12

Lagrangian and Hamiltonian techniques

The Lagrangian formulation of mechanics is popular in practically all modern treatments of the subject. The ideas date back to the pioneering work of Euler, Lagrange and Hamilton, who showed how the equations of Newtonian dynamics could be derived from variational principles. In these, the evolution of a system is viewed as a path in some parameter space. The path the system follows is one which extremises a quantity called the *action*, which is the integral of the Lagrangian with respect to the evolution parameter (usually time). The mathematics behind this approach was clear from the outset, but a thorough physical understanding had to wait until the arrival of quantum theory. In the path-integral formulation of quantum mechanics a particle is viewed as simultaneously following all possible paths. By assigning a phase factor to the action for each path and summing these, one obtains the amplitude for a quantum process. The classical limit can then be understood as resulting from trajectories that reinforce the amplitude. In this manner classical trajectories emerge as those which make the action stationary.

A closely related idea is the Hamiltonian formulation of dynamics. The advantage of this approach is that it produces a set of first-order equations, making it well suited to numerical methods. The Hamiltonian approach also exposes the appropriate geometry for classical dynamical systems, which is a symplectic manifold. The Lagrangian and Hamiltonian formulations are well suited to studying the role of symmetry in physics. Any symmetry present in the Lagrangian will remain present in the equations of motion, and will produce a set of possible paths all related by the appropriate symmetry group. In this chapter we will touch on many of these ideas, and provide a number of Lagrangians for systems of physical interest. We also show how the method can be extended to the case of a multivector Lagrangian, which establishes contact with the systems studied in pseudoclassical mechanics.

12.1 The Euler–Lagrange equations

Suppose that a system is described by the multivector variables X_i, $i = 1, \ldots, n$. (The use of multivector variables makes this derivation slightly more general than usually seen.) The Lagrangian L is a scalar-valued function of X_i and $\dot{X}_i$, and possibly time, where the dot denotes the derivative with respect to time. The action for the system is

$$S = \int_{t_1}^{t_2} dt\, L(X_i, \dot{X}_i, t), \tag{12.1}$$

and we seek the equations for a path for which the action is stationary. The solution to this problem is standard application of variational calculus. We write

$$X_i(t) = X_i^0(t) + \epsilon Y_i(t), \tag{12.2}$$

where Y_i is a multivector containing the same grades as X_i and which vanishes at the endpoints, ϵ is a scalar, and X_i^0 represents the extremal path. It follows that the action must satisfy

$$\left.\frac{dS}{d\epsilon}\right|_{\epsilon=0} = 0, \tag{12.3}$$

in order to ensure that X_i^0 is a stationary solution. The chain rule now gives

$$\begin{aligned}\left.\frac{dS}{d\epsilon}\right|_{\epsilon=0} &= \int_{t_1}^{t_2} dt \sum_{i=1}^{n} \left(Y_i * \partial_{X_i} L + \dot{Y}_i * \partial_{\dot{X}_i} L\right) \\ &= \int_{t_1}^{t_2} dt \sum_{i=1}^{n} Y_i * \left(\partial_{X_i} L - \frac{d}{dt}(\partial_{\dot{X}_i} L)\right),\end{aligned} \tag{12.4}$$

where $A * B = \langle AB \rangle$. This integral must equal zero for all paths Y_i, from which we can read off the Euler–Lagrange equations in the form

$$\frac{\partial L}{\partial X_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{X}_i}\right) = 0, \qquad \forall i = 1, \ldots, n. \tag{12.5}$$

The multivector derivative ensures that there are as many equations as there are grades present in the X_i, which implies we have precisely the same number of equations as there are degrees of freedom in the system.

12.1.1 Symmetries and conservation laws

Suppose now that we consider a scalar-parameterised transformation of the dynamical variables, so that we have

$$X_i' = X_i'(X_i, \alpha). \tag{12.6}$$

We further assume that $\alpha = 0$ corresponds to the identity transformation (this restriction can be removed if necessary). The first-order change in X_i is denoted by δX_i, where

$$\delta X_i = \left.\frac{\partial X_i'}{\partial \alpha}\right|_{\alpha=0}. \tag{12.7}$$

We define the new Lagrangian

$$L'(X_i, \dot{X}_i) = L(X_i', \dot{X}_i'), \tag{12.8}$$

which is obtained from L simply by replacing each of the dynamical variables by their transformed equivalent. The chain rule now gives

$$\left.\frac{dL'}{d\alpha}\right|_{\alpha=0} = \sum_{i=1}^{n}\Big((\delta X_i)*\partial_{X_i}L + (\delta \dot{X}_i)*\partial_{\dot{X}_i'}L\Big). \tag{12.9}$$

If we now suppose that the X_i satisfy the Euler–Lagrange equations, we can rewrite the right-hand side as a total derivative to obtain

$$\left.\frac{dL'}{d\alpha}\right|_{\alpha=0} = \frac{d}{dt}\sum_{i=1}^{n}\Big((\delta X_i)*\partial_{\dot{X}_i}L\Big). \tag{12.10}$$

This result applies for any transformation, and can be used in a number of ways.

If the transformation is a symmetry of the Lagrangian, then L' is independent of α. In this case we immediately establish that a conjugate quantity is conserved. That is, symmetries of the Lagrangian produce conjugate conserved quantities. This is Noether's theorem, and it is valuable for extracting conserved quantities from dynamical systems. The fact that the derivation of equation (12.10) assumed the equations of motion were satisfied means that the quantity is conserved 'on-shell'. Some symmetries can also be extended 'off-shell', which becomes an important issue in quantum and supersymmetric systems.

An important application of equation (12.10) is to the case of time translation,

$$X_i'(t, \alpha) = X_i(t + \alpha), \tag{12.11}$$

so that

$$\left.\frac{\partial X_i'}{\partial \alpha}\right|_{\alpha=0} = \dot{X}_i. \tag{12.12}$$

If there is no explicit time dependence in the Lagrangian, then equation (12.10) gives

$$\frac{dL}{dt} = \frac{d}{dt}\sum_{i=1}^{n}\Big(\dot{X}_i*\partial_{\dot{X}_i}L\Big). \tag{12.13}$$

We therefore define the conserved *Hamiltonian* by

$$H = \sum_{i=1}^{n}\dot{X}_i*\partial_{\dot{X}_i}L - L. \tag{12.14}$$

This is more often written in terms of the *generalised momenta*

$$P_i = \partial_{\dot{X}_i} L, \tag{12.15}$$

so that

$$H = \sum_{i=1}^{n} \dot{X}_i * P_i - L. \tag{12.16}$$

The Hamiltonian gives the total energy in the system, and is conserved for systems with no explicit time dependence.

12.1.2 Point particle actions

The simplest application of the Lagrangian framework is for a particle moving in three dimensions in an external potential $V(\boldsymbol{x})$. The Lagrangian is the difference between the kinetic and potential energies,

$$L = \frac{m\boldsymbol{v}^2}{2} - V(\boldsymbol{x}), \tag{12.17}$$

where $\boldsymbol{v} = \dot{\boldsymbol{x}}$. The Euler–Lagrange equations give

$$m\dot{\boldsymbol{v}} = -\boldsymbol{\nabla} V, \tag{12.18}$$

which identifies $-\boldsymbol{\nabla} V$ with the force on a particle. The Hamiltonian is

$$H = \frac{\boldsymbol{p}^2}{2m} + V, \tag{12.19}$$

where $\boldsymbol{p} = m\boldsymbol{v}$. The Hamiltonian is conserved if V is independent of time.

The relativistic action for a free point particle raises some new issues. We begin with the simplest form of the action, which is

$$S = -m \int dt\, (1 - \dot{\boldsymbol{x}}^2)^{1/2}, \tag{12.20}$$

where the overdot denotes the derivative with respect to time t, and we work in units with $c = 1$. The momentum is

$$\boldsymbol{p} = \frac{\partial L}{\partial \dot{\boldsymbol{x}}} = \frac{m\dot{\boldsymbol{x}}}{(1 - \dot{\boldsymbol{x}}^2)^{1/2}}, \tag{12.21}$$

and the equations of motion state that $\boldsymbol{p}$ is constant. The Hamiltonian is

$$H = \boldsymbol{p}\cdot\dot{\boldsymbol{x}} - L = (\boldsymbol{p}^2 + m^2)^{1/2}, \tag{12.22}$$

and is also conserved.

The fact that the energy and momentum are dealt with differently is unsatisfactory from the point of view of Lorentz invariance, so we seek an alternative

formulation which is manifestly covariant. This can be achieved from the observation that the action is equivalent to

$$S = -m \int d\lambda \, (x' \cdot x')^{1/2}, \tag{12.23}$$

where $x' = \partial_\lambda x(\lambda)$. This integral is unchanged under a reparameterisation of the trajectory. By identifying λ with t we recover equation (12.20), and setting λ equal to the proper time τ we see that the action is $-m$ times the proper time along the path. Variation with respect to the relativistic position x now produces

$$\frac{d}{d\lambda}\left(\frac{mx'}{(x' \cdot x')^{1/2}}\right) = 0. \tag{12.24}$$

If we now set λ equal to the proper time the left-hand side becomes m times the relativistic acceleration $\dot{v}$, where overdots now denote the derivative with respect to proper time.

Interaction with an electromagnetic field is included through a term in $-qx' \cdot A$, producing the action

$$S = \int d\lambda \left(-m(x' \cdot x')^{1/2} - qx' \cdot A(x)\right). \tag{12.25}$$

Variation with respect to x now produces

$$-q\nabla A(x) \cdot x' + \frac{d}{d\lambda}\left(m\frac{x'}{(x' \cdot x')^{1/2}} + qA(x)\right) = 0. \tag{12.26}$$

Setting λ equal to the proper time, we find that

$$\begin{aligned} m\dot{v} &= q\big(\nabla A(x) \cdot v - v \cdot \nabla A(x)\big) \\ &= qF \cdot v, \end{aligned} \tag{12.27}$$

where $F = \nabla \wedge A$. We therefore recover the Lorentz force law, as discussed in section 5.5.3.

The square root in the free-particle action of equation (12.23) is often inconvenient, and can be removed by the inclusion of an *einbein*. This is a scalar function $e(\lambda)$, which has the transformation property under reparameterisations that

$$e(\nu) = \frac{d\lambda}{d\nu} e(\lambda), \tag{12.28}$$

where $\nu(\lambda)$ denotes a new parameterisation for the trajectory. The action can now be written in the equivalent form

$$S = -\tfrac{1}{2} \int d\lambda \left(e^{-1} x' \cdot x' + m^2 e\right). \tag{12.29}$$

Variation of e produces

$$e = \frac{(x' \cdot x')^{1/2}}{m}, \tag{12.30}$$

and substitution of this back into the action recovers equation (12.23). A first-order form of the action can also be developed by introducing the momentum p and writing

$$S = \int d\lambda \left(-p \cdot x' + \frac{e}{2}(p^2 - m^2) \right). \tag{12.31}$$

Variation of e produces the constraint equation $p^2 = m^2$, and variation of p produces $x' = ep$. This ensures that e is again given by equation (12.30). Finally, the x variation determines

$$p' = \frac{d}{d\lambda} \left(\frac{mx'}{(x' \cdot x')^{1/2}} \right) = 0, \tag{12.32}$$

recovering the desired equation. In each of these cases interaction with an electromagnetic field is included through a term in $-qx' \cdot A$. Moving to a reparameterisation-invariant formulation ensures that Lorentz covariance is manifest, but it limits the use of Hamiltonian techniques. Hamiltonians deal with energy, so picking out a Hamiltonian almost always implies breaking manifest Lorentz covariance.

12.1.3 Rigid-body dynamics

As a further application, consider a rigid body as discussed in section 3.4.3. The configuration of the body is described by the variables $\boldsymbol{x}_0(t)$ and $R(t)$, where $\boldsymbol{x}_0$ is the position of the centre of mass, and R is a three-dimensional rotor. We will ignore the motion of the centre of mass and concentrate on the rotational degrees of freedom. We also assume for simplicity that the object is freely rotating, so the Lagrangian is given by the rotational energy,

$$L = -\tfrac{1}{2} \Omega_B \cdot \mathcal{I}(\Omega_B). \tag{12.33}$$

Here $\mathcal{I}(B)$ is the inertia tensor, and

$$\Omega_B = -2R^\dagger \dot{R}, \tag{12.34}$$

where the dagger denotes the reverse operation in three dimensions.

The fact that the degrees of freedom are described by a rotor presents a slight problem. Rotors belong to a Lie group, and so form a group manifold. The Lagrangian is then a function defined for paths on the group manifold, which makes the Euler–Lagrange equations slightly more difficult to write down. There are two main methods of proceeding. The first is to introduce an explicit parameterisation of R, such as the Euler angles, and to compute the Lagrangian in terms

of these. This has the disadvantage of introducing a fixed coordinate system, making it difficult to assemble the final equations into a coordinate-free form. The structure of the rotor group provides a more elegant alternative. We replace the rotor R by an arbitrary even element (a spinor) ψ. The constraint $\psi\psi^\dagger = 1$ is enforced through the inclusion of a Lagrange multiplier. This method allows us to use the coordinate-free apparatus of multivector calculus in the variational principle and leads quickly to the full set of Euler equations.

Our Lagrangian is now

$$L(\psi, \dot{\psi}) = -\tfrac{1}{2}\Omega_B \cdot \mathcal{I}(\Omega_B) - \lambda(\psi\psi^\dagger - 1), \tag{12.35}$$

where the dynamical variable is the spinor ψ, and λ is a Lagrange multiplier. The bivector Ω_B is determined from ψ by

$$\Omega_B = -\psi^\dagger\dot{\psi} + \dot{\psi}^\dagger\psi, \tag{12.36}$$

which is a bivector, as required. The Euler–Lagrange equations reduce to the single multivector equation

$$\partial_\psi L - \frac{d}{dt}\big(\partial_{\dot{\psi}} L\big) = 0. \tag{12.37}$$

The symmetry of the inertia tensor simplifies the derivatives, and we obtain

$$\begin{aligned} \partial_\psi\big(-\tfrac{1}{2}\Omega_B \cdot \mathcal{I}(\Omega_B)\big) &= -2\mathcal{I}(\Omega_B)\dot{\psi}^\dagger, \\ \partial_{\dot{\psi}}\big(-\tfrac{1}{2}\Omega_B \cdot \mathcal{I}(\Omega_B)\big) &= 2\mathcal{I}(\Omega_B)\psi^\dagger, \end{aligned} \tag{12.38}$$

where we have used the results of section 11.1. After reversing, the Euler–Lagrange equation for ψ is simply

$$\frac{d}{dt}\big(\psi\mathcal{I}(\Omega_B)\big) + \dot{\psi}\mathcal{I}(\Omega_B) = \lambda\psi. \tag{12.39}$$

Variation with respect to the Lagrange multiplier λ enforces the constraint that $\psi\psi^\dagger = 1$, which means we can now replace ψ with the rotor R. We therefore arrive at the equation

$$\mathcal{I}(\dot{\Omega}_B) - \Omega_B\mathcal{I}(\Omega_B) = \lambda. \tag{12.40}$$

The scalar part of this equation determines λ and shows that, in the absence of any applied couple, the rotational energy is a constant of the motion. The bivector part of equation (12.40) recovers the familiar equation

$$\mathcal{I}(\dot{\Omega}_B) - \Omega_B \times \mathcal{I}(\Omega_B) = 0, \tag{12.41}$$

as found in section 3.4.3. The Lagrange multiplier has avoided any need for handling the rotor group manifold.

12.2 Classical models for spin-1/2 particles

The use of non-relativistic spinors in describing the dynamics of a rigid body demonstrates that spinors are not necessarily restricted to applications in quantum mechanics. This is significant in addressing the question: what is the classical analogue of the Dirac equation? That is, what classical dynamical system produces the Dirac equation on quantisation? There have been many attempts to answer this question, and in the following sections we investigate two of them.

12.2.1 Rotor dynamics

For our first classical model of a fermion, we start with the Lagrangian for the Dirac field. Following the notation of section 8.2 this is

$$L_{Dirac} = \langle \nabla\psi I\gamma_3\tilde{\psi} - m\psi\tilde{\psi}\rangle. \tag{12.42}$$

The properties of this Lagrangian are studied in detail in chapter 13. Focusing on the first (kinetic) term, we can write this as

$$\langle \nabla\psi I\gamma_3\tilde{\psi}\rangle = \langle \nabla\psi I\boldsymbol{\sigma}_3\psi^{-1}\psi\gamma_0\tilde{\psi}\rangle = \langle J\nabla\psi I\boldsymbol{\sigma}_3\psi^{-1}\rangle, \tag{12.43}$$

where $J = \psi\gamma_0\tilde{\psi}$ is the Dirac current. The streamlines of J describe how the probability density flows through spacetime. To reduce to a point-particle model, we assume that only the derivatives along a streamline are important and that the density is concentrated entirely on one streamline. This streamline is then identified with the particle worldline, and the kinetic term becomes

$$\langle J\cdot\nabla\psi I\boldsymbol{\sigma}_3\psi^{-1}\rangle = \langle \psi' I\boldsymbol{\sigma}_3\psi^{-1}\rangle, \tag{12.44}$$

where the prime denotes the derivative with respect to some parameter along the worldline. Now recall from section 8.2 that a Dirac spinor decomposes into

$$\psi = \rho^{1/2}\mathrm{e}^{I\beta/2}R, \tag{12.45}$$

where ρ and β are scalars, and R is a Lorentz rotor (a member of the connected subgroup of the spin group). The inverse, ψ^{-1}, is therefore

$$\psi^{-1} = \rho^{-1/2}\mathrm{e}^{-I\beta/2}\tilde{R}. \tag{12.46}$$

Substituting this parameterisation into equation (12.44), we find that

$$\langle \psi' I\boldsymbol{\sigma}_3\psi^{-1}\rangle = \langle R' I\boldsymbol{\sigma}_3\tilde{R}\rangle. \tag{12.47}$$

The dynamics are now parameterised by a Lorentz rotor, as opposed to a full spinor. Given that the magnitude of a spinor is related to the quantum concept of probability density, it is sensible that the classical model should only depend on the rotor component.

To complete the model we need to impose the condition that the current $\psi\gamma_0\tilde{\psi}$

defines the tangent to the worldline. This is achieved by including a Lagrange multiplier to enforce the constraint that

$$x' = eR\gamma_0\tilde{R}, \tag{12.48}$$

where e is an einbein. Finally, the mass term $m\psi\tilde{\psi}$ becomes simply em, where again the einbein ensures reparameterisation invariance. The full Lagrangian is now

$$L(x, x', R, R', p, e) = \langle R' I\boldsymbol{\sigma}_3\tilde{R} - p(x' - eR\gamma_0\tilde{R}) - em\rangle, \tag{12.49}$$

and the action is formed by integrating this with respect to the evolution parameter λ. The p equation returns the constraint of equation (12.48), and the einbein e returns

$$p\cdot(R\gamma_0\tilde{R}) = m. \tag{12.50}$$

After variation we can choose the parameterisation such that $e = 1$, and x' is replaced by $\dot{x}$, with dots denoting the derivative with respect to proper time along the worldline $x(\tau)$. It follows that $p \cdot \dot{x} = m$. Clearly, then, we can identify p with the momentum. The x variation then says that the momentum is constant.

The final equation requires varying R, which lies on the group manifold of the rotor group $\mathrm{Spin}^+(1,3)$. This variation can be performed in a number of ways. We could extend the technique employed for rigid-body mechanics, and relax the normalisation constraint so that R becomes a full spinor. The normalisation is then enforced by a pair of Lagrange multipliers (one each for the scalar and pseudoscalar terms). However, we can avoid this by returning to the original form of the Lagrangian in terms of ψ and replacing the relevant terms by

$$\langle R' I\boldsymbol{\sigma}_3\tilde{R} + epR\gamma_0\tilde{R}\rangle = \langle\psi' I\boldsymbol{\sigma}_3\psi^{-1} + ep\psi\gamma_0\tilde{\psi}/\rho\rangle, \tag{12.51}$$

where $\rho = |\psi\gamma_0\tilde{\psi}|$. This form ensures that L is only dependent on the rotor component of ψ, but still allows us to vary L with respect to ψ. This is easier than constructing the derivative on the group manifold. To proceed we need a pair of additional results. The first is that

$$\partial_\psi\langle M\psi^{-1}\rangle = -\psi^{-1}M\psi^{-1}, \tag{12.52}$$

which holds for any even multivector M. The second is that

$$2\rho\partial_\psi\rho = \partial_\psi\langle\psi\gamma_0\tilde{\psi}\psi\gamma_0\tilde{\psi}\rangle = 4\gamma_0\tilde{\psi}\psi\gamma_0\tilde{\psi}, \tag{12.53}$$

which implies that

$$\partial_\psi\rho = 2\rho\psi^{-1}. \tag{12.54}$$

The ψ variation now produces (after setting λ equal to proper time τ)

$$-\psi^{-1}\dot{\psi}I\boldsymbol{\sigma}_3\psi^{-1} + \frac{e}{\rho}(2\gamma_0\tilde{\psi}p - 2\psi^{-1}\langle p\psi\gamma_0\tilde{\psi}\rangle) - \frac{d}{d\tau}(I\boldsymbol{\sigma}_3\psi^{-1}) = 0. \tag{12.55}$$

On multiplying through by ψ we obtain

$$\dot{S} + 2p \wedge \dot{x} = 0, \tag{12.56}$$

where

$$S = \psi I\boldsymbol{\sigma}_3\psi^{-1} = RI\boldsymbol{\sigma}_3\tilde{R}. \tag{12.57}$$

The rotor variation therefore produces an equation which states that the total angular momentum is conserved. This shows that the classical model has many of the desired features. Linear momentum is conserved, and the spin-1/2 nature of the particle is captured in the total angular momentum.

The simplest solution to the equations of motion has $m\dot{x} = p$, so that the particle is at rest in the p frame. The spin bivector is also constant, as one would expect in the absence of interaction. There are a range of further solutions, however, which are of interest. Suppose that we align γ_0 with momentum, and write

$$p = \frac{m}{\cosh(\alpha)}\gamma_0 = m^*\gamma_0, \tag{12.58}$$

which defines the 'effective mass' m^*. The equations of motion are then solved by

$$\begin{aligned} R &= \mathrm{e}^{I\boldsymbol{\sigma}_3 m^*\tau}\mathrm{e}^{\alpha\boldsymbol{\sigma}_2/2}, \\ x &= \tau\cosh(\alpha)\,\gamma_0 - \frac{\sinh(\alpha)}{2m^*}\gamma_1\mathrm{e}^{-2I\boldsymbol{\sigma}_3 m^*\tau}. \end{aligned} \tag{12.59}$$

The total angular momentum is

$$\tfrac{1}{2}S + p\wedge x = \tfrac{1}{2}\cosh(\alpha)\, I\boldsymbol{\sigma}_3, \tag{12.60}$$

which is constant. This solution describes a particle rotating at angular frequency $2m/\cosh(\alpha)$ (as measured by the proper time), and with a radius of

$$r_0 = \frac{1}{2m}\sinh(\alpha)\cosh(\alpha). \tag{12.61}$$

As α increases, the momentum goes 'off-shell', and the particle can 'borrow' energy to execute a circular motion and feel out its surroundings. This model therefore captures some aspects of fermionic quantum mechanics, exhibiting a form of zitterbewegung, while still describing a point-particle trajectory.

For many applications the model constructed here is unnecessarily complicated, and we instead choose to work with the somewhat simpler Lagrangian

$$L(x, x', \psi, \psi', p, e) = \langle \psi' I\boldsymbol{\sigma}_3\tilde{\psi} - p(x' - e\psi\gamma_0\tilde{\psi}) - em\psi\tilde{\psi}\rangle. \tag{12.62}$$

Global phase invariance of L ensures that $\langle\psi\tilde{\psi}\rangle$ is constant and can be set to 1. If the initial conditions are chosen suitably, one can also show that the 4-vector part of $\psi\tilde{\psi}$ remains zero, and the motion reduces to that of the previous model. An open question is whether either of these models produces the Dirac equation on

quantisation. The problem is that a path-integral quantisation involves the group manifold of $\mathrm{Spin}^+(1,3)$, which is non-compact. In addition, the Lagrangian is first order, which can give rise to complications in the path integral.

A deficiency of the classical model is exposed when we couple the particle to the electromagnetic field. If we consider the phase transformation

$$R \mapsto R\mathrm{e}^{I\boldsymbol{\sigma}_3\phi}, \tag{12.63}$$

then this introduces a term going as $-\partial_\lambda\phi = -x'\cdot(\nabla\phi)$ into the Lagrangian. Local phase invariance is therefore restored by modifying the Lagrangian to

$$L(x,x',R,R',p,e) = \langle R'I\boldsymbol{\sigma}_3\tilde{R} - p(x' - eR\gamma_0\tilde{R}) - qx'\cdot A - em\rangle, \tag{12.64}$$

where A is the electromagnetic vector potential. The $qx'\cdot A$ term is the natural point-particle equivalent of the interaction term $qJ\cdot A$ in the Dirac Lagrangian. Variation now modifies the p equation in the expected manner to read

$$\dot{p} = qF\cdot\dot{x}. \tag{12.65}$$

But the spin equation is not affected — we do not naturally pick up the $g = 2$ behaviour for the gyromagnetic ratio of a spin-1/2 particle. This is disappointing, given that the A term is all that is required to guarantee that $g = 2$ in Dirac theory. The problem can be rectified by introducing a further term into the Lagrangian, going as

$$L_g = \left\langle -\frac{q}{2m}FRI\boldsymbol{\sigma}_3\tilde{R}\right\rangle = \left\langle -\frac{q}{2m}F\psi I\boldsymbol{\sigma}_3\psi^{-1}\right\rangle. \tag{12.66}$$

This modifies both the R and p equations to give

$$\begin{aligned}\dot{S} &= 2\dot{x}\wedge p + \frac{q}{m}F\times S,\\ \dot{p} &= qF\cdot\dot{x} + \frac{q}{2m}\nabla F(x)\cdot S.\end{aligned} \tag{12.67}$$

These equations have the expected form for a particle with $g = 2$, but the value of the gyromagnetic ratio has been put in by hand.

12.2.2 Pseudoclassical mechanics

A quite different approach to the classical mechanics of a spin-1/2 particle is provided by pseudoclassical mechanics, which introduces the interesting new concept of a multivector-valued Lagrangian. We only consider the simplest case of a non-relativistic model. The model is motivated by the idea that the spin operators satisfy

$$\hat{s}_i\hat{s}_j + \hat{s}_j\hat{s}_i = \frac{\hbar^2}{2}\delta_{ij}. \tag{12.68}$$

The classical analogue of these relations should have zero on the right-hand side, so the particle is described by a set of anticommuting Grassmann variables.

This argument runs contrary to the viewpoint of this book, which is that there is nothing at all quantum-mechanical about a Clifford algebra, but the model itself is interesting. We introduce a set of three Grassmann variables $\{\zeta_i\}$ and define the Lagrangian

$$L = \tfrac{1}{2}\zeta_i\dot{\zeta}_i - \tfrac{1}{2}\epsilon_{ijk}\omega_i\zeta_j\zeta_k, \tag{12.69}$$

where the ω_i are constants. Following the prescription of section 11.2 we replace the Grassmann variables with a set of three vectors $\{e_i\}$ under the exterior product. The Lagrangian then becomes

$$L = \tfrac{1}{2}e_i\wedge\dot{e}_i - \omega, \tag{12.70}$$

where

$$\omega = \tfrac{1}{2}\epsilon_{ijk}\omega_i e_j e_k = \omega_1(e_2\wedge e_3) + \omega_2(e_3\wedge e_1) + \omega_3(e_1\wedge e_2). \tag{12.71}$$

The Lagrangian is now a *bivector*, and not simply a scalar. This raises an immediate question — how can the variational principle be applied to a multivector? The answer is that all components of the Lagrangian must remain stationary under variation. Suppose that we contract L with an arbitrary bivector B to form the scalar $\langle LB\rangle$. Variation of this produces the Euler–Lagrange equation

$$\partial_{e_i}\langle LB\rangle - \frac{d}{dt}(\partial_{\dot{e}_i}\langle LB\rangle) = 0. \tag{12.72}$$

Treating the $\{e_i\}$ as vector variables, we arrive at the equation

$$(\dot{e}_i + \epsilon_{ijk}\omega_j e_k)\cdot B = 0. \tag{12.73}$$

But we must demand that this vanishes for all possible B, from which we extract the equation

$$\dot{e}_i + \epsilon_{ijk}\omega_j e_k = 0. \tag{12.74}$$

This is the general method for handling multivector Lagrangians. The contraction with any constant multivector must result in a scalar Lagrangian which is stationary when the equations of motion are satisfied. Equation (12.73) illustrates a further feature. For a fixed B, equation (12.73) is not sufficient to extract the full set of equations. It is only by allowing B to vary, and hence treat the Lagrangian as a bivector, that the full equations are extracted.

To solve equation (12.74) we first establish that ω is constant,

$$\dot{\omega} = 0, \tag{12.75}$$

which follows immediately from the equation of motion. Next we introduce the reciprocal frame $\{e^i\}$ and write the equation of motion in the form

$$\dot{e}_i = e^i\cdot\omega. \tag{12.76}$$

Now suppose that we define the symmetric function g by

$$\mathsf{g}(a) = \sum_{i=1}^{3} a\cdot e_i\, e_i, \tag{12.77}$$

so that $\mathsf{g}(e^i) = e_i$. The function g is a form of metric for the non-orthonormal frame e_i. On differentiating $\mathsf{g}(a)$, holding a constant, we find that

$$\frac{d}{dt}\mathsf{g}(a) = \sum_{i=1}^{3}\big(a\cdot(e^i\cdot\omega)e_i + a\cdot e_i\, e^i\cdot\omega\big) = \omega\cdot a + a\cdot\omega = 0. \tag{12.78}$$

It follows that the function g is constant, even though the e_i vectors vary in time. The motion is found by introducing the square root of g, which satisfies

$$\bar{\mathsf{h}}\mathsf{h}(a) = \mathsf{g}(a), \qquad \bar{\mathsf{h}} = \mathsf{h}. \tag{12.79}$$

This function is found by diagonalising g and taking the square root of the eigenvalues. It follows that

$$\delta^j_i = e_i\cdot e^j = \mathsf{g}(e^i)\cdot e^j = \mathsf{h}(e^i)\cdot\mathsf{h}(e^j). \tag{12.80}$$

The vectors $\mathsf{h}(e^i)$ are therefore orthonormal, so we write

$$f_i = \mathsf{h}(e^i), \qquad f_i\cdot f_j = \delta_{ij}. \tag{12.81}$$

These vectors satisfy

$$\dot{f}_i = f_i\cdot\Omega, \tag{12.82}$$

where

$$\Omega = \omega_1 f_2 f_3 + \omega_2 f_3 f_1 + \omega_3 f_1 f_2. \tag{12.83}$$

Since $\mathsf{h}(\Omega) = \omega$, we see that Ω is a constant bivector. It follows that the $\{f_i\}$ frame simply rotates at a constant frequency in the Ω plane. The solution for the e_i vectors is therefore

$$e_i(t) = \mathsf{h}^{-1}\big(\mathrm{e}^{-\Omega t/2} f_i(0)\mathrm{e}^{\Omega t/2}\big). \tag{12.84}$$

The only motion taking place in this system is that a fixed set of orthonormal vectors is rotating in a constant plane, and the resulting frame is then distorted by a constant symmetric function. A simple picture of this type is fairly typical of pseudoclassical systems when analysed in this manner.

12.3 Hamiltonian techniques

The Hamiltonian formulation of mechanics is important in a range of applications, not least because of its superior handling of numerical issues. We start by forming Hamilton's equations in local coordinates, before placing Hamiltonian dynamics in a more geometric setting. Suppose that a dynamical system is

described in terms of a Lagrangian $L(q_i, \dot{q}_i, t)$, where the $\{q_i\}$ are a set of n coordinates for configuration space. The Euler–Lagrange equations are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i}. \tag{12.85}$$

These equations typically result in a set of n second-order equations that relate the generalised momenta to the forces in the system. The Euler–Lagrange equations are equivalent to the set of $2n$ first-order equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{12.86}$$

These are Hamilton's equations. The Hamiltonian $H(q_i, p_i, t)$ is given by

$$H(q_i, p_i, t) = \sum_{i=1}^{n} p_i \dot{q}_i - L(q_i, \dot{q}_i, t) \tag{12.87}$$

in which the $\dot{q}_i$ are expressed in terms of the p_i by inverting the equations

$$p_i = \frac{\partial L}{\partial \dot{q}_i}. \tag{12.88}$$

The transformation from a Lagrangian to a Hamiltonian framework is called a *Legendre transformation*. We move from considering dynamics in n-dimensional configuration space to a $2n$-dimensional *phase space*.

If the Hamiltonian is independent of time we can immediately see that it is conserved. That is, H gives the conserved energy in the system. The proof is straightforward:

$$\frac{dH}{dt} = \sum_{i=1}^{n}\left(\dot{q}_i \frac{\partial H}{\partial q_i} + \dot{p}_i \frac{\partial H}{\partial p_i}\right) = 0. \tag{12.89}$$

Phase space provides a very useful way of analysing the motion and stability of complicated systems. As a simple example, consider a pendulum consisting of a mass m attached to a rigid rod of length a. The configuration of the system is described by a single angle θ, and the Lagrangian is

$$L = \frac{ma^2\dot{\theta}^2}{2} + mga\cos(\theta). \tag{12.90}$$

The Hamiltonian is therefore

$$H = \frac{p_\theta^2}{2ma^2} - mga\cos(\theta), \tag{12.91}$$

and this is conserved. The trajectories of the system can be visualised in terms of a phase-space portrait, which plots surfaces of constant H in phase space. Sample trajectories are shown in figure 12.1. The figure illustrates how the phase portrait can capture global aspects of the system, such as the behaviour of the system as the energy gets close to value for which the pendulum can complete a full loop.

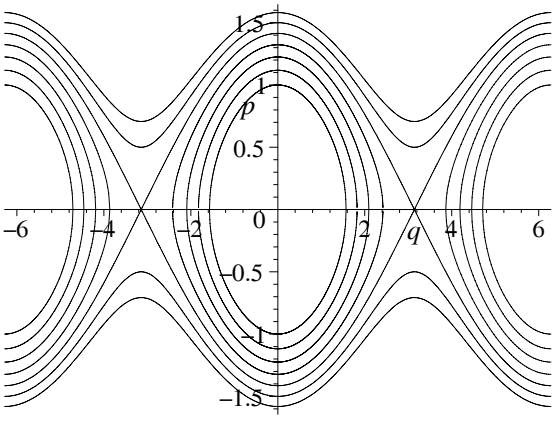

Figure 12.1 *A phase portrait.* The q coordinate represents the angle and p is the canonical momentum. The Hamiltonian is $p^2 - \cos(q)$. As H approaches 1 a bifurcation appears, corresponding to the energy for which the pendulum can complete a loop. The system is periodic, so the phase portrait can be thought of as wrapping up into a cylinder

12.3.1 Symplectic geometry

The natural setting for Hamilton's equations is provided by *symplectic geometry*. A symplectic manifold (M, Ω) consists of a $2n$-dimensional manifold M together with a closed, non-degenerate 2-form Ω. We will assume that n is finite so as to avoid discussion of the technicalities of infinite-dimensional spaces. We can analyse this structure using the apparatus of vector manifolds, described in section 6.5. A symplectic manifold does not have a metric structure, so we must take care not to employ the metric induced in the vector manifold by its embedding. This means we must distinguish tangent and cotangent spaces, as we can only apply the inner product between tangent and cotangent vectors. We denote the tangent space at x by $T_x M$, and the cotangent space $T_x^* M$.

The covariant vector derivative is denoted by ∇, and always results in a multivector that is intrinsic to the manifold (in section 6.5.3 this derivative was denoted D). The 2-form Ω is a bivector field evaluated in the cotangent space. The statement that Ω is closed is simply

$$\nabla \wedge \Omega = 0. \tag{12.92}$$

This is required in order that the Poisson bracket satisfies the Jacobi identity.

The condition that Ω is non-degenerate is simply that

$$\Omega(x)\cdot a \neq 0, \qquad \forall a \neq 0,\ a \in T_x M. \tag{12.93}$$

If we view $\Omega(x)\cdot a$ as a linear map from $T_x M$ to $T_x^* M$, then Ω being non-degenerate implies that the map has non-zero determinant, so is invertible. The inverse map is generated by a second bivector, which we label J. This second bivector lies in the tangent space, and can be viewed as the inverse of Ω. The two bivectors are related by the pair of equations

$$\begin{aligned} J\cdot(\Omega\cdot a) &= a, \quad \forall a \in T_x M, \\ \Omega\cdot(J\cdot a^*) &= a^*, \quad \forall a^* \in T_x^* M. \end{aligned} \tag{12.94}$$

The properties of Ω and J can be understood simply by introducing a set of local coordinates (p^i, q^i) over M. In terms of these we define the tangent vectors

$$e_i = \frac{\partial x}{\partial p^i}, \qquad f_i = \frac{\partial x}{\partial q^i}, \tag{12.95}$$

and the cotangent vectors

$$e^i = \nabla p^i, \qquad f^i = \nabla q^i. \tag{12.96}$$

We then set

$$\Omega = \sum_{i=1}^{n} f^i \wedge e^i, \qquad J = \sum_{i=1}^{n} e_i \wedge f_i. \tag{12.97}$$

So J and Ω both have a similar structure to the complex bivector introduced in section 11.4. By construction, Ω is clearly closed. It is also straightforward to verify the relations

$$\begin{aligned} \Omega\cdot e_i &= f^i, & \Omega\cdot f_i &= -e^i, \\ J\cdot e^i &= -f_i, & J\cdot f^i &= e_i, \end{aligned} \tag{12.98}$$

which confirm that equations (12.94) are satisfied.

The Hamiltonian $H(x,t)$ is a scalar field defined over M, and the dynamics of the system are governed by the equation

$$\dot{x} = (\nabla H)\cdot J. \tag{12.99}$$

This is an identity between tangent vectors. In terms of local coordinates this equation becomes

$$\dot{p}^i e_i + \dot{q}^i f_i = \left(e^i \frac{\partial H}{\partial p^i} + f^i \frac{\partial H}{\partial q^i} \right) \cdot J = f_i \frac{\partial H}{\partial p^i} - e_i \frac{\partial H}{\partial q^i}, \tag{12.100}$$

where repeated indices are summed over $1, \ldots, n$. We therefore recover Hamilton's equations in local coordinates.

12.3.2 Conservation theorems and the Poisson bracket

We now restrict to the case where H is independent of time t. Suppose that a scalar function $f(x)$ is defined over phase space. The evolution of this along a phase space trajectory $x(t)$ is determined by

$$\dot{f} = \dot{x}\cdot\nabla f = (\nabla f \wedge \nabla H)\cdot J. \tag{12.101}$$

It follows immediately that $\dot{H} = 0$. A further consequence follows if H is invariant along some direction a in phase space. If we form the directional derivative of H we obtain

$$a\cdot\nabla H = \big(J\cdot(\Omega\cdot a)\big)\cdot\nabla H = (\Omega\cdot a)\cdot\dot{x}. \tag{12.102}$$

So if H is unchanged in the a direction we have

$$(\Omega\cdot a)\cdot\dot{x} = 0, \tag{12.103}$$

so all flows are perpendicular to the cotangent vector $\Omega\cdot a$.

The equation for the evolution of f leads naturally to the definition of the Poisson bracket of a Hamiltonian system. Given two scalar fields f and g the Poisson bracket is defined by

$$\{f(x), g(x)\} = (\nabla f \wedge \nabla g)\cdot J. \tag{12.104}$$

In terms of local coordinates this takes the more familiar form

$$\{f(x), g(x)\} = \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \right). \tag{12.105}$$

The geometric form neatly brings out the antisymmetry of the Poisson bracket. It follows, for example, that the Poisson bracket with the Hamiltonian returns the time development of a scalar field:

$$\{f, H\} = (\nabla f \wedge \nabla H)\cdot J = \dot{f}. \tag{12.106}$$

Poisson brackets and the Hamiltonian formulation of dynamics provide a natural route through to quantum mechanics, where Poisson brackets are replaced by operator commutation relations.

An important property satisfied by the Poisson bracket is the Jacobi identity

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0, \tag{12.107}$$

which is easily confirmed in terms of local coordinates. This identity links the Poisson bracket structure to a Lie algebra structure. The identity is satisfied by any symplectic manifold, as we now establish. We first write

$$\{f, g\} = \big(\nabla\wedge(f\nabla g)\big)\cdot J = (J\cdot\nabla)\cdot(f\nabla g), \tag{12.108}$$

which follows from the identity that $\nabla\wedge\nabla = 0$ (every exact form is closed). The Jacobi identity now reduces to

$$(J\cdot\nabla)\cdot\big((\nabla f\wedge\nabla g)\cdot J\,\nabla h + \text{cyclic permutations}\big) = 0. \tag{12.109}$$

If we define

$$T = \nabla f\wedge\nabla g\wedge\nabla h, \tag{12.110}$$

then equation (12.109) simplifies to

$$(J\cdot\nabla)\cdot(J\cdot T) = 0. \tag{12.111}$$

To simplify this equation further we employ the identity

$$(B\wedge B)\cdot a = 2B\wedge(B\cdot a), \tag{12.112}$$

which holds for any bivector B and vector a. We can now write

$$(J\cdot\nabla)\cdot(J\cdot T) = (J\cdot\dot\nabla)\cdot(\dot J\cdot T) + \tfrac{1}{2}(J\wedge J)\cdot(\nabla\wedge T) \tag{12.113}$$

and the final term here vanishes as T is exact. It follows that the Jacobi identity reduces to the condition

$$(J\cdot\nabla)\wedge J = 0. \tag{12.114}$$

The final task is to demonstrate that this identity for J is equivalent to the statement that Ω is closed. Equation (12.114) is evaluated entirely in tangent space. If we use Ω to map each term into the cotangent space we arrive at the equivalent expression

$$e^\alpha\wedge e^\beta\wedge\dot\nabla\langle\dot J(\Omega\cdot e_\alpha)\wedge(\Omega\cdot e_\beta)\rangle = 0, \tag{12.115}$$

where Greek indices are summed from 1 to $2n$, with the first n covering the e_i frame, and the second n covering the f_i frame. This identity is equivalent to

$$\nabla\wedge(e^\alpha\wedge e^\beta\langle J(\Omega\cdot e_\alpha)\wedge(\Omega\cdot e_\beta)\rangle) - e^\alpha\wedge e^\beta\wedge\nabla\langle\check J(\Omega\cdot e_\alpha)\wedge(\Omega\cdot e_\beta)\rangle = 0, \tag{12.116}$$

where the check denotes that J is not differentiated in the second expression. The frame derivatives in this expression can all be shown to vanish, which leaves

$$2\nabla\wedge\Omega - e^\alpha\wedge e^\beta\wedge\dot\nabla\langle J(\dot\Omega\cdot e_\alpha)\wedge(\Omega\cdot e_\beta) + J(\Omega\cdot e_\alpha)\wedge(\dot\Omega\cdot e_\beta)\rangle = 0. \tag{12.117}$$

This is equivalent to

$$-2\nabla\wedge\Omega = 0, \tag{12.118}$$

which proves the main result. Any symplectic manifold admits a Poisson bracket structure that satisfies the Jacobi identity. As such, any symplectic manifold can form the basis for a Hamiltonian system.

12.3.3 The cotangent bundle

In practice, the phase space for a Hamiltonian system is often the cotangent bundle of configuration space. This works as follows. Suppose that Q denotes configuration space. It is a manifold with potentially non-trivial topology. At each point q in the manifold we define the cotangent space T_q^*Q. If q^i form a set of local coordinates in configuration space, then ∇q^i define a set of basis vectors for T_q^*Q. An arbitrary cotangent vector can be written as $p_i \nabla q^i$, so the p_i can be used as coordinates for T_q^*Q. Now consider the bundle of all tangent spaces, T^*Q. This is a manifold, and a general point in T^*Q is specified by the set of $2n$ coordinates (q^i, p_i). The first n of these locate the position over Q, and the second n locate a point within the cotangent space. The cotangent bundle T^*Q is a symplectic manifold, with the symplectic structure defined by

$$\Omega = \sum_{i=1}^{n} \nabla q^i \wedge \nabla p_i. \tag{12.119}$$

The reason this structure often arises is that, while there may be constraints placed on configuration space, there are usually no restrictions in momentum space. Returning to the case of the simple pendulum, Q is a circle since θ is a periodic coordinate. But there are no such constraints on $\dot{\theta}$, so the cotangent space is a line. The manifold T^*Q can therefore be visualised as a cylinder, which is the phase space for the pendulum.

12.3.4 Canonical transformations

Suppose that (M_1, Ω_1) and (M_2, Ω_2) are two symplectic manifolds. We let f denote a map from M_1 to M_2, which we write as

$$x' = f(x), \quad x \in M_1, \quad x' \in M_2. \tag{12.120}$$

This map is canonical if it respects the symplectic structure. That is, we must have

$$\bar{\mathsf{f}}(\Omega_2) = \Omega_1, \qquad \mathsf{f}(J_1) = J_2. \tag{12.121}$$

Here

$$\mathsf{f}(a) = a \cdot \nabla f(x) \tag{12.122}$$

is a map from $T_x M_1$ to $T_{x'} M_2$. The fact that Ω is non-degenerate means that we can define a volume form on either manifold by

$$V = \frac{1}{n!} \langle (\Omega)^n \rangle_{2n} = \frac{1}{n!} \langle \Omega \wedge \Omega \wedge \cdots \wedge \Omega \rangle_{2n}. \tag{12.123}$$

The map $\bar{\mathsf{f}}$ must preserve this volume form, so has non-zero determinant. It follows that f is invertible for a canonical transformation, and hence so too is f.

As a check, the Poisson bracket structure remains intact as

$$J_1 \cdot (\nabla f \wedge \nabla g) = J_2 \cdot \bar{\mathsf{f}}^{-1}(\nabla f \wedge \nabla g) = J_2 \cdot (\nabla_2 f \wedge \nabla_2 g), \tag{12.124}$$

where $\nabla_2 = \bar{\mathsf{f}}^{-1}(\nabla)$ is the vector derivative on M_2. It follows that the dynamics can be formulated on either M_1 or M_2, and the physical results will remain unchanged. This is potentially a very powerful result. The set of all possible symplectic transformations is large, and there may well by a suitable transformation which can dramatically simplify the dynamics. This is particularly evident when one notices that symplectic transformations can mix up the position and momentum coordinates in one space. These transformations are richer than simply converting between configuration spaces.

In some applications, phase space is simply R^{2n}, and the bivector J is constant. In this case we can consider canonical transformations which map phase space onto itself. For these the map f is canonical if and only if

$$\mathsf{f}(J) = J. \tag{12.125}$$

Linear transformations satisfying this identity define the *symplectic group*. This can be analysed using the spin group approach developed in section 11.4.

12.4 Lagrangian field theory

The Lagrangian approach to classical dynamics extends to field theory, which can be viewed as the dynamics of systems with an infinite number of degrees of freedom. There are some technical issues connected with the infinite size of the configuration space, but we will not discuss these here. Suppose that the system of interest depends on a field $\psi(x)$, where for simplicity we will assume that x is a spacetime vector. This does not restrict us to relativistic theories, as there is no need to restrict the Lagrangian to be Lorentz-invariant. The action is now defined as an integral over a region of spacetime by

$$S = \int d^4x \, \mathcal{L}(\psi, \partial_\mu \psi, x), \tag{12.126}$$

where $\mathcal{L}$ is the *Lagrangian density* and x^μ are a set of fixed orthonormal coordinates for spacetime. More general coordinate systems are easily accommodated with the inclusion of suitable factors of the Jacobian.

The derivation of the Euler–Lagrange equations proceeds precisely as in section 12.1. We assume that $\psi_0(x)$ represents the extremal path, satisfying the desired boundary conditions, and look for variations of the form

$$\psi(x) = \psi_0(x) + \epsilon\phi(x). \tag{12.127}$$

Here $\phi(x)$ is a field of the same form as $\psi(x)$, which vanishes over the boundary.

The first-order variation in the action is (summation convention in force)

$$\left.\frac{dS}{d\epsilon}\right|_{\epsilon=0} = \int d^4x \left(\phi(x) * \frac{\partial \mathcal{L}}{\partial \psi} + \frac{\partial \phi}{\partial x^\mu} * \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)}\right). \tag{12.128}$$

The final term is integrated by parts, and the boundary term vanishes. We therefore find that

$$\left.\frac{dS}{d\epsilon}\right|_{\epsilon=0} = \int d^4x\, \phi(x) * \left(\frac{\partial \mathcal{L}}{\partial \psi} - \frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)}\right)\right), \tag{12.129}$$

from which we can read off the variational equations as

$$\frac{\partial \mathcal{L}}{\partial \psi} - \frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)}\right) = 0. \tag{12.130}$$

If more fields are present we obtain an equation of this form for each field. Our main applications of these equations are in the following chapters, where we discuss gauge theories and gravitation. Here we illustrate the equations with a pair of examples concerned with elastic and fluid materials.

12.4.1 Hyperelastic materials

The equations of continuum mechanics, which govern an elastic body, were derived in section 6.6. For certain elastic materials it is possible to obtain these equations from a variational principle. We follow the notation of section 6.6, so f is the displacement field, $\mathsf{f}(\boldsymbol{a})$ is the directional derivative of f, and $\mathsf{C} = \bar{\mathsf{f}}\mathsf{f}$ is the Cauchy–Green tensor. The materials of interest here are called *hyperelastic*. These are defined by the property that, in the absence of external fields, their internal energy U is a function of C only. A suitable action for this system is

$$S = \int dt\, d^3x \left(\frac{\rho(x)}{2}\dot{f}^2 - U(\partial_i f)\right), \tag{12.131}$$

from which we can read off $\mathcal{L}$. Overdots denote the derivative with respect to time, and the integral runs over the space of the reference copy of the body.

The Euler–Lagrange equations are found entirely from the variation of the action with respect to the displacement field f. Since the Lagrangian depends on f through only its time and space derivatives, the Euler–Lagrange equations are

$$\frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{f}}\right) + \frac{\partial}{\partial x^i}\left(\frac{\partial \mathcal{L}}{\partial(\partial_i f)}\right) = 0. \tag{12.132}$$

These simplify to

$$\rho \dot{\boldsymbol{v}} = \frac{\partial}{\partial x^i}\left(\frac{\partial U}{\partial(\partial_i f)}\right), \tag{12.133}$$

where $\boldsymbol{v} = \dot{y}$ is the local velocity of the body. Comparison with equation (6.319) tells us that we must have

$$\mathsf{T}(\mathsf{e}^i) = \frac{\partial U}{\partial(\partial_i f)}, \tag{12.134}$$

where $\{\mathsf{e}_i\}$ is the (fixed) coordinate frame defined by the x^i coordinates. To simplify the right-hand side we employ the derivative with respect to a linear function, as defined in section 11.1.2. From the definition of equation (11.24) we have

$$\frac{\partial}{\partial \mathsf{f}(\mathsf{e}^i)} = \mathsf{e}_j \frac{\partial}{\partial \mathsf{f}_{ji}}. \tag{12.135}$$

The scalars f_{ji} are defined by

$$\mathsf{f}_{ji} = \mathsf{e}_j \cdot \mathsf{f}(\mathsf{e}_i) = \mathsf{e}_j \cdot (\partial_i f). \tag{12.136}$$

These are the components of the vector $\partial_i f$, so we can write

$$\frac{\partial}{\partial(\partial_i f)} = \mathsf{e}_j \frac{\partial}{\partial \mathsf{f}_{ji}} = \frac{\partial}{\partial \mathsf{f}(\mathsf{e}^i)}. \tag{12.137}$$

The notation for the derivative with respect to $\mathsf{f}(\mathsf{e}^i)$ is slightly misleading, as f is never evaluated on e^i. Instead, we have

$$\mathsf{T}(\boldsymbol{a}) = \partial_{\mathsf{f}(\boldsymbol{a})} U(\mathsf{C}). \tag{12.138}$$

The fact that U is a function of $\mathsf{C} = \bar{\mathsf{f}}\mathsf{f}$ ensures that the second Piola–Kirchoff stress $\mathcal{T} = \mathsf{f}^{-1}\mathsf{T}$ is a symmetric function.

To make further progress we must specify the precise form of U, which amounts to specifying the constitutive properties of the system. The simplest hyperelastic materials to consider are isotropic and homogeneous. For these the internal energy can only depend on the principal stretches:

$$W = W(\lambda_1, \lambda_2, \lambda_3). \tag{12.139}$$

Even within this class there are a large variety of models one can consider. To obtain a linear model the energy should be quadratic in the strains, where linear in this context refers to the relationship between the stress and strain tensors, and not to the underlying dynamics. A natural model to consider is to define the strain by

$$\mathcal{E}(\boldsymbol{a}) = \mathsf{C}^{1/2}(\boldsymbol{a}) - \boldsymbol{a}, \tag{12.140}$$

and set

$$\begin{aligned} U(\mathcal{E}) &= G\,\mathrm{tr}(\mathcal{E}^2) + (B/2 - G/3)\,\mathrm{tr}(\mathcal{E})^2 \\ &= G\big(\mathrm{tr}(\mathsf{C}) - 2\mathrm{tr}(\mathsf{C}^{1/2}) + 3\big) + (B/2 - G/3)\big(\mathrm{tr}(\mathsf{C}^{1/2}) - 3\big)^2, \end{aligned} \tag{12.141}$$

where B and G are respectively the (constant) bulk and the shear moduli. To

find the stress tensor we need the derivative of $\mathrm{tr}((\bar{\mathsf{f}}\mathsf{f})^{1/2})$. To evaluate this we first write

$$(\bar{\mathsf{f}}\mathsf{f})^{1/2} = \exp\big(\tfrac{1}{2}\ln(\mathsf{C})\big) = \exp(\mathcal{E}_{\ln}), \tag{12.142}$$

where

$$\mathcal{E}_{\ln} = \tfrac{1}{2}\ln(\mathsf{C}). \tag{12.143}$$

We can now make use of the result

$$\partial_{\mathsf{f}(\boldsymbol{a})}\mathrm{tr}(\mathcal{E}_{\ln}^{n}) = n\bar{\mathsf{f}}^{-1}\mathcal{E}_{\ln}^{n-1}(\boldsymbol{a}) \tag{12.144}$$

to prove that

$$\partial_{\mathsf{f}(\boldsymbol{a})}\mathrm{tr}\big((\bar{\mathsf{f}}\mathsf{f})^{1/2}\big) = \bar{\mathsf{f}}^{-1}(\bar{\mathsf{f}}\mathsf{f})^{1/2}(\boldsymbol{a}). \tag{12.145}$$

The stress tensor T therefore evaluates to

$$\begin{aligned}\mathsf{T} &= 2G(\mathsf{f} - \bar{\mathsf{f}}^{-1}(\bar{\mathsf{f}}\mathsf{f})^{1/2}) + (B - 2G/3)\mathrm{tr}(\mathcal{E})\bar{\mathsf{f}}^{-1}(\bar{\mathsf{f}}\mathsf{f})^{1/2}\\ &= \bar{\mathsf{f}}^{-1}(\bar{\mathsf{f}}\mathsf{f})^{1/2}\big(2G\mathcal{E} + (B - 2G/3)\mathrm{tr}(\mathcal{E})\mathsf{I}\big),\end{aligned} \tag{12.146}$$

where I is the identity transformation. The bracketed term is the expression we would expect to see in a linear theory. The extra pre-factor can be understood in terms of a singular-value decomposition of f. We write

$$\mathsf{f} = \mathsf{R}\mathsf{C}^{1/2}, \tag{12.147}$$

where R is a rotation. We then find that

$$\bar{\mathsf{f}}^{-1}\mathsf{C}^{1/2} = \mathsf{R}\mathsf{C}^{-1/2}\mathsf{C}^{1/2} = \mathsf{R}, \tag{12.148}$$

which recovers the rotation. We can now write

$$\mathsf{T}(\boldsymbol{a}) = \mathsf{R}\big((2G\mathcal{E}(\boldsymbol{a}) + (B - 2G/3)\mathrm{tr}(\mathcal{E})\boldsymbol{a}\big). \tag{12.149}$$

This can be understood as a linear function of $\mathcal{E}$, followed by a rotation to align the principal axes in the reference configuration with those in the body.

The definition of $\mathcal{E}_{\ln}$ raises the interesting prospect that this could be used as an alternative definition of the strain. For an isotropic, homogeneous media this amounts to choosing an energy density of

$$U_{\ln} = G\big((\ln\lambda_1)^2 + (\ln\lambda_2)^2 + (\ln\lambda_3)^2\big) + (B/2 - G/3)\big(\ln(\lambda_1\lambda_2\lambda_3)\big)^2. \tag{12.150}$$

This definition has the same behaviour under small deflection as the potential energy of equation (12.141), but differences emerge as the stresses build up. In essence, the logarithmic definition of energy defines a material which retains its elastic properties no matter what shape it is stretched into. This limits the application of $U_{\ln}$ for modelling physical objects, though it may well be of use in computer graphics simulations, as routines built on $U_{\ln}$ will not break down when forces become large.

12.4.2 Relativistic fluid dynamics

The field equations for a relativistic fluid can be formulated in a number of different ways. Here we give a fairly direct derivation, albeit from a slightly surprising starting point. We start with the action integral

$$S = \int d^4x \left(-\varepsilon + J\cdot(\nabla\lambda) - \mu J\cdot\nabla\eta\right), \tag{12.151}$$

where $J(x)$ is a spacetime current, ε is the total energy density and η is the entropy. The current can be written as

$$J = \rho v, \qquad v^2 = 1, \tag{12.152}$$

and we assume that ε is a function of ρ and η only, which we write as

$$\varepsilon = \rho(1 + e(\rho, \eta)). \tag{12.153}$$

The remaining terms λ and μ are Lagrange multipliers enforcing the two constraints

$$\nabla\cdot J = 0, \qquad v\cdot\nabla\eta = 0. \tag{12.154}$$

These are two of the four equations of motion. The first constraint is that the current is conserved, so the total number of particles in the system is constant. The second constraint says that entropy is constant along the field lines of J. The various constraints and assumptions ensure that we are describing a relativistic ideal fluid.

Variation with respect to η yields the equation

$$\frac{\partial e}{\partial \eta} = v\cdot\nabla\mu, \tag{12.155}$$

and variation with respect to J produces

$$v(1 + e) + v\rho\frac{\partial e}{\partial \rho} = \nabla\lambda - \mu\nabla\eta. \tag{12.156}$$

In the derivation of this equation we have employed the result that

$$\partial_J f(\rho) = v\frac{\partial f}{\partial \rho}. \tag{12.157}$$

Next, we define the pressure P by

$$P = \rho^2\frac{\partial e}{\partial \rho}, \tag{12.158}$$

so that equation (12.156) becomes

$$v(\varepsilon + P) = \rho(\nabla\lambda - \mu\nabla\eta). \tag{12.159}$$

The final step is to remove the Lagrange multipliers by employing the constraint equations. First we contract equation (12.159) with v to obtain

$$\varepsilon + P = \rho v \cdot \nabla \lambda = J \cdot (\nabla \lambda). \tag{12.160}$$

Next we differentiate equation (12.159) to obtain

$$\begin{aligned} J \cdot \nabla (\nabla \lambda - \mu \nabla \eta) &= v \cdot \nabla \big(v(\varepsilon + P) \big) + v(\varepsilon + P) J \cdot \nabla (\rho^{-1}) \\ &= v \cdot \nabla \big(v(\varepsilon + P) \big) + v(\varepsilon + P) \nabla \cdot v. \end{aligned} \tag{12.161}$$

The left-hand side is manipulated as follows:

$$\begin{aligned} J \cdot \nabla (\nabla \lambda - \mu \nabla \eta) &= \nabla (J \cdot \nabla \ \lambda) - \dot{\nabla} \dot{J} \cdot \nabla \lambda - \rho \frac{\partial e}{\partial \eta} \nabla \eta + \mu \dot{\nabla} \dot{J} \cdot \nabla \eta \\ &= \nabla (\varepsilon + P) - \rho \frac{\partial e}{\partial \eta} \nabla \eta - \dot{\nabla} \dot{J} \cdot v \frac{(\varepsilon + P)}{\rho} \\ &= \nabla P + \nabla \rho \, \frac{\partial \varepsilon}{\partial \rho} - \frac{(\varepsilon + P)}{\rho} \nabla \rho \\ &= \nabla P. \end{aligned} \tag{12.162}$$

We therefore arrive at the equation

$$v \cdot \nabla \big(v(\varepsilon + P) \big) + v(\varepsilon + P) \nabla \cdot v = \nabla P, \tag{12.163}$$

which describes a relativistic ideal fluid. This is more clearly seen if we introduce the relativistic stress-energy tensor $\mathsf{T}(a)$, which is defined by

$$\mathsf{T}(a) = (\varepsilon + P) a \cdot v \, v - P a. \tag{12.164}$$

The rest frame of the fluid is defined locally by v. We find that $\mathsf{T}(v) = \varepsilon v$, so that ε is the local energy density, as required. In any spacelike direction n perpendicular to v we have $\mathsf{T}(n) = -Pn$, which shows that the local stress is governed by an isotropic pressure P. These are the relativistic definitions of the stress-energy tensor for an ideal fluid. The field equations reduce to the single conservation equation

$$\dot{\mathsf{T}}(\dot{\nabla}) = 0, \tag{12.165}$$

which expresses relativistic conservation of the stress-energy tensor. Electromagnetic coupling is included simply with the addition of the term $-qJ \cdot A$ to the Lagrangian density.

12.5 Notes

The Lagrangian formulation of mechanics is described in a wide range of books. *Analytical Mechanics* by Hand & Finch (1998) contains a detailed introduction and, despite its name, *Introduction to Mechanics and Symmetry* by Marsden & Ratiu (1994) contains a more detailed description of Lagrangian and Hamiltonian

methods and symplectic geometry. Further applications, including relativistic fluid dynamics, are contained in *The Variational Principles of Dynamics* by Kupershmidt (1992).

Pseudoclassical mechanics was introduced by Berezin & Marinov (1977). Further references are contained in the notes to chapter 11. Similar ideas to those developed in this chapter have been applied in the supersymmetric setting by Heumann & Manton (2000). The section on spinor models of relativistic spin-1/2 point particles was motivated by the initial work of Barut & Zanghi (1984). The description given here contains a number of refinements, many of which are also discussed in Doran (1994). A detailed discussion of the complexities involved in performing a path-integral quantisation of such systems is given by Barut & Duru (1989).

12.6 Exercises

12.1 A relativistic action for a point particle is defined by

$$S = \int d\lambda \left(-p\cdot\dot{x} + \frac{e}{2}(p^2 - m^2) - qx'\cdot A(x) \right),$$

where A is an external field representing the electromagnetic vector potential. Vary S with respect to x, p and e to obtain the Lorentz force law.

12.2 Prove that

$$\partial_\psi \langle M\psi^{-1} \rangle = -\psi^{-1} M \psi^{-1},$$

where ψ and M are even multivectors.

12.3 The configuration of a rigid body is described by a rotor R. If we relax the normalisation of R and replace it by ψ, explain why we can write Ω_B as

$$\Omega_B = -2\tilde{R}\dot{R} = -\psi^{-1}\dot{\psi} + \dot{\psi}^\dagger \psi^{\dagger\,-1}.$$

Now define the Lagrangian

$$L(\psi, \dot{\psi}) = -\tfrac{1}{2}\Omega_B \cdot \mathcal{I}(\Omega_B),$$

where $\mathcal{I}$ is the inertia tensor. Find the Euler–Lagrange equation for variation with respect to ψ. Prove that this produces the equation of motion $\dot{J} = 0$, where J is the angular momentum. Why does this method work?

12.4 One classical model for a spin-1/2 particle describes the motion in terms of a rotor R and a momentum p. The rotor determines the quantities

$$\dot{x} = R\gamma_0\tilde{R}, \qquad S = RI\boldsymbol{\sigma}_3\tilde{R}$$

and the equations of motion are

$$\dot{p} = 0, \quad p \cdot \dot{x} = m, \quad \dot{S} + 2\dot{x} \wedge p = 0.$$

Verify that these are solved by

$$p = \frac{m}{\cosh(\alpha)}\gamma_0 = m^*\gamma_0, \qquad R = \mathrm{e}^{I\boldsymbol{\sigma}_3 m^* \tau}\mathrm{e}^{\alpha\boldsymbol{\sigma}_2/2}.$$

Integrate $\dot{x}$ to find the trajectory of the particle and comment on its properties.

12.5 Find the equations of motion for the Lagrangian

$$L = \langle \psi' I\boldsymbol{\sigma}_3 \tilde{\psi} - p(x' - e\psi\gamma_0\tilde{\psi}) - em\psi\tilde{\psi} - qx' \cdot A \rangle,$$

where ψ is a spinor and $A(x)$ is an external electromagnetic vector potential. Comment on the form of the solutions.

12.6 A set of vectors satisfy the equations

$$\dot{e}_i + \epsilon_{ijk}\omega_j e_k = 0,$$

where the ω_i are constant. Prove that the volume element E is constant, where

$$E = e_1 \wedge e_2 \wedge e_3.$$

12.7 The relativistic Hamiltonian for a charged particle in three dimensions is defined by

$$H(\boldsymbol{p}, \boldsymbol{x}, t) = \left((\boldsymbol{p} - q\boldsymbol{A})^2 + m^2\right)^{1/2} + q\phi,$$

where $\phi + \boldsymbol{A} = A\gamma_0$ and the vector potential A is a function of $\boldsymbol{x}$ and t. Find Hamilton's equations and prove that these recover the Lorentz force law.

12.8 Fill in the missing steps in the proof that a closed non-degenerate 2-form in a symplectic manifold guarantees that the Poisson bracket satisfies the Jacobi identity.

12.9 A system is described by the Hamiltonian

$$H(p, q) = \frac{1}{2}\left(\frac{1}{q^2} + p^2 q^4\right).$$

Find a canonical transformation which maps this onto the Hamiltonian for a simple harmonic oscillator.

12.10 The total energy in a hyperelastic medium is given by

$$E = \int d^3x \left(\frac{\rho(x)}{2}\dot{f}^2 + U\right).$$

Prove that the energy flow per unit area perpendicular to $\boldsymbol{n}$ is given by $\dot{f} \cdot \mathsf{T}(\boldsymbol{n})$, where $\boldsymbol{n}$ is a vector in the reference configuration.

12.11 An incompressible elastic material is one for which $\det \mathsf{f} = 1$. The Mooney–Rivlin model for rubber is as an incompressible material with internal energy

$$U = \alpha(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3) + \beta((\lambda_2\lambda_3)^2 + (\lambda_3\lambda_1)^2 + (\lambda_1\lambda_2)^2 - 3),$$

where the λ_i denote the principal strains. Analyse the properties of this material under uniform pressure. What happens when two of the λ_i pass through $4^{1/3}$?

12.12 A hyperelastic material is defined with an energy density

$$U_{\ln} = G\big((\ln\lambda_1)^2 + (\ln\lambda_2)^2 + (\ln\lambda_3)^2\big) + (B/2 - G/3)\big(\ln(\lambda_1\lambda_2\lambda_3)\big)^2.$$

Prove that when this system is placed under isotropic pressure P we have

$$-3B\ln\lambda = P.$$