mathbf{r}_a}, \tag{16.26}$$

where the given function $U : \mathbb{R}^{3N} \to r$ is known as the **potential energy** of the system.

A **constrained system** consists of a Newtonian dynamical system together with a manifold M of dimension $n \leq 3N$, and a map $C : M \to \mathbb{R}^{3N}$, called the **constraint**. In a local system of coordinates $(V; (q^1, q^2, \ldots, q^n))$ on M, the constraint can be written as a set of functions

$$\mathbf{r}_a = \mathbf{r}_a(q^1, q^2, \ldots, q^n)$$

and n is called the **number of degrees of freedom** of the constrained system. The coordinates q^i $(i = 1, \ldots, n)$ are commonly called **generalized coordinates** for the constrained system. They may be used even for an unconstrained system, in which $n = 3N$ and V is an open submanifold of $\mathbb{R}^{3N}$; in this case we are essentially expressing the original Newtonian dynamical system in terms of general coordinates. It will always be assumed that (M, C) is an embedded submanifold of $\mathbb{R}^{3N}$, so that the tangent map C_* is injective everywhere. This implies that the matrix $[\partial \mathbf{r}_a / \partial q^i]$ has rank n everywhere (no critical points).

Using the chain rule

$$\dot{\mathbf{r}}_a = \frac{\partial \mathbf{r}_a}{\partial q^i}\dot{q}^i \quad \text{where} \quad \dot{q}^i = \frac{\mathrm{d}q^i}{\mathrm{d}t}$$

the kinetic energy for a constrained system may be written

$$T = \tfrac{1}{2} g_{ij} \dot{q}^i \dot{q}^j \tag{16.27}$$

where

$$g_{ij} = \sum_{a=1}^{N} m_a \frac{\partial \mathbf{r}_a}{\partial q^i} \cdot \frac{\partial \mathbf{r}_a}{\partial q^j}. \tag{16.28}$$

This is a tensor field of type $(0, 2)$ over the coordinate neighbourhood V, since

$$g'_{i'j'} = \sum_{a=1}^{N} m_a \frac{\partial \mathbf{r}_a}{\partial q'^{i'}} \cdot \frac{\partial \mathbf{r}_a}{\partial q'^{j'}} = g_{ij} \frac{\partial q^i}{\partial q'^{i'}} \frac{\partial q^j}{\partial q'^{j'}}.$$

At each point of $q \in M$ we can define an inner product on the tangent space T_q,

$$g(u, v) \equiv u \cdot v = g_{ij} u^i v^j \quad \text{where} \quad u = u^i \frac{\partial}{\partial q^i}, \quad v = v^j \frac{\partial}{\partial q^j},$$

which is positive definite since

$$g(u, u) = \sum_{a=1}^{N} m_a \left(\frac{\partial \mathbf{r}_a}{\partial q^i} u^i \right)^2 \geq 0$$

and the value 0 is only possible if $u^i = 0$ since the constraint map is an embedding and has no critical points. A manifold M with a positive definite inner product defined everywhere is called a **Riemannian manifold**; further discussion of such manifolds will be found in Chapter 18. The associated symmetric tensor field $g = g_{ij}\,\mathrm{d}q^i \otimes \mathrm{d}q^j$ is called the **metric tensor**. These remarks serve as motivation for the following definition.

A **Lagrangian mechanical system** consists of an n-dimensional Riemannian manifold (M, g) called **configuration space**, together with a function $L : TM \to \mathbb{R}$ called the **Lagrangian** of the system. The Lagrangian will be assumed to have the form $L = T - U$ where, for any $u = (q^i, \dot{q}^j) \in TM$

$$T(u) = \tfrac{1}{2}g(u, u) = \tfrac{1}{2}g_{ij}(q^1, \ldots, q^n)\dot{q}^i\dot{q}^j$$

and

$$U(u) = U(\pi(u)) = U(q^1, \ldots, q^n).$$

As for the calculus of variations it will be common to write $L(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$.

The previous discussion shows that every constrained system can be considered as a Lagrangian mechanical system with $U(q^1, \ldots, q^n) = U\big(\mathbf{r}_1(q^i), \ldots, \mathbf{r}_N(q^i)\big)$. In place of Newton's law (16.26) we postulate **Hamilton's principle**, that every motion $t \mapsto \gamma(t) \equiv q^i(t)$ of the system is an extremal of the action determined by the Lagrangian L

$$\delta S = \int_{t_0}^{t_1} \delta L\,\mathrm{d}t = 0.$$

The equations of motion are then the second-order differential equations (16.25),

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}^i}\Big) - \frac{\partial L}{\partial q^i} = 0, \qquad L = T - U \tag{16.29}$$

known as **Lagrange's equations**.

Example 16.6 A Newtonian system of N unconstrained particles, $\mathbf{r}_a = (x_a, y_a, z_a)$ can be considered also as a Lagrangian system with $3N$ degrees of freedom if we set

$$q^1 = x_1,\ q^2 = y_1,\ q^3 = z_1,\ q^4 = x_2,\ \ldots,\ q^{3N} = z_N.$$

The metric tensor is diagonal with $g_{11} = m_1$, $g_{22} = m_1$, $\ldots$, $g_{44} = m_2$, $\ldots$, etc. Lagrange's equations (16.29) read, for $i = 3a - 2$ $(a = 1, 2, \ldots, N)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{x}_a}\Big) - \frac{\partial L}{\partial x_a} = \frac{\mathrm{d}}{\mathrm{d}t}\big(m_a\dot{x}_a\big) + \frac{\partial U}{\partial x_a} = 0,$$

that is,

$$m_a\ddot{x}_a = -\frac{\partial U}{\partial x_a}$$

and similarly

$$m_a\ddot{y}_a = -\frac{\partial U}{\partial y_a}, \qquad m_a\ddot{z}_a = -\frac{\partial U}{\partial z_a}$$

in agreement with Eq. (16.26).

For a single particle, in spherical polar coordinates, $q^1 = r > 0$, $0 < q^2 = \theta < \pi$, $0 < q^3 = \phi < 2\pi$,

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta$$

the kinetic energy is

$$T = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2\right).$$

Hence the metric tensor g has components

$$[g_{ij}] = \begin{pmatrix} m\dot{r}^2 & 0 & 0 \\ 0 & mr^2\dot{\theta}^2 & 0 \\ 0 & 0 & mr^2\sin^2\theta\dot{\phi}^2 \end{pmatrix}$$

and Lagrange's equations for a *central potential* $U = U(r)$ read

$$m\ddot{r} - mr\dot{\theta}^2 - mr\sin^2\theta\dot{\phi}^2 + \frac{\mathrm{d}U}{\mathrm{d}r} = 0,$$
$$m\frac{\mathrm{d}}{\mathrm{d}t}\left(r^2\dot{\theta}\right) - mr^2\sin\theta\cos\theta\dot{\phi}^2 = 0,$$
$$m\frac{\mathrm{d}}{\mathrm{d}t}\left(r^2\sin^2\theta\dot{\phi}\right) = 0.$$

Exercise: Write out the equations of motion for a particle constrained to the plane $z = 0$ in polar coordinates, $x = r\cos\theta$, $y = r\sin\theta$.

Example 16.7 The plane pendulum has configuration space $M = S^1$, the one-dimensional circle, which can be covered with two charts, $0 < \phi_1 < 2\pi$ and $-\pi < \phi_2 < \pi$, such that on the overlaps they are related by

$$\theta_2 = \theta_1 \qquad \text{for } 0 < \theta_1 \leq \pi$$
$$\theta_2 = \theta_1 - 2\pi \qquad \text{for } \pi \leq \theta_1 < 2\pi$$

and constraint functions embedding this manifold in $\mathbb{R}^3$ are

$$x = 0, \qquad y = -a\sin\theta_1, \qquad z = -a\cos\theta_1;$$
$$x = 0, \qquad y = -a\sin\theta_2, \qquad z = -a\cos\theta_2.$$

For $\theta = \theta_1$ or $\theta = \theta_2$ we have

$$T = \frac{m}{2}a^2\dot{\theta}^2, \qquad U = mgz = -mga\cos\theta, \qquad L(\theta,\dot{\theta}) = T - U$$

and subsituting in Lagrange's equations (16.29) with $q^1 = \theta$ gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial\dot{\theta}}\right) - \frac{\partial L}{\partial\theta} = ma^2\ddot{\theta} + mga\sin\theta = 0.$$

For small values of θ, the pendulum hanging near vertical, the equation approximates the simple harmonic oscillator equation

$$\ddot{\theta} + \frac{g}{a}\theta = 0$$

with period $\tau = 2\pi\sqrt{g/a}$.

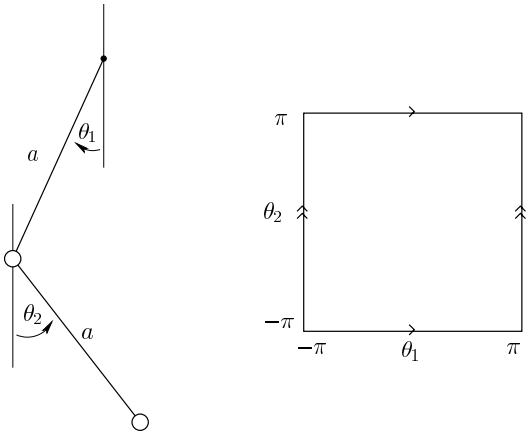

Figure 16.3 Double pendulum

Example 16.8 The spherical pendulum is similar to the plane pendulum, but the configuration manifold is the 2-sphere S^2. In spherical polars the constraint is

$$x = a\sin\theta\cos\phi, \qquad y = a\sin\theta\sin\phi, \qquad z = -a\cos\theta$$

and as for Example 16.6 we find

$$T = \frac{m}{2}a^2\big(\dot\theta^2 + \sin^2\theta\dot\phi^2\big), \qquad U = mgz = -mga\cos\theta.$$

Exercise: Write out Lagrange's equations for the spherical pendulum.

Example 16.9 The double pendulum consists of two plane pendula, of lengths a, b and equal mass m, one suspended from the end of the other (see Fig. 16.3). The configuration manifold is the 2-torus $M = S^1 \times S^1 = T^2$, and constraint functions are

$$x_1 = x_2 = 0, \qquad y_1 = -a\sin\theta_1, \qquad z_1 = -a\cos\theta_1,$$
$$y_2 = -a\sin\theta_1 - b\sin\theta_2, \qquad z_2 = -a\cos\theta_1 - b\cos\theta_2.$$

The kinetic energy is

$$\begin{aligned} T &= \frac{m}{2}\big(\dot y_1^2 + \dot z_1^2 + \dot y_2^2 + \dot z_2^2\big) \\ &= \frac{m}{2}\big(2a^2\dot\theta_1{}^2 + b^2\dot\theta_2{}^2 + 2ab\cos(\theta_1 - \theta - 2)\dot\theta_1\dot\theta_2\big) \end{aligned}$$

and the potential energy is $U = -2mga\cos\theta_1 - mgb\cos\theta_2$.

Exercise: Write out Lagrange's equations for the double pendulum of this example.

Exercise: Write out the Lagrangian for a double pendulum with unequal masses, m_1 and m_2.

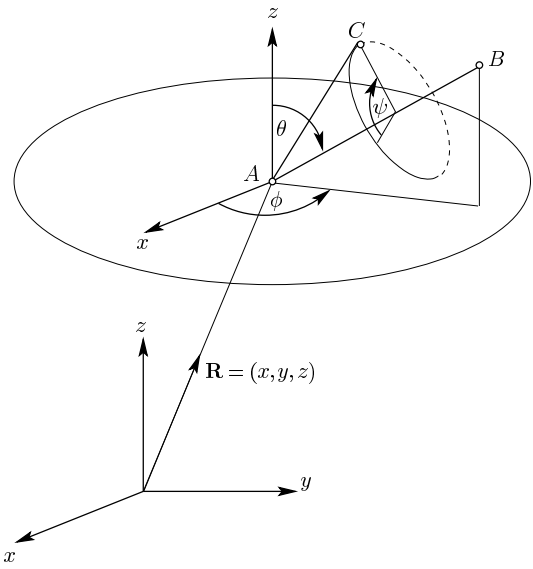

Figure 16.4 Degrees of freedom of a rigid body

Example 16.10 A rigid body is a system of particles subject to the constraint that all distances between particles are constant, $|\mathbf{r}_a - \mathbf{r}_b| = c_{ab} = \text{const}$. These equations are not independent since their number is considerably greater in general than the number of components in the $\mathbf{r}_a$. The number of degrees of freedom is in general six, as can be seen from the following argument. Fix a point in the object A, such as its centre of mass, and assign to it three rectangular coordinates $\mathbf{R} = (X, Y, Z)$. Any other point B of the body is at a fixed distance from A and therefore is constrained to move on a sphere about A. It can be assigned two spherical angles θ, ϕ as for the spherical pendulum. The only remaining freedom is a rotation by an angle ψ, say, about the axis AB. Every point of the rigid body is now determined once these three angles are specified (see Fig. 16.4). Thus the configuration manifold of the rigid body is the six-dimensional manifold $\mathbb{R}^3 \times S^2 \times S^1$. Alternatively the freedom of the body about the point A may be determined by a member of the rotation group $SO(3)$, which can be specified by three Euler angles. These are the most commonly used generalized coordinates for a rigid body. Details may be found in [12, chap. 6].

Given a tangent vector $u = \dot{\gamma} = \dot{q}^i \partial_{q^i}$, the **momentum 1-form conjugate to** u is defined by

$$\langle \omega_u, v \rangle = g(u, v) = g_{ij} \dot{q}^i v^j.$$

Setting $\omega_u = p_i \, \mathrm{d}q^i$ we see that

$$p_i = g_{ij} \dot{q}^j = \frac{\partial L}{\partial \dot{q}^i}. \tag{16.30}$$

The last step follows either by direct differentiation of $L = \frac{1}{2} g_{ij} \dot{q}^i \dot{q}^j - U(q)$ or by applying Euler's theorem on homogeneous functions to $T(q^j, \lambda \dot{q}^i) = \lambda^2 T(q^j, \dot{q}^i)$. The components p_i of the momentum 1-form, given by Eq. (16.30), are called the **generalized momenta** conjugate to the generalized coordinates q^i.

Exercise: For a general Lagrangian L, not necessarily of the form $T - U$, show that $\omega = (\partial L / \partial \dot{q}^i)\, \mathrm{d}q^i$ is a well-defined 1-form on M.

Example 16.11 The generalized momentum for an unconstrained particle $L = \frac{1}{2} m \dot{\mathbf{r}}^2 - U(\mathbf{r})$ given by

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x}, \qquad p_y = \frac{\partial L}{\partial \dot{y}} = m\dot{y}, \qquad p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z},$$

which are the components of standard momentum $\mathbf{p} = (p_x, p_y, p_z) = m\dot{\mathbf{r}}$.

In spherical polar coordinates

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

whence

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = mr^2 \dot{\phi} \sin^2\theta.$$

This can be identified with the z-component of angular momentum,

$$L_z = \mathbf{r} \times \mathbf{p} \cdot \hat{\mathbf{z}} = m(x\dot{y} - y\dot{x}) = m\dot{\phi} r^2 \sin^2\theta.$$

It is a general result that the momentum conjugate to an angular coordinate about a fixed axis is the angular momentum about that axis.

Exercise: The angle θ in the previous example does not have a fixed axis of definition unless $\phi = \text{const}$. In this case show that $p_\theta = \mathbf{L} \cdot (-\cos\phi, \sin\phi, 0)$ and interpret geometrically.

Example 16.12 If the Lagrangian has no explicit dependence on a particular generalized coordinate q^k, so that $\partial L / \partial q^k = 0$, it is called an **ignorable** or **cyclic** coordinate, The corresponding generalized momentum p_k is then a constant of the motion, for the kth Lagrange's equation reads

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^k} - \frac{\partial L}{\partial q_k} = \frac{\mathrm{d}p_k}{\mathrm{d}t}.$$

This is a particular instance of a more general statement, known as **Noether's theorem**.

Let $\varphi_s : M \to M$ be a local one-parameter group of motions on M, generating the vector field X by

$$X_q f = \left.\frac{\partial f\big(\varphi_s(q)\big)}{\partial s}\right|_{s=0}.$$

The tangent map φ_{*s} induces a local flow on the tangent bundle, since $\varphi_{*s} \circ \varphi_{*t} = \big(\varphi_s \circ \varphi_t\big)_* = \varphi_{*(s+t)}$, and the Lagrangian is said to be **invariant** under this local one-parameter group if $L(\varphi_{*s} u) = L(u)$ for all $u \in TM$. Noether's theorem asserts that the quantity $\langle \omega_u, X \rangle$

is then a constant of the motion. The result is most easily proved in natural coordinates on TM.

Let $q^i(t)$ be any solution of Lagrange's equations, and set $\mathbf{q}(s,t) = \varphi_s \mathbf{q}(t)$. On differentiation with respect to t we have

$$\dot{\mathbf{q}}(s,t) = \varphi_s \dot{\mathbf{q}}(t) = \frac{\partial \mathbf{q}(s,t)}{\partial t}$$

and invariance of the Lagrangian implies, using Lagrange's equations at $s = 0$,

$$\begin{aligned} 0 = \frac{\partial L}{\partial s}\Big|_{s=0} &= \frac{\partial}{\partial s}\Big(L(\mathbf{q}(s,t), \dot{\mathbf{q}}(s,t))\Big)\Big|_{s=0} \\ &= \Big(\frac{\partial L}{\partial q^i}\frac{\partial q^i}{\partial s} + \frac{\partial L}{\partial \dot{q}^i}\frac{\partial \dot{q}^i}{\partial s}\Big)\Big|_{s=0} \\ &= \Big(\frac{\partial}{\partial t}\Big(\frac{\partial L}{\partial \dot{q}^i}\Big)\frac{\partial q^i}{\partial s} + \frac{\partial L}{\partial \dot{q}^i}\frac{\partial^2 q^i}{\partial s \partial t}\Big)\Big|_{s=0} \\ &= \frac{\partial}{\partial t}\Big(\frac{\partial L}{\partial \dot{q}^i}\frac{\partial q^i}{\partial s}\Big)\Big|_{s=0} \\ &= \frac{\partial}{\partial t}\big(p_i X^i\big). \end{aligned}$$

Hence, along any solution of Lagrange's equations, we have an integral of the motion

$$\langle \omega_u, X\rangle = g_{ij}\dot{q}^i X^j = p_i X^j = \text{const.}$$

The one-parameter group is often called a **symmetry group** of the system, and Noether's theorem exhibits the relation between symmetries and conservation laws.

If q^k is an ignorable coordinate then the one-parameter group of motions

$$\varphi_s(\mathbf{q}) = (q^1, \ldots, q^{k-1}, q^k + s, q^{k+1}, \ldots, q^n)$$

is an invariance group of the Lagrangian. It generates the vector field $X^i = \partial(\varphi_s(\mathbf{q}))^i/\partial s\big|_{s=0} = \delta^i_k$, and the associated constant of the motion is the generalized momentum $p_i X^i = p_i \delta^i_k = p_k$ conjugate to the ignorable coordinate.

Hamiltonian mechanics

A 2-form Ω is said to be **non-degenerate** at $p \in M$ if

$$\Omega_p(X_p, Y_p) = 0 \quad \text{for all } Y_p \in T_p(M) \Longrightarrow X_p = 0.$$

As for the concept of non-singularity for inner products (Chapter 5), this is true if and only if

$$\Omega_p = A_{ij}(\mathrm{d}x^i)_p \wedge (\mathrm{d}x^j)_p \quad \text{where} \quad A_{ij} = -A_{ji},\ \det[A_{ij}] \neq 0.$$

Exercise: Prove this statement.

The manifold M must necessarily be of even dimension $m = 2n$ if there exists a non-degenerate 2-form, since $\det \mathsf{A} = \det \mathsf{A}^T = \det(-\mathsf{A}) = (-1)^m \det \mathsf{A}$. A **symplectic structure** on a $2n$-dimensional manifold M is a closed differentiable 2-form Ω that is everywhere non-degenerate. Recall that *closed* means that $\mathrm{d}\Omega = 0$ everywhere. An even-dimensional manifold M with a symplectic structure is called a **symplectic manifold**.

As in Examples 7.6 and 7.7, a symplectic form Ω induces an isomorphic map $\bar{\Omega} : T_p(M) \to T_p^*(M)$ where the covector $\overline{X_p} = \bar{\Omega} X_p$ is defined by

$$\langle \overline{X_p}, Y_p \rangle \equiv \langle \bar{\Omega} X_p, Y_p \rangle = \Omega(X_p, Y_p). \tag{16.31}$$

We may naturally extend this correspondence to one between vector fields and differential 1-forms $X \leftrightarrow \overline{X}$ such that for any vector field Y

$$\langle \overline{X}, Y \rangle = \Omega(X, Y).$$

By Eq. (16.10), we find for any vector field

$$\overline{X} = \tfrac{1}{2} \mathrm{i}_X \Omega. \tag{16.32}$$

In components $\overline{X}_i = A_{ji} X^j$.

We will write the vector field corresponding to a 1-form by the same notation $\overline{\omega} = \bar{\Omega}^{-1} \omega$, such that

$$\langle \omega, Y \rangle = \Omega(\overline{\omega}, Y) \tag{16.33}$$

for all vector fields Y. A vector field X is said to be a **Hamiltonian vector field** if there exists a function H on M such that $X = \overline{\mathrm{d}H}$, or equivalently $\overline{X} = \mathrm{d}H$. The function H is called the **Hamiltonian** generating this vector field. A function f is said to be a **first integral of the phase flow** generated by the Hamiltonian vector field $X = \overline{\mathrm{d}H}$ if $Xf = 0$. The Hamiltonian H is a first integral of the phase flow, for

$$X(H) = \langle \mathrm{d}H, X \rangle = \langle \mathrm{d}H, \overline{\mathrm{d}H} \rangle = \Omega(\overline{\mathrm{d}H}, \overline{\mathrm{d}H}) = 0$$

on setting $\omega = \mathrm{d}H$ and $Y = \overline{\mathrm{d}H}$ in Eq. (16.33) and using the antisymmetry of Ω.

Any function $f : M \to \mathbb{R}$ is known as a **dynamical variable**. For any dynamical variable f we set $X_f = \overline{\mathrm{d}f}$ to be the Hamiltonian vector field generated by f. Then for any vector field Y,

$$\Omega(X_f, Y) = \langle \overline{X_f}, Y \rangle = \langle \mathrm{d}f, Y \rangle = Y(f),$$

and we have the identity

$$\mathrm{i}_{X_f} \Omega = 2\, \mathrm{d}f.$$

Define the **Poisson bracket** of two dynamical variables f and g to be

$$(f, g) = \Omega(X_f, X_g), \tag{16.34}$$

from which

$$(f, g) = \langle \mathrm{d}f, X_g \rangle = X_g f = -X_f g = -(f, g).$$

In these and other conventions, different authors adopt almost random sign conventions – so beware of any discrepencies between formulae given here and those in other books!

From Eq. (16.13) we have that $\mathrm{d}\Omega = 0$ implies

$$\mathcal{L}_X\Omega = \mathrm{i}_X\,\mathrm{d}\Omega + \mathrm{d}\circ\mathrm{i}_X\Omega = \mathrm{d}(\mathrm{i}_X\Omega),$$

whence the Lie derivative of the symplectic form in any Hamiltonian direction vanishes,

$$\mathcal{L}_{X_f}\Omega = 2\,\mathrm{d}(\mathrm{d}f) = 2\,\mathrm{d}^2 f = 0.$$

Using Eq. (16.12) with $X = X_f$ and $Y = Y_g$ we obtain

$$\mathrm{i}_{[X_f,X_g]}\Omega = \mathcal{L}_{X_f}\big(\mathrm{i}_{X_g}\Omega\big).$$

By (16.8),

$$\mathrm{i}_{[X_f,X_g]}\Omega = 2\mathcal{L}_{X_f}\mathrm{d}g = 2\,\mathrm{d}\mathcal{L}_{X_f}g = 2\,\mathrm{d}(X_f g) = \mathrm{i}_{X_{(g,f)}}\Omega,$$

whence

$$[X_f, X_g] = X_{(g,f)} = -X_{(f,g)}. \tag{16.35}$$

From the Jacobi identity (15.24) it then follows that

$$\big((f,g),h\big) + \big((g,h),f\big) + \big((h,f),g\big) = 0. \tag{16.36}$$

Exercise: Prove Eq. (16.36).

Exercise: Show that $(f,g)+(f,h) = (f,g+h)$ and $(f,gh) = g(f,h)+h(f,g)$.

The rate of change of a dynamical variable f along a Hamiltonian flow is given by

$$\dot{f} = \frac{\mathrm{d}f}{\mathrm{d}t} = X_H f = (f,H). \tag{16.37}$$

Thus f is a first integral of the phase flow generated by the Hamiltonian vector field X_H if and only if it 'commutes' with the Hamiltonian, in the sense that its Poisson bracket with H vanishes, $(f,H) = 0$. The analogies with quantum mechanics (Chapter 14) are manifest.

Exercise: Show that if f and g are first integrals then so is (f,g).

Example 16.13 Let $M = \mathbb{R}^{2n}$ with coordinates labelled $(q^1,\ldots,q^n,\,p_1,\ldots,p_n)$. The 2-form $\Omega = 2\mathrm{d}q^i\wedge\mathrm{d}p_i = \mathrm{d}q^i\otimes\mathrm{d}p_i - \mathrm{d}p_i\otimes\mathrm{d}q^i$, having constant components

$$\mathsf{A} = \begin{pmatrix} \mathsf{O} & \mathsf{I} \\ -\mathsf{I} & \mathsf{O} \end{pmatrix}$$

is a symplectic structure, since $\det\mathsf{A} = 1$ (a simple exercise!) and it is closed,

$$\mathrm{d}\Omega = 2\,\mathrm{d}^2q^i\wedge\mathrm{d}p_i - \mathrm{d}q^i\wedge\mathrm{d}^2p_i = 0.$$

If X and Y are vector fields having components

$$X = \xi^i\frac{\partial}{\partial q^i} + \xi_j\frac{\partial}{\partial p_j}, \qquad Y = \eta^i\frac{\partial}{\partial q^i} + \eta_j\frac{\partial}{\partial p_j},$$

then

$$\langle \overline{X}, Y\rangle = \Omega(X, Y) = (\mathrm{d}q^i \otimes \mathrm{d}p_i - \mathrm{d}p_i \otimes \mathrm{d}q^i)(X, Y) = \xi^i \eta_i - \xi_i \eta^i$$

so that the 1-form $\overline{X}$ has components

$$\overline{X} = -\xi_j \, \mathrm{d}q^j + \xi^i \mathrm{d}p_i .$$

A Hamiltonian vector field X has $\xi_j = -\partial H/\partial q^i$ and $\xi^i = \partial H/\partial p_i$, so that

$$X = \overline{\mathrm{d}H} = X_H = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^j}\frac{\partial}{\partial p_j}.$$

A curve $\gamma : \mathbb{R} \to M$ is an integral curve of this vector field if the functions $q^i = q^i(t)$, $p_j = p_j(t)$ satisfy the differential equations known as **Hamilton's equations**:

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\mathrm{d}p_j}{\mathrm{d}t} = -\frac{\partial H}{\partial q^j}. \tag{16.38}$$

The Poisson bracket is given by

$$(f, g) = X_g f = \Big(\frac{\partial g}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial g}{\partial q^j}\frac{\partial}{\partial p_j}\Big) f = \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q^j}.$$

For any dynamical variable f it is straightforward to verify the Poisson bracket relations

$$(q^i, f) = \frac{\partial f}{\partial p_i}, \qquad (p_i, f) = -\frac{\partial f}{\partial q^i}, \tag{16.39}$$

from which the *canonical relations* are immediate

$$(q^i, q^j) = 0, \qquad (p_i, p_j) = 0, \qquad (q^i, p_j) = \delta^i{}_j.$$

Connection between Lagrangian and Hamiltonian mechanics

If M is a manifold of any dimension n its cotangent bundle T^*M, consisting of all covectors at all points, is a $2n$-dimensional manifold. If $(U; q^i)$ is any coordinate chart on M, a chart is generated on T^*M by assigning coordinates $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ to any covector $\omega_q = p_i(\mathrm{d}q^i)_q$ at $q \in M$. The natural projection map $\pi : T^*M \to M$ has the effect of sending any covector to its base point, $\pi(\omega_q) = q$. The tangent map corresponding to this projection map, $\pi_* : T_{\omega_q}(T^*M) \to T_q(M)$, maps every tangent vector $X_{\omega_q} \in T_{\omega_q}(T^*M)$ to a tangent vector $\pi_* X_{\omega_q} \in T_q(M)$. In canonical coordinates, set

$$X_{\omega_q} = \xi^i \frac{\partial}{\partial q^i} + \xi_j \frac{\partial}{\partial p_j}$$

and for any function $f : M \to \mathbb{R}$, written in coordinates as $f(q^1, \ldots, q^n)$, we have

$$(\pi_* X_{\omega_q}) f(\mathbf{q}) = X_{\omega_q}(f \circ \pi)(\mathbf{q}, \mathbf{p}) = \xi^i \frac{\partial f(\mathbf{q})}{\partial q^i} + \xi_i \frac{\partial f(\mathbf{q})}{\partial p_i}$$

so that

$$\pi_* X_{\omega_q} = \xi^i \frac{\partial}{\partial q^i}.$$

This defines a **canonical 1-form** θ on T^*M by setting

$$\theta_{\omega_q}(X_{\omega_q}) \equiv \langle \theta_{\omega_q}, X_{\omega_q} \rangle = \langle \omega_q, \pi_* X_{\omega_q} \rangle.$$

Alternatively, we can think of θ as the pullback $\theta_{\omega_q} = \pi^* \omega_q \in T *_{\omega_q} (T^*M)$, for

$$\langle \pi^* \omega_q, X_{\omega_q} \rangle = \langle \omega_q, \pi_* X_{\omega_q} \rangle = \theta_{\omega_q}(X_{\omega_q})$$

for arbitrary $X_{\omega_q} \in T_{\omega_q}(T^*M)$. Writing $\omega_q = p_i \, \mathrm{d}q^i$, we thus have $\langle \theta_{\omega_q}, X_{\omega_q} \rangle = p_i \xi^i$, so that in any canonical chart $(U \times \mathbb{R}^n; q^1, \ldots, q^n, p_1, \ldots, p_n)$

$$\theta = p_i \, \mathrm{d}q^i. \tag{16.40}$$

The 2-form

$$\Omega = -2 \, \mathrm{d}\theta = 2 \, \mathrm{d}q^i \wedge \mathrm{d}p_i \tag{16.41}$$

is of the same form as that in Example 16.13, and provides a natural symplectic structure on the cotangent bundle of any manifold M.

Given a Lagrangian system having configuration space (M, g) and Lagrangian function $L = T - U : TM \to \mathbb{R}$ where

$$T(\mathbf{q}, \dot{\mathbf{q}}) = \tfrac{1}{2} g_{ij}(\mathbf{q}) \dot{q}^i \dot{q}^j, \qquad U = U(\mathbf{q})$$

the cotangent bundle T^*M, consisting of momentum 1-forms on M, is known as the **phase space** of the system. The coordinates p_i and $\dot{q}^j$ are related by Eq. (16.30), so that velocity components can be expressed in terms of generalized momenta, $\dot{q}^j = g^{jk} p_k$ where $g^{jk} g_{ki} = \delta^j_i$, and Lagrange's equations (16.29) can be written

$$\dot{p}_i = \frac{\partial L}{\partial q^i}.$$

Our first task is to find a Hamiltonian function $H : T^*M \to \mathbb{R}$, written $H(q^1, \ldots, q^n, p_1, \ldots, p_n)$, such that the equations of motion of the system in phase space have the form of Hamiltonian equations (16.38) in Example 16.13. The Hamiltonian function H must then have exterior derivative

$$\begin{aligned}
\mathrm{d}H &= \frac{\partial H}{\partial q^i} \mathrm{d}q^i + H p_i \, \mathrm{d}p_i \\
&= -\dot{p}_i \mathrm{d}q^i + \dot{q}^i \mathrm{d}p_i \\
&= -\frac{\partial L}{\partial q^i} \mathrm{d}q^i + \mathrm{d}(\dot{q}^i p_i) - p_i \, \mathrm{d}\dot{q}^i \\
&= \mathrm{d}(\dot{q}^i p_i) - \left(\frac{\partial L}{\partial q^i} \mathrm{d}q^i + \frac{\partial L}{\partial \dot{q}^j} \mathrm{d}\dot{q}^j \right) \\
&= \mathrm{d}(\dot{q}^i p_i - L)
\end{aligned}$$

whence, within an arbitrary constant

$$H = \dot{q}^i p_i - L = g_{ij} \dot{q}^i \dot{q}^j - L = 2T - (T - U) = T + U = E.$$

The Hamiltonian is thus the energy of the system expressed in terms of canonical coordinates on T^*M.

Apart from expressing the equations of mechanics as a first-order system of equations, one of the advantages of the Hamiltonian view is that coordinates in which the symplectic form takes the form given in Example 16.13 need not be restricted to the canonical coordinates generated by the tangent bundle construction. For example, let $(q^i, p_j) \to (\bar{q}^i, \bar{p}_j)$ be any coordinate transformation such that the canonical 1-forms $\theta = p_i \, \mathrm{d}q^i$ and $\bar{\theta} = \bar{p}_i \, \mathrm{d}\bar{q}^i$ generate the same symplectic form,

$$\bar{\Omega} = -2 \, \mathrm{d}\bar{\theta} = \Omega = -2 \, \mathrm{d}\theta$$

so that $\bar{\theta} = \theta - \mathrm{d}F$ for some function F on T^*M.

Exercise: Show that $\bar{p}_i \dfrac{\partial \bar{q}^i}{\partial p_j} = -\dfrac{\partial F}{\partial p_j}$, $\bar{p}_i \dfrac{\partial \bar{q}^i}{\partial q^j} = p_j - \dfrac{\partial F}{\partial q^j}$.

Since $\Omega = \bar{\Omega}$ the Hamiltonian vector fields generated by any dynamical variable f, are identical for the two forms, $X_f = \bar{X}_f$, since for any vector field Y on T^*M,

$$\bar{\Omega}(\bar{X}_f, Y) = Yf = \Omega(X_f, Y).$$

Hence, Poisson brackets are invariant with respect to this change of coordinates, for

$$(f, g)_{\bar{q},\bar{p}} = \bar{X}_g f = X_g f = (f, g)_{g,p}.$$

This result is easy to prove directly by change of variables, as is done in some standard books on analytic mechanics. Using Eqs. (16.37) and (16.39) we have then

$$\frac{\mathrm{d}\bar{q}^i}{\mathrm{d}t} = (\bar{q}^i, H)_{q,p} = (\bar{q}^i, H)_{\bar{q},\bar{p}} = \frac{\partial H}{\partial \bar{p}_i},$$
$$\frac{\mathrm{d}\bar{p}_i}{\mathrm{d}t} = (\bar{p}_i, H)_{q,p} = (\bar{p}_i, H)_{\bar{q},\bar{p}} = -\frac{\partial H}{\partial \bar{q}^i},$$

and Hamilton's equations are preserved under such transformations. These are called **homogeneous contact transformations**.

More generally, let $H(q^1, \ldots, q^n, p_1, \ldots, p_n, t)$ be a time-dependent Hamiltonian, defined on **extended phase space** $T^*M \times \mathbb{R}$, where $\mathbb{R}$ represents the time variable t, and let λ be the **contact 1-form**,

$$\lambda = p_i \, \mathrm{d}q^i - H \, \mathrm{d}t.$$

If $T^*\bar{M} \to \mathbb{R}$ is another extended phase of the same dimension with canonical coordinates $\bar{q}^i$, $\bar{p}_i$ and Hamiltonian $\bar{H}(\bar{\mathbf{q}}, \bar{\mathbf{p}}, t)$, then a diffeomorphism $\phi : T^*M \times \mathbb{R} \to T^*\bar{M} \to \mathbb{R}$ is called a **contact transformation** if $\phi^* \, \mathrm{d}\bar{\lambda} = \mathrm{d}\lambda$. Since $\phi^* \circ \mathrm{d} = \mathrm{d} \circ \phi$ there exists a function F on $T^*M \times \mathbb{R}$ in the neighbourhood of any point such that $\phi^*\bar{\lambda} = \lambda - \mathrm{d}F$. If we write the function F as depending on the variables q^i and $\bar{q}^i$, which is generally possible locally,

$$\bar{p}_i \, \mathrm{d}\bar{q}^i - \bar{H} \, \mathrm{d}t = p_i \, \mathrm{d}q^i - H \, \mathrm{d}t - \frac{\partial F}{\partial q^i}\mathrm{d}q^i - \frac{\partial F}{\partial \bar{q}^i}\mathrm{d}\bar{q}^i - \frac{\partial F}{\partial t}\mathrm{d}t$$

and we arrive at the classical canonical transformation equations

$$\bar{p}_i = -\frac{\partial F}{\partial \bar{q}^i}, \qquad p_i = \frac{\partial F}{\partial q^i}, \qquad \bar{H} = H + \frac{\partial F}{\partial t}.$$

If $\bar{H} = 0$ then the solution of the Hamilton equations trivially of the form $\bar{q}^i = \text{const.}$, $\bar{p}_i = \text{const.}$ To find the function F for a transformation to this system we seek the general solution of the first-order partial differential equation known as the **Hamilton–Jacobi equation**,

$$\frac{\partial S(q^1, \ldots, q^n, c^1, \ldots, c^n, t)}{\partial t} + H\Big(q^1, \ldots, q^n, \frac{\partial S}{\partial q^1}, \ldots, \frac{\partial S}{\partial q^n}\Big) = 0, \tag{16.42}$$

and set $F(q^1, \ldots, q^n, \bar{q}^1, \ldots, \bar{q}^n) = S(q^1, \ldots, q^n, \bar{q}^1, \ldots, \bar{q}^n)$.

References

[1] R. W. R. Darling. *Differential Forms and Connections*. New York, Cambridge University Press, 1994.
[2] H. Flanders. *Differential Forms*. New York, Dover Publications, 1989.
[3] S. I. Goldberg. *Curvature and Homology*. New York, Academic Press, 1962.
[4] L. H. Loomis and S. Sternberg. *Advanced Calculus*. Reading, Mass., Addison-Wesley, 1968.
[5] M. Spivak. *Calculus on Manifolds*. New York, W. A. Benjamin, 1965.
[6] W. H. Chen, S. S. Chern and K. S. Lam. *Lectures on Differential Geometry*. Singapore, World Scientific, 1999.
[7] S. Sternberg. *Lectures on Differential Geometry*. Englewood Cliffs, N.J., Prentice-Hall, 1964.
[8] F. W. Warner. *Foundations of Differential Manifolds and Lie Groups*. New York, Springer-Verlag, 1983.
[9] C. de Witt-Morette, Y. Choquet-Bruhat and M. Dillard-Bleick. *Analysis, Manifolds and Physics*. Amsterdam, North-Holland, 1977.
[10] T. Frankel. *The Geometry of Physics*. New York, Cambridge University Press, 1997.
[11] R. Abraham. *Foundations of Mechanics*. New York, W. A. Benjamin, 1967.
[12] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. New York, Springer-Verlag, 1978.
[13] G. W. Mackey. *Mathematical Foundations of Quantum Mechanics*. New York, W. J. Benjamin, 1963.
[14] W. Thirring. *A Course in Mathematical Physics, Vol. 1: Classical Dynamical Systems*. New York, Springer-Verlag, 1978.
[15] F. P. Hildebrand *Methods of Applied Mathematics*. Englewood Cliffs, N.J., Prentice-Hall, 1965.