IOP Publishing

A Modern Course in Quantum Field Theory, Volume 1
Fundamentals
Badis Ydri

Appendix A

Exercises[1]

Exercise 1: We consider the two Euclidean integrals

$$I(m^2) = \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2 + m^2}.$$

$$J(p^2, m^2) = \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2 + m^2} \frac{1}{(p - k)^2 + m^2}.$$

- Determine in each case the divergent behavior of the integral.
- Use dimensional regularization to compute the above integrals. Determine in each case the divergent part of the integral. In the case of $J(p^2, m^2)$ assume for simplicity zero external momentum $p = 0$.

Exercise 2: The two integrals in exercise 1 can also be regularized using a cut-off Λ. First we perform Laplace transform as follows

$$\frac{1}{k^2 + m^2} = \int_0^\infty d\alpha e^{-\alpha(k^2+m^2)}.$$

- Do the integral over k in $I(m^2)$ and $J(p^2, m^2)$. In the case of $J(p^2, m^2)$ assume for simplicity zero external momentum $p = 0$.

[1] These exercises were given as QFT examinations.

- The remaining integral over α is regularized by replacing the lower bound $\alpha = 0$ by $\alpha = 1/\Lambda^2$. Perform the integral over α explicitly. Determine the divergent part in each case.

  Hint: Use the exponential-integral function

$$Ei(-x) = \int_{-\infty}^{-x} \frac{e^t}{t} dt = \mathbf{C} + \ln x + \int_0^x dt \frac{e^{-t} - 1}{t}.$$

Exercise 3: Let z_i be a set of complex numbers, θ_i be a set of anticommuting Grassmann numbers and let M be a hermitian matrix. Perform the following integrals

$$\int \prod_i dz_i^+ dz_i e^{-M_{ij} z_i^+ z_j - z_i^+ j_i - j_i^+ z_i}.$$

$$\int \prod_i d\theta_i^+ d\theta_i e^{-M_{ij} \theta_i^+ \theta_i - \theta_i^+ \eta_i - \eta_i^+ \theta_i}.$$

Exercise 4: Let $S(r, \theta)$ be an action dependent on two degrees of freedom r and θ which is invariant under two-dimensional rotations, i.e. $\vec{r} = (r, \theta)$. We propose to gauge fix the following two-dimensional path integral

$$W = \int e^{iS(\vec{r})} d^2\vec{r}.$$

We will impose the gauge condition

$$g(r, \theta) = 0.$$

- Show that

$$\left| \frac{\partial g(r, \theta)}{\partial \theta} \right|_{g=0} \int d\phi \delta(g(r, \theta + \phi)) = 1.$$

- Use the above identity to gauge fix the path integral W.

Exercise 5: The gauge fixed path integral of quantum electrodynamics is given by

$$Z[J] = \int \prod_\mu \mathcal{D}A_\mu \exp\left(-i \int d^4x \frac{(\partial_\mu A^\mu)^2}{2\xi} - \frac{i}{4} \int d^4x F_{\mu\nu} F^{\mu\nu} - i \int d^4x J_\mu A^\mu\right).$$

- Derive the equations of motion.
- Compute $Z[J]$ in a closed form.
- Derive the photon propagator.

Exercise 6: We consider phi-four interaction in four dimensions. The action is given by

$$S[\phi] = \int d^4x\left[\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}(\phi^2)^2\right].$$

- Write down Feynman rules in momentum space.
- Use Feynman rules to derive the 2-point proper vertex $\Gamma^2(p)$ up to the 1-loop order. Draw the corresponding Feynman diagrams.
- Use Feynman rules to derive the 4-point proper vertex $\Gamma^4(p_1, p_2, p_3, p_4)$ up to the 1-loop order. Draw the corresponding Feynman diagrams.
- By assuming that the momentum loop integrals are regularized perform 1-loop renormalization of the theory. Impose the two conditions

  $$\Gamma^2(0) = m_R^2, \quad \Gamma^4(0, 0, 0, 0) = \lambda_R.$$

  Determine the bare coupling constants m^2 and λ in terms of the renormalized coupling constants m_R^2 and λ_R.
- Determine $\Gamma^2(p)$ and $\Gamma^4(p_1, p_2, p_3, p_4)$ in terms of the renormalized coupling constants.

Exercise 7:

- Write down an expression of the free scalar field in terms of creation and annihilation.
- Compute the 2-point function

  $$D_F(x_1 - x_2) = \langle 0|T\hat{\phi}(x_1)\hat{\phi}(x_2)|0\rangle.$$

- Compute in terms of D_F the 4-point function

  $$D(x_1, x_2, x_3, x_4) = \langle 0|T\hat{\phi}(x_1)\hat{\phi}(x_2)\hat{\phi}(x_3)\hat{\phi}(x_4)|0\rangle.$$

- Without calculation what is the value of the 3-point function $\langle 0|T\hat{\phi}(x_1)\hat{\phi}(x_2)\hat{\phi}(x_3)|0\rangle$. Explain.

Exercise 8: The electromagnetic field is a vector in four-dimensional Minkowski spacetime denoted by

$$A^\mu = (A^0, \vec{A}).$$

A^0 is the electric potential and $\vec{A}$ is the magnetic vector potential. The Dirac Lagrangian density with non-zero external electromagnetic field is given

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - e\bar{\psi}\gamma_\mu\psi A^\mu.$$

Derive the Euler–Lagrange equation of motion.

Exercise 9: Compute the integral over p^0:

$$\int d^3\vec{p} \int dp^0 \ \delta(p^2 - m^2).$$

What do you conclude for the action of Lorentz transformations on

$$\frac{d^3\vec{p}}{2E_p}.$$

Exercise 10: The Yukawa Lagrangian density describes the interaction between spinorial and scalar fields. It is given by

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi - \phi^2) - g\phi\bar{\psi}\psi.$$

Derive the Euler–Lagrange equation of motion.

Exercise 11: Show that the Feynman propagator in one dimension is given by

$$G_{\vec{p}}(t - t') = \int \frac{dE}{2\pi} \frac{i}{E^2 - E_{\vec{p}}^2 + i\epsilon} e^{-iE(t-t')} = \frac{e^{-iE_{\vec{p}}|t-t'|}}{2E_{\vec{p}}}.$$

Exercise 12:
- What is the condition satisfied by the Dirac matrices in order for the Dirac equation to be covariant.
- Write down the spin representation of the infinitesimal Lorentz transformations

$$\Lambda = 1 - \frac{i}{2}\epsilon_{\mu\nu}\mathcal{L}^{\mu\nu}.$$

Exercise 13:
- Show that gamma matrices in two dimensions are given by

$$\gamma^0 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

- Write down the general solution of Dirac equation in two dimensions in the massless limit.

Exercise 14:
- Write down the vacuum stability condition.
- Write down Gell-Mann–Low formulas.

- Write down the scattering S-matrix.
- Write down the Lehmannn–Symanzik–Zimmermann (LSZ) reduction formula which expresses the transition probability amplitude between 1-particle states in terms of the 2-point function.
- Write down the Lehmannn–Symanzik–Zimmermann (LSZ) reduction formula which expresses the transition probability amplitude between 2-particle states in terms of the 4-point function.
- Write down Wick's theorem. Apply for 2, 4 and 6 fields.

Exercise 15: We consider phi-cube theory in four dimensions where the interaction is given by the Lagrangian density

$$\mathcal{L}_{\mathrm{int}} = -\frac{\lambda}{3!}\phi^3.$$

- Compute the 0-point function up to the second order of perturbation theory and express the result in terms of Feynman diagrams.
- Compute the 1-point function up to the second order of perturbation theory and express the result in terms of Feynman diagrams.
- Compute the 2-point function up to the second order of perturbation theory and express the result in terms of Feynman diagrams.
- Compute the connected 2-point function up to the second order of perturbation theory and express the result in terms of Feynman diagrams.

Exercise 16: We consider phi-four theory in four dimensions where the interaction is given by the Lagrangian density

$$\mathcal{L}_{\mathrm{int}} = -\frac{\lambda}{4!}\phi^4.$$

Compute the 4-point function up to the first order of perturbation theory and express the result in terms of Feynman diagrams.

Exercise 17: Show that

$$\langle 0|T\hat{\phi}(x)\hat{\phi}(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2 - m^2 + i\epsilon}e^{-ip(x-y)}.$$

We give

$$\hat{\phi}(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E(\vec{p})}}(\hat{a}(\vec{p})e^{-ipx} + \hat{a}(\vec{p})^+e^{ipx}).$$

Exercise 18: Show that the scalar field operator $\hat{\phi}_I(x)$ and the conjugate momentum field operator $\hat{\pi}_I(x)$ (operators in the interaction picture) are free field operators.

Exercise 19: Calculate the 2-point function $\langle 0|T(\hat{\phi}(x_1)\hat{\phi}(x_2))|0\rangle$ in ϕ-four theory up to the second order in perturbation theory using the Gell-Mann Low formula and Wick's theorem. Express each order in perturbation theory in terms of Feynman diagrams.

Exercise 20: We consider a single forced harmonic oscillator given by the equation of motion

$$(\partial_t^2 + E^2)Q(t) = J(t).$$

- Show that the S-matrix defined by the matrix elements $S_{mn} = \langle m \ \ \mathrm{out}|n \ \ \mathrm{in}\rangle$ is unitary.
- Determine S from solving the equation

$$S^{-1}\hat{a}_{\mathrm{in}}S = \hat{a}_{\mathrm{out}} = \hat{a}_{\mathrm{in}} + \frac{i}{\sqrt{2E}}j(E).$$

- Compute the probability $|\langle n \ \ \mathrm{out}|0 \ \ \mathrm{in}\rangle|^2$.
- Determine the evolution operator in the interaction picture $\Omega(t)$ from solving the Schrödinger equation

$$i\partial_t\Omega(t) = \hat{V}_I(t)\Omega(t), \quad \hat{V}_I(t) = -J(t)\hat{Q}_I(t).$$

- Deduce from (4) the S-matrix and compare with the result of (2).

Exercise 21: The probability amplitudes for a Dirac particle (antiparticle) to propagate from the spacetime point y (x) to the spacetime x (y) are

$$S_{ab}(x - y) = \langle 0|\hat{\psi}_a(x)\bar{\hat{\psi}}_b(y)|0\rangle.$$

$$\bar{S}_{ba}(y - x) = \langle 0|\bar{\hat{\psi}}_b(y)\hat{\psi}_a(x)|0\rangle.$$

- Compute S and $\bar{S}$ in terms of the Klein–Gordon propagator $D(x - y)$ given by

$$D(x - y) = \int \frac{d^3p}{(2\pi\hbar)^3}\frac{1}{2E(\vec{p})}e^{-\frac{i}{\hbar}p(x-y)}.$$

- Show that the retarded Green's function of the Dirac equation is given by
$$(S_R)_{ab}(x - y) = \langle 0|\{\hat{\psi}_a(x), \bar{\hat{\psi}}_b(y)\}|0\rangle.$$
- Verify that S_R satisfies the Dirac equation
$$(i\hbar\gamma^\mu\partial_\mu^x - mc)_{ca}(S_R)_{ab}(x - y) = i\frac{\hbar}{c}\delta^4(x - y)\delta_{cb}.$$
- Derive an expression of the Feynman propagator in terms of the Dirac fields $\hat{\psi}$ and $\bar{\hat{\psi}}$ and then write down its Fourier expansion.

Exercise 22:

- Compute the electron 2-point function in configuration space up to 1-loop using the Gell-Mann Low formula and Wick's theorem. Write down the corresponding Feynman diagrams.
- Compute the electron 2-point function in momentum space up to 1-loop using Feynman rules.
- Use dimensional regularization to evaluate the electron self-energy. Add a small photon mass to regularize the IR behavior. What is the UV behavior of the electron self-energy.
- Determine the physical mass of the electron at 1-loop.
- Determine the wave-function renormalization Z_2 and the counter term $\delta_2 = 1 - Z_2$ up to 1-loop.

Exercise 23:

- Write down all Feynman diagrams up to 1-loop which contribute to the probability amplitude of the process $e^-(p) + \mu^-(k) \longrightarrow e^-(p') + \mu^-(k')$.
- Write down using Feynman rules the tree-level probability amplitude of the process $e^-(p) + \mu^-(k) \longrightarrow e^-(p') + \mu^-(k')$. Write down the probability amplitude of this process at 1-loop due to the electron vertex correction.
- Use Feynman parameters to express the product of propagators as a single propagator raised to some power of the form
$$\frac{1}{[L^2 - \Delta + i\epsilon]^q}.$$
Determine the shifted momentum L, the effective mass Δ and the power q. Add a small photon mass μ^2.
- Verify the relations
$$(\gamma \cdot p)\gamma^\mu = 2p^\mu - \gamma^\mu(\gamma \cdot p)$$
$$\gamma^\mu(\gamma \cdot p) = 2p^\mu - (\gamma. p)\gamma^\mu$$
$$(\gamma \cdot p)\gamma^\mu(\gamma \cdot p') = 2p^\mu(\gamma \cdot p') - 2\gamma^\mu p \cdot p' + 2p'^\mu(\gamma \cdot p) - (\gamma \cdot p')\gamma^\mu(\gamma \cdot p).$$

- We work in d dimensions. Use Lorentz invariance, the properties of the gamma matrices in d dimensions and the results of question (4) to show that we can replace

$$\gamma^\lambda \cdot i(\gamma \cdot l' + m_e) \cdot \gamma^\mu \cdot i(\gamma \cdot l + m_e)\gamma_\lambda \longrightarrow \gamma^\mu A + (p + p')^\mu B + (p - p')^\mu C.$$

  Determine the coefficients A, B and C.
- Use Gordon's identity to show that the vertex function $\Gamma(p', p)$ is of the form

$$\Gamma^\mu(p', p) = \gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m_e}F_2(q^2).$$

  Determine the form factors F_1 and F_2.
- Compute the integrals

$$\int \frac{d^dL_E}{(2\pi)^d}\frac{L_E^2}{\left(L_E^2 + \Delta\right)^3}, \quad \int \frac{d^dL_E}{(2\pi)^d}\frac{1}{\left(L_E^2 + \Delta\right)^3}.$$

- Calculate the form factor $F_1(q^2)$ explicitly in dimensional regularization. Determine the UV behavior.
- Compute the renormalization constant Z_1 or equivalently the counter term $\delta_1 = Z_1 - 1$ at 1-loop.
- Prove the Ward identity $\delta_1 = \delta_2$.

Exercise 24:
- Write down using Feynman rules the photon self-energy $i\Pi_2^{\mu\nu}(q)$ at one-loop.
- Use dimensional regularization to show that

$$\Pi_2^{\mu\nu}(q) = \Pi_2(q^2)(q^2\eta^{\mu\nu} - q^\mu q^\nu). \tag{A.1}$$

  Determine $\Pi_2(q^2)$. What is the UV behavior.
- Compute at one-loop the counter term $\delta_3 = Z_3 - 1$.
- Compute at one-loop the effective charge e_{eff}^2. How does the effective charge behave at high energies.

Exercise 25: Compute the unpolarized differential cross section of the process $e^-{+}e^+ \longrightarrow \mu^-{+}\mu^+$ in the center of mass system.

IOP Publishing

A Modern Course in Quantum Field Theory, Volume 1
Fundamentals
Badis Ydri

Appendix B

Classical mechanics

B.1 D'Alembert principle

We consider a system of many particles and let $\vec{r}_i$ and m_i be the radius vector and the mass, respectively, of the ith particle. Newton's second law of motion for the ith particle reads

$$\vec{F}_i = \vec{F}_i^{(e)} + \sum_j \vec{F}_{ji} = \frac{d\vec{p}_i}{dt}. \tag{B.1}$$

The external force acting on the ith particle is $\vec{F}_i^{(e)}$, whereas $\vec{F}_{ji}$ is the internal force on the ith particle due to the jth particle ($\vec{F}_{ii} = 0$ and $\vec{F}_{ij} = -\vec{F}_{ji}$). The momentum vector of the ith particle is $\vec{p}_i = m_i\vec{v}_i = m_i\frac{d\vec{r}_i}{dt}$. Thus we have

$$\vec{F}_i = \vec{F}_i^{(e)} + \sum_j \vec{F}_{ji} = m_i\frac{d^2\vec{r}_i}{dt^2}. \tag{B.2}$$

By summing over all particles we get

$$\sum_i \vec{F}_i = \sum_i \vec{F}_i^{(e)} = \sum_i m_i\frac{d^2\vec{r}_i}{dt^2} = M\frac{d^2\vec{R}}{dt^2}. \tag{B.3}$$

The total mass M is $M = \sum_i m_i$ and the average radius vector $\vec{R}$ is $\vec{R} = \sum_i m_i\vec{r}_i/M$. This is the radius vector of the center of mass of the system. Thus the internal forces if they obey Newton's third law of motion will have no effect on the motion of the center of mass.

doi:10.1088/2053-2563/ab0547ch10

The goal of mechanics is to solve the set of second-order differential equations (B.2) for $\vec{r}_i$ given the forces $\vec{F}_i^{(e)}$ and $\vec{F}_{ji}$. This task is, in general, very difficult and it is made even more complicated by the possible presence of constraints which limit the motion of the system. As an example, we take the class of systems known as rigid bodies in which the motion of the particles is constrained in such a way that the distances between the particles are kept fixed and do not change in time. It is clear that constraints correspond to forces which cannot be specified directly but are only known via their effect on the motion of the system. We will only consider holonomic constraints which can be expressed by equations of the form

$$f(\vec{r}_1, \vec{r}_2, \vec{r}_3, \dots, t) = 0. \tag{B.4}$$

The constraints which cannot be expressed in this way are called non-holonomic. In the example of rigid bodies, the constraints are holonomic since they can be expressed as

$$(\vec{r}_i - \vec{r}_j)^2 - c_{ij}^2 = 0. \tag{B.5}$$

The presence of constraints means that not all the vectors $\vec{r}_i$ are independent, i.e. not all the differential equations (B.2) are independent. We assume that the system contains N particles and that we have k holonomic constraints. Then there must exist $3N - k$ independent degrees of freedom q_i which are called generalized coordinates. We can therefore express the vectors $\vec{r}_i$ as functions of the independent generalized coordinates q_i as

$$\begin{aligned} \vec{r}_1 &= \vec{r}_1(q_1, q_2, \dots, q_{3N-k}, t) \\ &\vdots \\ \vec{r}_N &= \vec{r}_N(q_1, q_2, \dots, q_{3N-k}, t). \end{aligned} \tag{B.6}$$

Let us compute the work done by the forces $\vec{F}_i^{(e)}$ and $\vec{F}_{ji}$ in moving the system from an initial configuration 1 to a final configuration 2. We have

$$W_{12} = \sum_i \int_1^2 \vec{F}_i d\vec{s}_i = \sum_i \int_1^2 \vec{F}_i^{(e)} d\vec{s}_i + \sum_{i,j} \int_1^2 \vec{F}_{ji} d\vec{s}_i. \tag{B.7}$$

We have on one hand

$$\begin{aligned} W_{12} = \sum_i \int_1^2 \vec{F}_i d\vec{s}_i &= \sum_i \int_1^2 m_i \frac{d\vec{v}_i}{dt} \vec{v}_i dt \\ &= \sum_i \int_1^2 d\left(\frac{1}{2} m_i v_i^2\right) \\ &= T_2 - T_1. \end{aligned} \tag{B.8}$$

The total kinetic energy is defined by

$$T = \sum_i \frac{1}{2} m_i v_i^2. \tag{B.9}$$

We assume that the external forces $\vec{F}_i^{(e)}$ are conservative, i.e. they are derived from potentials V_i such that

$$\vec{F}_i^{(e)} = -\vec{\nabla}_i V_i. \tag{B.10}$$

Then we compute

$$\sum_i \int_1^2 \vec{F}_i^{(e)} d\vec{s}_i = -\sum_i \int_1^2 \vec{\nabla}_i V_i d\vec{s}_i = -\sum_i V_i|_1^2. \tag{B.11}$$

We also assume that the internal forces $\vec{F}_{ji}$ are derived from potentials V_{ij} such that

$$\vec{F}_{ji} = -\vec{\nabla}_i V_{ij}. \tag{B.12}$$

Since we must have $\vec{F}_{ij} = -\vec{F}_{ji}$ we must take V_{ij} as a function of the distance $|\vec{r}_i - \vec{r}_j|$ only, i.e. $V_{ij} = V_{ji}$. We can also check that the force $\vec{F}_{ij}$ lies along the line joining the particles i and j.

We define the difference vector by $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$. We have then $\vec{\nabla}_i V_{ij} = -\vec{\nabla}_j V_{ij} = \vec{\nabla}_{ij} V_{ij}$. We then compute

$$\begin{aligned}
\sum_{i,j} \int_1^2 \vec{F}_{ji} d\vec{s}_i &= -\frac{1}{2}\sum_{i,j} \int_1^2 (\vec{\nabla}_i V_{ij} d\vec{s}_i + \vec{\nabla}_j V_{ij} d\vec{s}_j) \\
&= -\frac{1}{2}\sum_{i,j} \int_1^2 \vec{\nabla}_{ij} V_{ij}(d\vec{s}_i - d\vec{s}_j) \\
&= -\frac{1}{2}\sum_{i,j} \int_1^2 \vec{\nabla}_{ij} V_{ij} d\vec{r}_{ij} \\
&= -\frac{1}{2}\sum_{i\neq j} V_{ij}|_1^2.
\end{aligned} \tag{B.13}$$

Thus the work done is found to be given by

$$W_{12} = -V_2 + V_1. \tag{B.14}$$

The total potential is given by

$$V = \sum_i V_i + \frac{1}{2}\sum_{i\neq j} V_{ij}. \tag{B.15}$$

From the results $W_{12} = T_2 - T_1$ and $W_{12} = -V_2 + V_1$ we conclude that the total energy $T + V$ is conserved. The term $\frac{1}{2}\sum_{i\neq j} V_{ij}$ in V is called the internal potential energy of the system.

For rigid bodies the internal energy is constant since the distances $|\vec{r}_i - \vec{r}_j|$ are fixed. Indeed, in rigid bodies the vectors $d\vec{r}_{ij}$ can only be perpendicular to $\vec{r}_{ij}$ and therefore perpendicular to $\vec{F}_{ij}$ and as a consequence the internal forces do no work and the internal energy remains constant. In this case the forces $\vec{F}_{ij}$ are precisely the forces of constraints, i.e. the forces of constraint do no work.

We consider infinitesimal virtual displacements $\delta\vec{r}_i$ which are consistent with the forces of constraints imposed on the system at time t. A virtual displacement $\delta\vec{r}_i$ is to be compared with a real displacement $d\vec{r}_i$ which occurs during a time interval dt. Thus during a real displacement the forces and constraints imposed on the system may change. To be more precise, an actual displacement is given in general by the equation

$$d\vec{r}_i = \frac{\partial\vec{r}_i}{\partial t}dt + \sum_{j=1}^{3N-k}\frac{\partial\vec{r}_i}{\partial q_j}dq_j. \tag{B.16}$$

A virtual displacement is given on the other hand by an equation of the form

$$\delta\vec{r}_i = \sum_{j=1}^{3N-k}\frac{\partial\vec{r}_i}{\partial q_j}\delta q_j. \tag{B.17}$$

The effective force on each particle is zero, i.e. $\vec{F}_{i\ \mathrm{eff}} = \vec{F}_i - \frac{d\vec{p}_i}{dt} = 0$. The virtual work of this effective force in the displacement $\delta\vec{r}_i$ is therefore trivially zero. Summed over all particles we get

$$\sum_i\left(\vec{F}_i - \frac{d\vec{p}_i}{dt}\right)\delta\vec{r}_i = 0. \tag{B.18}$$

We decompose the force $\vec{F}_i$ into the applied force $\vec{F}_i^{(a)}$ and the force of constraint $\vec{f}_i$, viz $\vec{F}_i = \vec{F}_i^{(a)} + \vec{f}_i$. Thus we have

$$\sum_i\left(\vec{F}_i^{(a)} - \frac{d\vec{p}_i}{dt}\right)\delta\vec{r}_i + \sum_i\vec{f}_i\,\delta\vec{r}_i = 0. \tag{B.19}$$

We restrict ourselves to those systems for which the net virtual work of the forces of constraints is zero. In fact, virtual displacements which are consistent with the constraints imposed on the system are precisely those displacements which are perpendicular to the forces of constraints in such a way that the net virtual work of the forces of constraints is zero. We get then

$$\sum_i\left(\vec{F}_i^{(a)} - \frac{d\vec{p}_i}{dt}\right)\delta\vec{r}_i = 0. \tag{B.20}$$

This is the principle of virtual work of D'Alembert. The forces of constraints, which as we have said are generally unknown but only their effect on the motion is known, do not appear explicitly in D'Alembert principle which is our goal. Their only effect in the equation is to make the virtual displacements $\delta\vec{r}_i$ not all independent.

B.2 Lagrange's equations

We compute

$$\begin{aligned}\sum_i \vec{F}_i^{(a)}\delta\vec{r}_i &= \sum_{i,j}\vec{F}_i^{(a)}\frac{\partial\vec{r}_i}{\partial q_j}\delta q_j\\ &= \sum_j Q_j\delta q_j.\end{aligned}\tag{B.21}$$

The Q_j are the components of the generalized force. They are defined by

$$Q_j = \sum_i \vec{F}_i^{(a)}\frac{\partial\vec{r}_i}{\partial q_j}.\tag{B.22}$$

Let us note that since the generalized coordinates q_i need not have the dimensions of length the components Q_i of the generalized force need not have the dimensions of force.

We also compute

$$\begin{aligned}\sum_i \frac{d\vec{p}_i}{dt}\delta\vec{r}_i &= \sum_{i,j} m_i\frac{d^2\vec{r}_i}{dt^2}\frac{\partial\vec{r}_i}{\partial q_j}\delta q_j\\ &= \sum_{i,j} m_i\left[\frac{d}{dt}\left(\frac{d\vec{r}_i}{dt}\frac{\partial\vec{r}_i}{\partial q_j}\right) - \frac{d\vec{r}_i}{dt}\frac{d}{dt}\left(\frac{\partial\vec{r}_i}{\partial q_j}\right)\right]\delta q_j\\ &= \sum_{i,j} m_i\left[\frac{d}{dt}\left(\vec{v}_i\frac{\partial\vec{r}_i}{\partial q_j}\right) - \vec{v}_i\frac{\partial\vec{v}_i}{\partial q_j}\right]\delta q_j.\end{aligned}\tag{B.23}$$

By using the result $\frac{\partial\vec{v}_i}{\partial\dot{q}_j} = \frac{\partial\vec{r}_i}{\partial q_j}$ we obtain

$$\begin{aligned}\sum_i \frac{d\vec{p}_i}{dt}\delta\vec{r}_i &= \sum_{i,j} m_i\left[\frac{d}{dt}\left(\vec{v}_i\frac{\partial\vec{v}_i}{\partial\dot{q}_j}\right) - \vec{v}_i\frac{\partial\vec{v}_i}{\partial q_j}\right]\delta q_j\\ &= \sum_j\left[\frac{d}{dt}\left(\frac{\partial T}{\partial\dot{q}_j}\right) - \frac{\partial T}{\partial q_j}\right]\delta q_j.\end{aligned}\tag{B.24}$$

The total kinetic term is $T = \sum_i \frac{1}{2}m_i v_i^2$. Hence D'Alembert's principle becomes

$$\sum_i\left(\vec{F}_i^{(a)} - \frac{d\vec{p}_i}{dt}\right)\delta\vec{r}_i = -\sum_j\left[Q_j - \frac{d}{dt}\left(\frac{\partial T}{\partial\dot{q}_j}\right) + \frac{\partial T}{\partial q_j}\right]\delta q_j = 0.\tag{B.25}$$

Since the generalized coordinates q_i for holonomic constraints can be chosen such that they are all independent we get the equations of motion

$$-Q_j + \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = 0. \tag{B.26}$$

Above $j = 1, \dots, n$ where $n = 3N - k$ is the number of independent generalized coordinates. For conservative forces we have $\vec{F}_i^{(a)} = -\vec{\nabla}_i V$, i.e.

$$Q_j = -\frac{\partial V}{\partial q_j}. \tag{B.27}$$

Hence we get the equations of motion

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0. \tag{B.28}$$

These are Lagrange's equations of motion where the Lagrangian L is defined by

$$L = T - V. \tag{B.29}$$

B.3 Hamilton's principle: the principle of least action

In the previous section we have derived Lagrange's equations from considerations involving virtual displacements around the instantaneous state of the system using the differential principle of D'Alembert. In this section we will rederive Lagrange's equations from considerations involving virtual variations of the entire motion between times t_1 and t_2 around the actual entire motion between t_1 and t_2 using the integral principle of Hamilton.

The instantaneous state or configuration of the system at time t is described by the n generalized coordinates $q_1, q_2, \dots, q_n$. This is a point in the n-dimensional configuration space with axes given by the generalized coordinates q_i. As time evolves the system changes and the point $(q_1, q_2, \dots, q_n)$ moves in configuration space, tracing out a curve called the path of motion of the system.

Hamilton's principle is less general than D'Alembert's principle in that it describes only systems in which all forces (except the forces of constraints) are derived from generalized scalar potentials U. The generalized potentials are velocity-dependent potentials which may also depend on time, i.e. $U = U(q_i, \dot{q}_i, t)$. The generalized forces are obtained from U as

$$Q_j = -\frac{\partial U}{\partial q_j} + \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_j}\right). \tag{B.30}$$

Such systems are called monogenic where Lagrange's equations of motion will still hold with Lagrangians given by $L = T - U$. The systems become conservative if the potentials depend only on coordinates. We define the action between times t_1 and t_2 by the line integral

$$S[q] = \int_{t_1}^{t_2} Ldt, \quad L = T - V. \tag{B.31}$$

The Lagrangian is a function of the generalized coordinates and velocities q_i and $\dot{q}_i$ and of time t, i.e. $L = L(q_1, q_2, \dots, q_n, \dot{q}_1, \dot{q}_2, \dots, \dot{q}_n, t)$. The action I is a functional.

Hamilton's principle can be stated as follows. The line integral I has a stationary value, i.e. it is an extremum for the actual path of the motion. Therefore, any first-order variation of the actual path results in a second-order change in I so that all neighboring paths which differ from the actual path by infinitesimal displacements have the same action. This is a variational problem for the action functional which is based on one single function which is the Lagrangian. Clearly I is invariant to the system of generalized coordinates used to express L and as a consequence the equations of motion, which will be derived from I, will be covariant. We write Hamilton's principle as follows

$$\frac{\delta}{\delta q_i} S[q] = \frac{\delta}{\delta q_i} \int_{t_1}^{t_2} L(q_1, q_2, \dots, q_n, \dot{q}_1, \dot{q}_2, \dots, \dot{q}_n, t)dt. \tag{B.32}$$

For systems with holonomic constraints it can be shown that Hamilton's principle is a necessary and sufficient condition for Lagrange's equations. Thus we can take Hamilton's principle as the basic postulate of mechanics rather than Newton's laws when all forces (except the forces of constraints) are derived from potentials which can depend on the coordinates, velocities and time.

Let us denote the solutions of the extremum problem by $q_i(t, 0)$. We write any other path around the correct path $q_i(t, 0)$ as $q_i(t, \alpha) = q_i(t, 0) + \alpha\eta_i(t)$ where the η_i are arbitrary functions of t which must vanish at the end points t_1 and t_2 and are continuous through the second derivative and α is an infinitesimal parameter which labels the set of neighboring paths which have the same action as the correct path. For this parametric family of curves the action becomes an ordinary function of α given by

$$S(\alpha) = \int_{t_1}^{t_2} L(q_i(t, \alpha), \dot{q}_i(t, \alpha), t)dt. \tag{B.33}$$

We define the virtual displacements δq_i by

$$\delta q_i = \left(\frac{\partial q_i}{\partial \alpha}\right)|_{\alpha=0} d\alpha = \eta_i d\alpha. \tag{B.34}$$

Similarly the infinitesimal variation of I is defined by

$$\delta S = \left(\frac{dS}{d\alpha}\right)|_{\alpha=0} d\alpha. \tag{B.35}$$

We compute

$$
\begin{aligned}
\frac{dS}{d\alpha} &= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_i} \frac{\partial q_i}{\partial \alpha} + \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial \alpha} \right) dt \\
&= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_i} \frac{\partial q_i}{\partial \alpha} + \frac{\partial L}{\partial \dot{q}_i} \frac{\partial}{\partial t} \frac{\partial q_i}{\partial \alpha} \right) dt \\
&= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_i} \frac{\partial q_i}{\partial \alpha} + \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \frac{\partial q_i}{\partial \alpha} \right) dt \\
&= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_i} \frac{\partial q_i}{\partial \alpha} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \frac{\partial q_i}{\partial \alpha} \right) dt + \left(\frac{\partial L}{\partial \dot{q}_i} \frac{\partial q_i}{\partial \alpha} \right)_{t_1}^{t_2}.
\end{aligned} \tag{B.36}
$$

The last term vanishes since all varied paths pass through the points $(t_1, y_i(t_1, 0))$ and $(t_2, y_i(t_2, 0))$. Thus we get

$$
\delta S = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right) \delta q_i dt. \tag{B.37}
$$

Hamilton's principle reads

$$
\frac{\delta S}{d\alpha} = \left(\frac{dS}{d\alpha} \right) \Big|_{\alpha=0} = 0. \tag{B.38}
$$

This leads to the equations of motion

$$
\int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right) \eta_i dt = 0. \tag{B.39}
$$

This should hold for any set of functions η_i. Thus by the fundamental lemma of the calculus of variations we must have

$$
\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = 0. \tag{B.40}
$$

Formally we write Hamilton's principle as

$$
\frac{\delta S}{\delta q_i} = \frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = 0. \tag{B.41}
$$

These are Lagrange's equations.

B.4 The Hamilton equations of motion

Again we will assume that the constraints are holonomic and the forces are monogenic, i.e. they are derived from generalized scalar potentials as in equation (B.30). For a system with n degrees of freedom we have n Lagrange's equations of motion. Since Lagrange's equations are second-order differential equations the motion of the system can be completely determined only after we also supply $2n$ initial conditions. As an example of initial conditions we can provide the n q_is and the n $\dot{q}_i$'s at an initial time t_0.

In the Hamiltonian formulation we want to describe the motion of the system in terms of first-order differential equations. Since the number of initial conditions must remain $2n$ the number of first-order differential equation which are needed to describe the system must be equal $2n$, i.e. we must have $2n$ independent variables. It is only natural to choose the first half of the $2n$ independent variables to be the n generalized coordinates q_i. The second half will be chosen to be the n generalized momenta p_i defined by

$$p_i = \frac{\partial L(q_j, \dot{q}_j, t)}{\partial \dot{q}_i}. \tag{B.42}$$

The pairs (q_i, p_i) are known as canonical variables. The generalized momenta p_i are also known as canonical or conjugate momenta.

In the Hamiltonian formulation the state or configuration of the system is described by the point $(q_1, q_2, \dots, q_n, p_1, p_2, \dots, p_n)$ in the $2n$-dimensional space known as the phase space of the system with axes given by the generalized coordinates and momenta q_i and p_i. The $2n$ first-order differential equations will describe how the point $(q_1, q_2, \dots, q_n, p_1, p_2, \dots, p_n)$ moves inside the phase space as the configuration of the system evolves in time.

The transition from the Lagrangian formulation to the Hamiltonian formulation corresponds to the change of variables $(q_i, \dot{q}_i, t) \longrightarrow (q_i, p_i, t)$ which is an example of a Legendre transformation. Instead of the Lagrangian which is a function of q_i, $\dot{q}_i$ and t, viz $L = L(q_i, \dot{q}_i, t)$ we will work in the Hamiltonian formulation with the Hamiltonian H which is a function of q_i, p_i and t defined by

$$H(q_i, p_i, t) = \sum_i \dot{q}_i p_i - L(q_i, \dot{q}_i, t). \tag{B.43}$$

We compute on one hand

$$dH = \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial t} dt. \tag{B.44}$$

On the other hand we compute

$$\begin{aligned} dH &= \dot{q}_i dp_i + p_i d\dot{q}_i - \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial t} dt \\ &= \dot{q}_i dp_i - \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial t} dt \\ &= \dot{q}_i dp_i - \dot{p}_i dq_i - \frac{\partial L}{\partial t} dt. \end{aligned} \tag{B.45}$$

By comparison we get the canonical equations of motion of Hamilton

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, -\dot{p}_i = \frac{\partial H}{\partial q_i}. \tag{B.46}$$

We also get

$$-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t}. \tag{B.47}$$

For a large class of systems and sets of generalized coordinates the Lagrangian can be decomposed as $L(q_i, \dot{q}_i, t) = L_0(q_i, t) + L_1(q_i, \dot{q}_i, t) + L_2(q_i, \dot{q}_i, t)$ where L_2 is a homogeneous function of degree 2 in $\dot{q}_i$, whereas L_1 is a homogeneous function of degree 1 in $\dot{q}_i$. In this case we compute

$$\dot{q}_i p_i = \dot{q}_i \frac{\partial L_1}{\partial \dot{q}_i} + \dot{q}_i \frac{\partial L_2}{\partial \dot{q}_i} = L_1 + 2L_2. \tag{B.48}$$

Hence

$$H = L_2 - L_0. \tag{B.49}$$

If the transformation equations which define the generalized coordinates do not depend on time explicitly, i.e. $\vec{r}_i = \vec{r}_i(q_1, q_2, \ldots, q_n)$ then $\vec{v}_i = \sum_j \frac{\partial \vec{r}_i}{\partial q_j} \dot{q}_j$ and as a consequence $T = T_2$ where T_2 is a function of q_i and $\dot{q}_i$ which is quadratic in the $\dot{q}_i$'s. In general, the kinetic term will be of the form $T = T_2(q_i, \dot{q}_i, t) + T_1(q_i, \dot{q}_i, t) + T_0(q_i, t)$. Further, if the potential does not depend on the generalized velocities $\dot{q}_i$ then $L_2 = T$, $L_1 = 0$ and $L_0 = -V$. Hence we get

$$H = T + V. \tag{B.50}$$

This is the total energy of the system. It is not difficult to show using Hamilton's equations that $\frac{dH}{dt} = \frac{\partial H}{\partial t}$. Thus if V does not depend on time explicitly then L will not depend on time explicitly and as a consequence H will be conserved.

B.5 Canonical transformations

A change of coordinates in configuration space is given by $q_i \longrightarrow Q_i = Q_i(q_i, t)$. This is known as a point transformation. A change of coordinates in phase space is given by $q_i \longrightarrow Q_i = Q_i(q_j, p_j, t)$ and $p_i \longrightarrow P_i = P_i(q_j, p_j, t)$. The q_i's and p_i's are assumed to solve Hamilton's equations of motion, i.e.

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, -\dot{p}_i = \frac{\partial H}{\partial q_i}. \tag{B.51}$$

These equations can be derived from the *modified* Hamilton's principle

$$\delta \int_{t_1}^{t_2} (p_i \dot{q}_i - H(q, p, t)) = 0. \tag{B.52}$$

The transformation $q_i \longrightarrow Q_i = Q_i(q_j, p_j, t)$, $p_i \longrightarrow P_i = P_i(q_j, p_j, t)$ is known as a canonical transformation if the new Q_i's and P_i's are canonical variables. This means that there must exist a function $K(Q, P, t)$ such that

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, -\dot{P}_i = \frac{\partial K}{\partial Q_i}. \tag{B.53}$$

Clearly these equations can also be derived from a *modified* Hamilton's principle given by

$$\delta \int_{t_1}^{t_2} (P_i \dot{Q}_i - K(Q, P, t)) = 0. \tag{B.54}$$

Thus one must have

$$\delta \int_{t_1}^{t_2} (p_i \dot{q}_i - H(q, p, t)) = \delta \int_{t_1}^{t_2} (P_i \dot{Q}_i - K(Q, P, t)) = 0. \tag{B.55}$$

Or equivalently

$$\lambda(p_i \dot{q}_i - H(q, p, t)) = P_i \dot{Q}_i - K(Q, P, t) + \frac{dF}{dt}. \tag{B.56}$$

The transformations of canonical coordinates for which $\lambda \neq 1$ are called extended canonical transformations. The transformations for which $\lambda = 1$ are called canonical transformations. Thus canonical transformations are such that

$$p_i \dot{q}_i - H(q, p, t) = P_i \dot{Q}_i - K(Q, P, t) + \frac{dF}{dt}. \tag{B.57}$$

The canonical transformations which do not depend on time explicitly, viz $Q_i = Q_i(q_j, p_j)$ and $P_i = P_i(q_j, p_j)$ are called restricted canonical transformations.

By a scale transformation such as $Q_i \longrightarrow Q_i' = \mu Q_i$, $P_i \longrightarrow P_i' = \nu P_i$ we obtain $\mu\nu(P_i \dot{Q}_i - K) = P_i' \dot{Q}_i' - K'$, i.e. $K' = \mu\nu K$. Thus, any extended canonical transformation $q_i \longrightarrow Q_i'$, $p_i \longrightarrow P_i'$ with $\lambda \neq 1$, i.e. $\lambda(p_i \dot{q}_i - H(q, p, t)) = P_i' \dot{Q}_i' - K'(Q', P', t) + \frac{dF'}{dt}$ can be composed of the canonical transformation $q_i \longrightarrow Q_i$, $p_i \longrightarrow P_i$ given by equation (B.57) followed by a scale transformation $Q_i \longrightarrow Q_i' = \mu Q_i$, $P_i \longrightarrow P_i' = \nu P_i$ with $\mu\nu = \lambda$ and $F' = \mu\nu F$.

The function F is a function of the phase space coordinates q_i, Q_i, p_i and P_i and time with continuous second derivatives. By using $Q_i = Q_i(q_j, p_j, t)$ and $P_i = P_i(q_j, p_j, t)$ and their inverses we can express F in terms partly of half of the old set of canonical variables and partly of half of the new set of canonical variables. Assuming that this can be done, the function F will act precisely as the generating function of the canonical transformation. We consider in some detail the following two general types of generating functions

$$F = F_1(q_i, Q_i, t). \tag{B.58}$$

$$F = F_2(q, P, t) - Q_i P_i. \tag{B.59}$$

In the first case we compute

$$p_i \dot{q}_i - H = P_i \dot{Q}_i - K + \frac{\partial F_1}{\partial t} + \frac{\partial F_1}{\partial q_i}\dot{q}_i + \frac{\partial F_1}{\partial Q_i}\dot{Q}_i. \tag{B.60}$$

Since q_i and Q_i are separately independent we must have

$$p_i = \frac{\partial F_1}{\partial q_i}, \quad P_i = -\frac{\partial F_1}{\partial Q_i}. \tag{B.61}$$

$$K = H + \frac{\partial F_1}{\partial t}. \tag{B.62}$$

In the second case we compute

$$p_i \dot{q}_i - H = -Q_i \dot{P}_i - K + \frac{\partial F_2}{\partial t} + \frac{\partial F_2}{\partial q_i}\dot{q}_i + \frac{\partial F_2}{\partial P_i}\dot{P}_i. \tag{B.63}$$

Again since q_i and P_i are separately independent we must have

$$p_i = \frac{\partial F_2}{\partial q_i}, \quad Q_i = \frac{\partial F_2}{\partial P_i}. \tag{B.64}$$

$$K = H + \frac{\partial F_2}{\partial t}. \tag{B.65}$$

There are two more general types of generating functions given by

$$F = F_3(p_i, Q_i, t) + q_i p_i. \tag{B.66}$$

$$F = F_4(p_i, P_i, t) + q_i p_i - Q_i P_i. \tag{B.67}$$

B.6 Poisson brackets

For restricted canonical transformations the generating function does not depend on time explicitly and as a consequence $K = H$. Let η be the $2n$-dimensional column vector constructed out of q_i and p_i and ξ be the $2n$-dimensional column vector constructed out of Q_i and P_i. The equations of restricted canonical transformations $Q_i = Q_i(q_j, p_j)$ and $P_i = P_i(q_j, p_j)$ can be rewritten as $\xi = \xi(\eta)$. The Hamilton's equations of motion in the η variables read

$$\dot{\eta} = J\frac{\partial H}{\partial \eta}. \tag{B.68}$$

The $2n \times 2n$ matrix J is given explicitly by

$$J = \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}. \tag{B.69}$$

Similarly, the Hamilton's equations of motion in the ξ variables read

$$\dot{\xi} = J\frac{\partial H}{\partial \xi}. \tag{B.70}$$

We define the matrix M by

$$M_{ij} = \frac{\partial \xi_i}{\partial \eta_j}. \tag{B.71}$$

We have

$$\begin{aligned} \dot{\xi}_i &= M_{ij}\dot{\eta}_j \\ &= M_{ij}J_{jk}\frac{\partial H}{\partial \eta_k} \\ &= M_{ij}J_{jk}M_{lk}\frac{\partial H}{\partial \xi_l} \\ &= (MJM^T)_{il}\frac{\partial H}{\partial \xi_l}. \end{aligned} \tag{B.72}$$

We must then have

$$MJM^T = J. \tag{B.73}$$

This is the symplectic condition. The matrix M is a symplectic matrix. The symplectic condition is a necessary and sufficient condition for all canonical transformations, even those which depend explicitly on time. Further, the symplectic condition implies the existence of a generating function. The symplectic or the generator formalisms can be used to show that the set of all canonical transformations form a group.

Let us introduce infinitesimal canonical transformations. First we note that $F_2 = q_i P_i$ generates the canonical transformation which acts as the identity. Indeed, this transformation gives $Q_i = q_i$, $P_i = p_i$ and $K = H$. An infinitesimal canonical transformation corresponds to

$$F_2 = q_i P_i + \epsilon G(q_j, P_j, t). \tag{B.74}$$

We compute $P_i = p_i - \epsilon\frac{\partial G}{\partial q_i}$, $Q_i = q_i + \epsilon\frac{\partial G}{\partial P_i} = q_i + \epsilon\frac{\partial G}{\partial p_i}$. In other words, we can think of G as a function of q and p (instead of q and P) and time. The function G is called the generating function of the infinitesimal canonical transformation. We write $\delta p_i = P_i - p_i = -\epsilon\frac{\partial G}{\partial q_i}$, $\delta q_i = Q_i - q_i = \epsilon\frac{\partial G}{\partial p_i}$ in a compact form as

$$\delta\eta = \epsilon J \frac{\partial G}{\partial \eta}. \tag{B.75}$$

We also introduce the notion of Poisson brackets. The Poisson bracket of any two functions u and v with respect to the canonical variables q_i and p_i is defined by

$$\begin{aligned} [u, v]_\eta &= \sum_i \left(\frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial v}{\partial q_i} \right) \\ &= \left(\frac{\partial u}{\partial \eta} \right)^T J \frac{\partial v}{\partial \eta}. \end{aligned} \tag{B.76}$$

We compute

$$\begin{aligned} [u, v]_\eta &= \frac{\partial u}{\partial \eta_i} J_{ij} \frac{\partial v}{\partial \eta_j} \\ &= \frac{\partial u}{\partial \xi_k} \frac{\partial \xi_k}{\partial \eta_i} J_{ij} \frac{\partial \xi_l}{\partial \eta_j} \frac{\partial v}{\partial \xi_l} \\ &= \frac{\partial u}{\partial \xi_k} (MJM^T)_{kl} \frac{\partial v}{\partial \xi_l} \\ &= [u, v]_\xi. \end{aligned} \tag{B.77}$$

In other words, the Poisson brackets are canonical invariant. This is the single most important property of Poisson brackets. We also write down the fundamental Poisson brackets

$$[\eta, \eta]_\eta = J. \tag{B.78}$$

In components we have

$$[q_i, q_j]_\eta = 0, \quad [p_i, p_j]_\eta = 0, \quad [q_i, p_j]_\eta = \delta_{ij}. \tag{B.79}$$

Let u be some function of the canonical variables q_i, p_i and time, i.e. $u = u(q_i, p_i, t)$. The total time derivative of u is given by

$$\begin{aligned} \frac{du}{dt} &= \sum_i \left(\frac{\partial u}{\partial q_i} \dot{q}_i + \frac{\partial u}{\partial p_i} \dot{p}_i \right) + \frac{\partial u}{\partial t} \\ &= \sum_i \left(\frac{\partial u}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial H}{\partial q_i} \right) + \frac{\partial u}{\partial t} \\ &= [u, H]_\eta + \frac{\partial u}{\partial t}. \end{aligned} \tag{B.80}$$

This is the equation of motion of the function u. Hamilton's equation (B.68) can be obtained as a special case. Indeed, if we choose $u = q_i, p_i$ then $\dot{q}_i = [q_i, H]_\eta$,

$\dot{p}_i = [p_i, H]_\eta$. In symplectic notation these equations can be rewritten as $\dot{\eta} = [\eta, H]_\eta = J\frac{\partial H}{\partial \eta}$ which is Hamilton's equation of motion (B.68).

The infinitesimal canonical transformation (B.75) can also be expressed in terms of Poisson brackets. By choosing $u = \eta$ and $v = G$ in equation (B.76) we get $[\eta, G]_\eta = J\frac{\partial G}{\partial \eta}$. The infinitesimal canonical transformation (B.75) can then be put in the form

$$\delta\eta = \epsilon[\eta, G]_\eta. \tag{B.81}$$

Let us choose $\epsilon = dt$ and $G = H$ then $\delta\eta = \dot{\eta}dt = d\eta$. In other words, the Hamiltonian is the generator of the evolution of the system in time. As a second example let us choose $\epsilon = dx$ and $G = p_j$ then $\delta q_i = dx[q_i, p_j]_\eta = \delta_{ij}dx$ and $\delta p_i = dx[p_i, p_j]_\eta = 0$ and as a consequence translation in the jth direction is generated by the momentum p_j.

Finally, we note that canonical transformations can be understood either passively or actively. In the passive view of a canonical transformation we change from the phase space η with coordinates q_i and p_i to the phase space ξ with coordinates Q_i and P_i. Thus the system at some time t which is described by the configuration $A = (q_i, p_i)$ can also be described by the transformed configuration $A' = (Q_i, P_i)$. In other words, any function u of the system variables should have the same value in the two phase spaces, i.e. $u(A) = u(A')$ although the functional dependence of u on q_i and p_i is in general different from its functional dependence on Q_i and P_i.

In the active interpretation of a canonical transformation the coordinates Q_i and P_i should be thought of as the coordinates of a point B in the same phase space as the point A. Thus the canonical transformation moves the system point from $A = (q_i, p_i)$ to $B = (Q_i, P_i)$ in the sense that it re-expresses the configuration B in terms of the configuration A and vice versa. Hence, under this view the value of a function u of the system variables will change when we go from A to B although in this case the functional dependence is the same. The change ∂u in the value of the function when we go from A to B is

$$\begin{aligned}\partial u &= u(B) - u(A)\\ &= u(\eta + \delta\eta) - u(\eta)\\ &= \frac{\partial u}{\partial \eta}\delta\eta\\ &= \epsilon\frac{\partial u}{\partial \eta}J\frac{\partial G}{\partial \eta}\\ &= \epsilon[u, G]_\eta.\end{aligned} \tag{B.82}$$

For the Hamiltonian the situation is more involved. Even under the passive view of a canonical transformation the Hamiltonian will change from $H(A)$ to $K(A')$ as we go from A to A' where $K = H + \frac{\partial F_2}{\partial t} = H + \epsilon\frac{\partial G}{\partial t}$. In this case ∂H will be given by the difference in the value of the Hamiltonian under the two interpretations, viz

$$\begin{aligned}\partial H &= (H(B) - H(A)) - (K(A') - H(A)) \\ &= H(B) - K(A') \\ &= H(B) - H(A') - \epsilon\frac{\partial G}{\partial t} \\ &= H(B) - H(A) - \epsilon\frac{\partial G}{\partial t} \\ &= \epsilon[H,\, G]_\eta - \epsilon\frac{\partial G}{\partial t} \\ &= -\,\epsilon\frac{dG}{dt}.\end{aligned} \tag{B.83}$$

The crucial conclusion is that if G is a constant of the motion then G will generate an infinitesimal canonical transformation which does not change the value of the Hamiltonian, i.e. it leaves the Hamiltonian invariant.

B.7 Hamilton–Jacobi equation

We consider a canonical transformation from $(q_i,\, p_i)$ to $(Q_i,\, P_i)$ where Q_i and P_i are constant in time, i.e. $Q_i = \beta_i$ and $P_i = \alpha_i$. This can be achieved by requiring that the transformed Hamiltonian $K(Q,\, P,\, t)$ vanishes identically. Since $K(Q,\, P,\, t) = H(q,\, p,\, t) + \frac{\partial F}{\partial t}$ we must then have

$$H(q,\, p,\, t) + \frac{\partial F}{\partial t} = 0. \tag{B.84}$$

We take $F = F_2(q_i,\, P_i,\, t)$. Since $p_i = \frac{\partial F_2}{\partial q_i}$ we can write the above action as

$$H\left(q_1,\, q_2,\, \dots\,,\, q_n,\, \frac{\partial F_2}{\partial q_1},\, \frac{\partial F_2}{\partial q_2},\, \dots\,,\, \frac{\partial F_2}{\partial q_n},\, t\right) + \frac{\partial F_2}{\partial t} = 0. \tag{B.85}$$

This is the Hamilton–Jacobi equation. It is a partial differential equation in the $n + 1$ variables $q_1,\, \dots\,,\, q_n$ and t for the generating function F_2. We denote the solution by $F_2 = S = S(q_1,\, \dots\,,\, q_n,\, \alpha_1,\, \dots\,,\, \alpha_n,\, \alpha_{n+1},\, t)$ and call it Hamilton's principal function. The $n + 1$ numbers α_i are the constants of integration. Clearly if S is a solution then $S + \alpha$ is also a solution. In other words, one of the constants of integration is irrelevant since it appears only additively and thus will drop from the partial derivatives. Further, we are at liberty to choose the new n momenta P_i which are constants such that $P_i = \alpha_i$. A complete solution of the above first-order partial differential equation is therefore given by

$$F_2 = S = S(q_1,\, \dots\,,\, q_n,\, P_1,\, \dots\,,\, P_n,\, t). \tag{B.86}$$

From $p_i = \frac{\partial F_2}{\partial q_i} = \frac{\partial S(q,\alpha,t)}{\partial q_i}$ at time t_0 we can fix α_i in terms of the initial values of q_i and p_i, whereas from $Q_i = \frac{\partial F_2}{\partial P_i} = \frac{\partial S(q,\alpha,t)}{\partial \alpha_i} = \beta_i$ at time $t = t_0$ we can determine β_i in terms of α_i and the initial values of q_i. We can then invert $\frac{\partial S(q,\alpha,t)}{\partial \alpha_i} = \beta_i$ to provide the q_i in terms of α_i, β_i and time, viz $q_i = q_i(\alpha, \beta, t)$. Then by substituting $q_i = q_i(\alpha, \beta, t)$ in $p_i = \frac{\partial S(q,\alpha,t)}{\partial q_i}$ we can find the p_i in terms of α_i, β_i and time, viz $p_i = p_i(\alpha, \beta, t)$.

Therefore, we conclude that finding Hamilton's principal function $S = S(q, \alpha, t)$ through solving the Hamilton–Jacobi equation is equivalent to finding a solution to the original Hamilton's equations of motion. We also compute

$$\begin{aligned} \frac{dS}{dt} &= \frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial t} \\ &= p_i q_i - H \\ &= L. \end{aligned} \tag{B.87}$$

In other words, S is essentially the action, viz

$$S = \int L dt + \text{constant}. \tag{B.88}$$

If the Hamiltonian does not depend on time explicitly then the Hamilton–Jacobi equation will read

$$H\left(q_i, \frac{\partial S}{\partial q_i}\right) + \frac{\partial S}{\partial t} = 0. \tag{B.89}$$

We can then separate time by writing

$$S(q_i, \alpha_i, t) = W(q_i, \alpha_i) - \alpha_1 t. \tag{B.90}$$

The Hamilton–Jacobi equation reduces to

$$H\left(q_i, \frac{\partial W}{\partial q_i}\right) = \alpha_1. \tag{B.91}$$

The function W is known as Hamilton's characteristic function. It generates a canonical transformation in which all new coordinates are cyclic, i.e. they do not appear in the transformed Hamiltonian. Indeed, let us consider the canonical transformation $(q_i, p_i) \longrightarrow (Q_i, P_i)$ where the new momenta P_i are constants of the motion α_i and with a generating function $W(q_i, P_i)$ which does not depend explicitly on time and hence $K(Q_i, P_i) = H(q_i, p_i)$. Let $H(q_i, p_i)$ be equal to the constant of the motion α_1. As before we must have $p_i = \frac{\partial W}{\partial q_i}$ and $Q_i = \frac{\partial W}{\partial P_i} = \frac{\partial W}{\partial \alpha_i}$ and thus the requirement $H(q_i, p_i) = \alpha_1$ is identical to equation (B.91). Let us note that under this canonical transformation we have $K(Q_i, P_i) = P_1$, i.e. the transformed Hamiltonian is independent of the new coordinates Q_i so that they are all cyclic. Further, we can derive from Hamilton's equations that $Q_1 = t + \beta_1$ and $Q_i = \beta_i$ for $i \neq 1$, i.e. all new coordinates with the exception of Q_1 are constants of the motion.

Appendix C

Classical electrodynamics

C.1 Coulomb's and Gauss's laws

Electrostatics is the theory of stationary charges. Coulomb's law, together with the superposition principle, are the two main foundations of electrostatics.

Coulomb's law states that the force $\vec{F}$ on a test charge Q placed at a point P due to a stationary single point charge q a distance R away is proportional to the product of the charges qQ and inversely proportional to the square of the separation distance R^2. It is given by

$$\vec{F} = \frac{1}{4\pi\epsilon_0}\frac{qQ}{R^2}\hat{u}. \tag{C.1}$$

The force points along the line from the source charge q to the test charge Q. Let $\vec{r}$ and $\vec{r}_q$ be the position vectors of Q and q, respectively, then

$$\vec{R} = \vec{r} - \vec{r}_q = R\hat{u}. \tag{C.2}$$

The force is attractive if the two charges have opposite signs and it is repulsive if the two charges have the same sign. The permittivity of the vacuum ϵ_0 is given by

$$\epsilon_0 = 8.85 \times 10^{-12}\, \mathrm{C}^2\, N^{-1}\, m^{-2}. \tag{C.3}$$

In the case of N point charges $q_1, q_2, \ldots, q_N$ the total force $\vec{F}$ on Q is obtained using the superposition principle. It is given by

doi:10.1088/2053-2563/ab0547ch11

$$\begin{aligned}\vec{F} &= \sum_{i=1}^{N} \vec{F}_i \\ &= \frac{Q}{4\pi\epsilon_0} \sum_{i=1}^{N} \frac{q_i}{R_i^2} \hat{u}_i \\ &= Q\vec{E}.\end{aligned} \tag{C.4}$$

The vector $\vec{E}$ is the electric field of the source charges. It depends on the position vector $\vec{r}$ of the field point P and not on the test charge Q. It is given by

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N} \frac{q_i}{R_i^2} \hat{u}_i. \tag{C.5}$$

For a continuous charge distribution the sum will be replaced by an integral, viz

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{dq}{R^2} \hat{u}. \tag{C.6}$$

In this formula $\vec{R}$ is the vector from the infinitesimal source charge dq to the the field point P, i.e. $\vec{R} = \vec{r} - \vec{r}_{dq} = \vec{r} - \vec{r}\,' = R\hat{u}$. For a continuous charge distribution contained inside a volume V with a charge density ρ the above equation can be put in the form

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_V dV' \frac{\rho(\vec{r}\,')}{R^2} \hat{u}. \tag{C.7}$$

Next we compute the divergence $\vec{\nabla}\vec{E}$ of $\vec{E}$ where

$$\vec{\nabla} = \hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}. \tag{C.8}$$

First, we extend the integral in equation (C.7) to all space since the charge density ρ vanishes outside the volume V anyway. We have

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int dV' \frac{\rho(\vec{r}\,')}{R^2} \hat{u}. \tag{C.9}$$

We then compute

$$\vec{\nabla}\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int dV' \rho(\vec{r}\,') \vec{\nabla}\left(\frac{\hat{u}}{R^2}\right). \tag{C.10}$$

In spherical coordinates we have

$$\vec{\nabla}\vec{v} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 v_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta v_\theta) + \frac{1}{r\sin\theta}\frac{\partial v_\phi}{\partial\phi}. \tag{C.11}$$

We consider the vector

$$\vec{v} = \frac{\hat{r}}{r^2}. \tag{C.12}$$

We get immediately that

$$\vec{\nabla}\vec{v} = 0, \quad \text{for any } \vec{r} \neq 0. \tag{C.13}$$

The divergence theorem states

$$\int_V dV \vec{\nabla}\vec{X} = \oint_S \vec{X} \cdot \vec{dS}. \tag{C.14}$$

The closed surface S is the boundary of the volume V. We apply this theorem to the vector $\vec{X} = \vec{v}$ with S being the surface of a sphere with radius r. We get

$$\begin{aligned}\int_V dV \vec{\nabla}\vec{v} &= \oint_S \vec{v} \cdot \vec{dS} \\ &= \oint_S \frac{1}{r^2} \cdot r^2 \sin\theta d\theta d\phi \\ &= 4\pi.\end{aligned} \tag{C.15}$$

The vector $\vec{\nabla}\vec{v}/4\pi$ vanishes for all $\vec{r} \neq 0$ and its integral over any volume containing the origin is 1. This is precisely the behavior of the Dirac delta function, viz

$$\vec{\nabla}\vec{v} = 4\pi\delta^3(\vec{r}). \tag{C.16}$$

Hence

$$\begin{aligned}\vec{\nabla}\vec{E}(\vec{r}) &= \frac{1}{4\pi\epsilon_0}\int dV' \rho(\vec{r}\,')4\pi\delta^3(\vec{R}) \\ &= \frac{1}{\epsilon_0}\rho(\vec{r}).\end{aligned} \tag{C.17}$$

This is Gauss's law in differential form. We apply now the divergence theorem to the electric field $\vec{E}$. We obtain

$$\begin{aligned}\oint_S \vec{E} \cdot \vec{dS} &= \int_V dV \vec{\nabla}\vec{E} \\ &= \int_V dV \frac{1}{\epsilon_0}\rho(\vec{r}) \\ &= \frac{1}{\epsilon_0} q_{\text{enc}}.\end{aligned} \tag{C.18}$$

This is Gauss's law in integral form. The integral $\oint_S \vec{E} \cdot \vec{ds}$ is the flux of the electric field through the surface S.

Next we compute the curl of $\vec{E}$. We have

$$\vec{\nabla} \times \vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0}\int dV' \rho(\vec{r}\,')\vec{\nabla} \times \left(\frac{\hat{u}}{R^2}\right). \tag{C.19}$$

In spherical coordinates we have

$$\begin{aligned}\vec{\nabla}\times\vec{v} = & \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}(\sin\theta v_\phi) - \frac{\partial v_\theta}{\partial\phi}\right]\hat{r} + \frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial v_r}{\partial\phi} - \frac{\partial(rv_\phi)}{\partial r}\right]\hat{\theta} \\ & + \frac{1}{r}\left[\frac{\partial(rv_\theta)}{\partial r} - \frac{\partial v_r}{\partial\theta}\right]\hat{\phi}.\end{aligned} \tag{C.20}$$

We can immediately conclude that

$$\vec{\nabla}\times\frac{\hat{r}}{r^2} = 0. \tag{C.21}$$

Hence

$$\vec{\nabla}\times\vec{E}(\vec{r}) = 0. \tag{C.22}$$

Stokes' theorem states

$$\int_S d\vec{S}\cdot\vec{\nabla}\times\vec{X} = \oint_l \vec{X}\cdot d\vec{l}. \tag{C.23}$$

The closed line l is the boundary of the surface S. If we apply this theorem to the electric field $\vec{E}$ we get

$$\begin{aligned}\oint_l \vec{E}\cdot d\vec{l} &= \int_S d\vec{S}\cdot\vec{\nabla}\times\vec{E} \\ &= 0.\end{aligned} \tag{C.24}$$

C.2 Lorentz, Biot–Savart's and Ampère's laws

Magnetostatics is the theory of steady currents. The Lorentz force law and the Bito–Savart's law together with the superposition principle are three main foundations of magnetostatics.

A current at a given point in a one-dimensional wire is the charge per unit time which passes that point, viz

$$I = \frac{dq}{dt} = \lambda\frac{dl}{dt} = \lambda v. \tag{C.25}$$

A steady current is a current which is the same all along the wire, viz

$$\frac{\partial I}{\partial l} = 0. \tag{C.26}$$

By charge conservation the charge per unit time leaving a segment l is equal to the decrease per unit time of the charge inside l. In other words

$$\begin{aligned}\frac{dQ_{\text{leaving}}}{dt} &= -\frac{dQ_{\text{inside}}}{dt} \\ &= -\frac{d}{dt}\int_l \lambda dl \\ &= -\int_l \frac{\partial\lambda}{\partial t}dl.\end{aligned} \tag{C.27}$$

The charge $dQ_{\rm leaving}$ is the charge which leaves the segment l from both endpoints in a time interval dt. Let $l = [a, b]$. We write $dQ_{\rm leaving} = dQ^b_{\rm leaving} - dQ^a_{\rm leaving}$ where $dQ^b_{\rm leaving}$ is the charge which exits through the endpoint b and $-dQ^a_{\rm leaving}$ is the charge which exits through the endpoint a. Clearly we have

$$\frac{dQ_{\rm leaving}}{dt} = \int_l \frac{\partial I}{\partial l} dl. \tag{C.28}$$

Hence we get the one-dimensional continuity equation

$$\frac{\partial I}{\partial l} = -\frac{\partial \lambda}{\partial t}. \tag{C.29}$$

Thus, for a steady current we must have

$$\frac{\partial \lambda}{\partial t} = 0. \tag{C.30}$$

Thus, for a steady current the charge cannot accumulate at, or dissipate from, any point on the wire. In other words given a segment l the charge leaving l is equal to the charge entering l.

Now we generalize to three dimensions. We assume that the flow of charge is distributed throughout a three-dimensional region. Thus

$$I = \frac{dq}{dt} = \rho \frac{dV}{dt}. \tag{C.31}$$

Let $dS_\perp$ be the cross-section of an infinitesimal tube which runs parallel to the flow of charge and dI be the current in this tube. Then $dI = \rho dS_\perp dl/dt$ where dl is length of the infinitesimal tube. The quantity $\rho dS_\perp dl$ is the charge which passes in a time interval dt across any given section of the tube. Thus dl/dt is precisely the speed of the charge. The volume current density is defined by

$$J = \frac{dI}{dS_\perp} = \rho v. \tag{C.32}$$

In other words J is the current per unit area perpendicular to the flow. Clearly the volume current density is a vector

$$\vec{J} = \frac{d\vec{I}}{dS_\perp} = \rho \vec{v}. \tag{C.33}$$

The total current crossing a surface S is

$$I = \int_S \vec{J} \cdot \vec{dS} \tag{C.34}$$

Thus the total charge per unit time leaving a volume V is

$$\frac{dQ_{\rm leaving}}{dt} = \oint_S \vec{J} \cdot \vec{dS} = \int_V \vec{\nabla} \vec{J} dV. \tag{C.35}$$

The conservation of electric charge gives

$$\begin{aligned}\frac{dQ_{\text{leaving}}}{dt} &= -\frac{dQ_{\text{inside}}}{dt}\\ &= -\frac{d}{dt}\int_V \rho dV\\ &= -\int_V \frac{\partial\rho}{\partial t}dV.\end{aligned} \tag{C.36}$$

We get therefore the continuity equation

$$\vec{\nabla}\vec{J} = -\frac{\partial\rho}{\partial t}. \tag{C.37}$$

For a steady current the charge cannot accumulate at, or dissipate from, any point. This means that given a volume V the charge leaving V is equal to the charge entering V. Hence we must have

$$\frac{\partial\rho}{\partial t} = 0. \tag{C.38}$$

Thus, for a steady current the volume current density is constant throughout the current distribution in the sense that

$$\vec{\nabla}\vec{J} = 0. \tag{C.39}$$

The magnetic field due to a steady current I at a point P with position vector $\vec{r}$ is given by the Biot–Savart law:

$$\vec{B} = \frac{\mu_0}{4\pi}\int \frac{\vec{J}(\vec{r}\,') \times \hat{u}}{R^2}dV'. \tag{C.40}$$

The integration is over the region in which the volume current density $\vec{J}$ does not vanish. As before, $\vec{R} = \vec{r} - \vec{r}\,' = R\hat{u}$ where $\vec{r}\,'$ is the position vector of the infinitesimal current $\vec{J}(\vec{r}\,')da'_{\perp}$. The permeability of the vacuum μ_0 is given by

$$\mu_0 = 4\pi \times 10^{-7}N/A^2. \tag{C.41}$$

The magnetic force exerted by this magnetic field $\vec{B}$ on another volume current density $\vec{J}_0$ is given by Lorentz force law:

$$\begin{aligned}\vec{F} &= \int dq_0(\vec{v}_0 \times \vec{B})\\ &= \int \rho_0(\vec{v}_0 \times \vec{B})dV\\ &= \int (\vec{J}_0 \times \vec{B})dV.\end{aligned} \tag{C.42}$$

The integration is now over the region in which the volume current density $\vec{J}_0$ does not vanish. For a single point charge q_0 with velocity $\vec{v}_0$ the Lorentz force law reads

$$\vec{F} = q_0(\vec{v}_0 \times \vec{B}). \tag{C.43}$$

Next we compute the curl of the magnetic field $\vec{B}$. We have

$$\vec{\nabla} \times \vec{B} = \frac{\mu_0}{4\pi} \int \vec{\nabla} \times \left(\frac{\vec{J}(\vec{r}\,') \times \hat{u}}{R^2} \right) dV'. \tag{C.44}$$

Using the identities

$$\vec{\nabla} \times (\vec{A} \times \vec{B}) = (\vec{B}.\vec{\nabla})\vec{A} - (\vec{A}.\vec{\nabla})\vec{B} + \vec{A}(\vec{\nabla}.\vec{B}) - \vec{B}(\vec{\nabla}.\vec{A}). \tag{C.45}$$

$$\vec{\nabla}(f\vec{A}) = f\vec{\nabla}\vec{A} + \vec{A}\vec{\nabla}f. \tag{C.46}$$

We get (using also the fact that $\vec{J}(\vec{r}\,')$ does not depend on $\vec{r}$ and $\vec{\nabla}'\vec{J}(\vec{r}\,') = 0$)

$$\begin{aligned}
\vec{\nabla}\left(\frac{\vec{J}(\vec{r}\,') \times \hat{u}}{R^2} \right) &= -(\vec{J} \cdot \vec{\nabla})\frac{\hat{u}}{R^2} + \vec{J}\left(\vec{\nabla} \cdot \frac{\hat{u}}{R^2} \right) \\
&= (\vec{J} \cdot \vec{\nabla}')\frac{x - x'}{R^3}\hat{i} + (\vec{J} \cdot \vec{\nabla}')\frac{y - y'}{R^3}\hat{j} \\
&\quad + (\vec{J} \cdot \vec{\nabla}')\frac{z - z'}{R^3}\hat{k} + \vec{J}(4\pi\delta^3(\vec{R})) \\
&= \vec{\nabla}'\left(\vec{J}\frac{x - x'}{R^3} \right)\hat{i} + \vec{\nabla}'\left(\vec{J}\frac{y - y'}{R^3} \right)\hat{j} \\
&\quad + \vec{\nabla}'\left(\vec{J}\frac{z - z'}{R^3} \right)\hat{k} + \vec{J}(4\pi\delta^3(\vec{R}))
\end{aligned} \tag{C.47}$$

The first three terms give boundary integrals which are zero. The last term gives

$$\vec{\nabla} \times \vec{B} = \mu_0\vec{J}. \tag{C.48}$$

This is the differential form of Ampère's law. Using Stokes' theorem we have

$$\begin{aligned}
\oint_l \vec{B} \cdot \vec{dl} &= \int_S \vec{\nabla} \times \vec{B} \cdot \vec{dS} \\
&= \mu_0 \int_S \vec{J} \cdot \vec{dS} \\
&= \mu_0 I_{\text{enc}}.
\end{aligned} \tag{C.49}$$

The current I_{enc} is the total current passing through the surface S, i.e. the total current enclosed by the loop l which is the boundary of the surface S. This is the integral form of Ampère's law.

Similarly we compute the divergence of the magnetic field $\vec{B}$. We have

$$\vec{\nabla}\vec{B} = \frac{\mu_0}{4\pi} \int \vec{\nabla}\left(\frac{\vec{J}(\vec{r}\,') \times \hat{u}}{R^2} \right) dV'. \tag{C.50}$$

Now we use the identity

$$\vec{\nabla}(\vec{A} \times \vec{B}) = \vec{B}(\vec{\nabla} \times \vec{A}) - \vec{A}(\vec{\nabla} \times \vec{B}). \tag{C.51}$$

We get that

$$\vec{\nabla}\left(\frac{\vec{J}(\vec{r}\,') \times \hat{u}}{R^2}\right) = 0. \tag{C.52}$$

Hence

$$\vec{\nabla}\vec{B} = 0. \tag{C.53}$$

C.3 Electromagnetic induction and Faraday's laws

Electromotive force

In a closed circuit, because of resistivity (electrical friction), there must be some force which we call electromotive force or emf to maintain a steady current. An ideal source of emf will provide a constant voltage between two terminals. An example of a source of emf is a battery.

We consider an electric circuit consisting of a battery connected to a resistor. Let a and b be the negative and positive terminals, respectively, of the battery. The current I generated outside the battery will flow from the positive terminal b to the negative terminal a opposite the direction of flow of electrons. Equivalently we can pretend that actually positive charges move in the direction of the current from b to a.

The chemical force per unit charge $\vec{F}_s$ generated within the battery is directed from negative to positive terminals and it is only confined to the battery. From Ohm's law $\vec{J} = \sigma\vec{E}$ where σ is the conductivity we see that a current density is non-zero outside the battery only if an electrostatic field $\vec{E}$ exists. Therefore, there must exist outside the battery an electrostatic field $\vec{E}$ which helps to maintain the flow of the charges. The electric potential is defined for an electrostatic field such as $\vec{E}$ by

$$\vec{E} = -\vec{\nabla}V. \tag{C.54}$$

The potential difference between the terminals a and b is

$$V_+ - V_- = -\int_a^b \vec{E}\vec{dl} = \int_a^b \vec{\nabla}V\vec{dl} = \mathcal{E}. \tag{C.55}$$

Thus, when a positive charge passes from the negative terminal a to the positive terminal b within the battery its potential will increase by the amount $\mathcal{E}$. By conservation of energy the chemical energy in the battery will decrease by the amount $\mathcal{E}$. The work done per unit charge by the battery is therefore equal to $\mathcal{E}$, viz

$$\mathcal{E} = \int_a^b \vec{F}_s\vec{dl}. \tag{C.56}$$

This means in particular that within the battery $\vec{F}_s = -\vec{E}$.

This can also be seen as follows. The chemical force $\vec{F}_s$ inside the battery will cause charges to be displaced which in turn will create an electrostatic field $\vec{E}$. Thus by Ohm's law the current within the battery is $\vec{J} = \sigma(\vec{E} + \vec{F}_s)$ or equivalently $\vec{E} + \vec{F}_s = \rho\vec{J}$ where $\rho = 1/\sigma$ is the resistivity of the battery. For an ideal battery $\rho = 0$ and hence $\vec{E} + \vec{F}_s = 0$.

The quantity $\mathcal{E}$ is called the electromotive force or emf which can also be rewritten as

$$\mathcal{E} = \oint \vec{F}_s \vec{dl}. \tag{C.57}$$

We think of $\vec{F}_s$ as an electric field but it is not electrostatic since its curl is non-zero. In summary, the battery or any other source of emf will establish and maintain a constant voltage difference equal to the emf $\mathcal{E}$ between two terminals.

Motional emf

The generator is another source of emf in a circuit. The emf in this case is known as a motional emf since it arises from the motion of the circuit in a magnetic field. Let us consider a rectangular loop in the xy plane placed in a uniform magnetic field $\vec{B}$ which is pointing along the positive z direction. The circuit consists only of a resistor. The segment $cbad$ where $y_a = y_b$, $y_c = y_d$, $x_a - x_b = x_d - x_c = h$ and $y_c - y_b = y_d - y_a = s$ is in the region where $B \neq 0$. Clearly, if we decrease s by pulling the entire loop with a velocity v along the positive y direction the magnetic flux through the rectangular loop will change and as a consequence an electric current will be induced in the loop. Indeed, the magnetic force per unit charge in the segments $\vec{ba}$, $\vec{ad}$ and $\vec{cb}$ given by $\vec{F}_{\text{mag}} = \vec{v} \times \vec{B} = vB\hat{i}$ will drive a current in the segment $\vec{ba}$ and not in the segments $\vec{ad}$ and $\vec{cb}$.

The motional emf $\mathcal{E}$ is the constant voltage difference $V_a - V_b$. In other words, as a positive charge moves from b to a its potential will increase by the amount $\mathcal{E}$. Thus by conservation of energy $\mathcal{E}$ must be equal to the work of the mechanical force $\vec{F}_{\text{pull}}$ which is pulling with a velocity v, i.e.

$$\mathcal{E} = \int_b^a \vec{F}_{\text{pull}} \vec{dl}. \tag{C.58}$$

The existence of a current means that positive charges will have, in addition to the velocity $\vec{v}$, another velocity $\vec{u}$ which is always in the direction of the current. The total magnetic force is therefore $\vec{F}_{\text{mag}} = (\vec{v} + \vec{u}) \times \vec{B}$. In the segment $\vec{ba}$ the magnetic force $\vec{F}_{\text{mag}}$ will have a horizontal component given by $-uB\hat{j}$. The mechanical force which is pulling with a velocity v is therefore equal to $\vec{F}_{\text{pull}} = uB\hat{j}$. Let θ be the angle which the velocity $\vec{w} = \vec{v} + \vec{u}$ makes with the x-axis, i.e. $w\cos\theta = u$ and $w\sin\theta = v$. The actual displacement of the charges in the segment $\vec{ba}$ will be in the direction of $\vec{w}$. The integration path for the calculation of the work of $\vec{F}_{\text{pull}}$ is this displacement which makes an angle θ with the x-axis and which is of length $h/\cos\theta$. Thus the work of $\vec{F}_{\text{pull}}$ is

$$\mathcal{E} = \int_b^a \vec{F}_{\text{pull}} \vec{d}l = (uB)\frac{h}{\cos\theta}\cos\left(\frac{\pi}{2} - \theta\right) = vBh. \tag{C.59}$$

It is not difficult to check that

$$\begin{aligned}\mathcal{E} &= \oint \vec{F}_{\text{mag}} \vec{d}l \\ &= \int_b^a \vec{F}_{\text{mag}} \vec{d}l \\ &= \int vBdx \\ &= vBh.\end{aligned} \tag{C.60}$$

The motional emf $\mathcal{E}$ is not the work of $\vec{F}_{\text{mag}}$ since magnetic forces never do work. Indeed, the integration in the last equation above is done around the loop at a given instant of time.

Now we relate the motional emf with the flux of the magnetic field. The flux Φ of the magnetic field $\vec{B}$ through the loop is given by

$$\Phi = \int \vec{B}\vec{d}S = \int Bdxdy = Bhs. \tag{C.61}$$

As we decrease s the flux decreases so $d\Phi/dt$ must be negative. By using the fact that $v = -ds/dt$ since ds/dt is negative we obtain the result

$$\frac{d\Phi}{dt} = -Bhv. \tag{C.62}$$

In other words

$$\mathcal{E} = -\frac{d\Phi}{dt}. \tag{C.63}$$

This is the flux rule which applies quite generally to non-rectangular loops moving in arbitrary directions in non-uniform magnetic fields.

Transformer emf

Another source of emf is the transformer. The emf in this case may be called transformer emf. Let us consider the previous setup, only now the rectangular loop is kept stationary. Next we either move the electromagnet which created the magnetic field $\vec{B} = B\hat{k}$ with a velocity v along the negative y direction or we vary the current in the coil of the electromagnet so that the strength of the magnetic field $\vec{B}$ changes. In both cases a current will flow in the loop.

In these cases the loop is stationary and therefore the force responsible for the flow of the current is not magnetic since stationary charges cannot experience a magnetic force. Faraday concluded that there must exist an electric field $\vec{E}$ in the loop which causes the current to flow. This electric field, which was induced by changing the magnetic field, is not electrostatic. The work done by the induced electric field $\vec{E}$ around the loop is the transformer emf $\mathcal{E}$, i.e.

$$\mathcal{E} = \oint \vec{E}\vec{d}l. \tag{C.64}$$

Empirically we find that $\mathcal{E}$ is again given by the flux rule, viz

$$\mathcal{E} = -\frac{d\Phi}{dt}. \tag{C.65}$$

In other words

$$\oint \vec{E}\vec{d}l = -\int \frac{\partial\vec{B}}{\partial t}\vec{d}S. \tag{C.66}$$

This is Faraday's law in integral form. Using Stokes' theorem we obtain Faraday's law in differential form, viz

$$\vec{\nabla} \times \vec{E} = -\frac{\partial\vec{B}}{\partial t}. \tag{C.67}$$

C.4 Maxwell's equations

In summary, we have obtained the following laws:

$$\vec{\nabla}\vec{E} = \frac{\rho}{\epsilon_0}. \tag{C.68}$$

$$\vec{\nabla}\vec{B} = 0. \tag{C.69}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial\vec{B}}{\partial t}. \tag{C.70}$$

$$\vec{\nabla} \times \vec{B} = \mu_0\vec{J}. \tag{C.71}$$

However, we know that for any vector $\vec{X}$ the identity $\vec{\nabla}(\vec{\nabla} \times \vec{X}) = 0$ must hold. This identity holds for $\vec{X} = \vec{E}$. Indeed we have

$$\vec{\nabla}(\vec{\nabla} \times \vec{E}) = -\vec{\nabla}\left(\frac{\partial\vec{B}}{\partial t}\right) = -\frac{\partial}{\partial t}(\vec{\nabla} \cdot \vec{B}) = 0. \tag{C.72}$$

But for $\vec{X} = \vec{B}$ we have

$$\vec{\nabla}(\vec{\nabla} \times \vec{B}) = \vec{\nabla}(\mu_0\vec{J}) = \mu_0\vec{\nabla}\vec{J}. \tag{C.73}$$

This is zero only for steady currents, i.e. when $\vec{\nabla}\vec{J} = 0$. Therefore, either the identity $\vec{\nabla}(\vec{\nabla} \times \vec{X}) = 0$ is not true, which is simply impossible, or Ampère's law (C.71) is wrong for non-steady currents.

For non-steady currents we must use the continuity equation

$$\vec{\nabla}\vec{J} = -\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial t}(\epsilon_0 \vec{\nabla}\vec{E}) = -\vec{\nabla}\left(\epsilon_0 \frac{\partial \vec{E}}{\partial t}\right). \tag{C.74}$$

In other words

$$\vec{\nabla}\left(\mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}\right) = 0. \tag{C.75}$$

Therefore, Ampère's law must be modified such that

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}. \tag{C.76}$$

Now clearly $\vec{\nabla}(\vec{\nabla} \times \vec{B}) = 0$. The quantity $\vec{J}_D = \epsilon_0 \partial \vec{E}/\partial t$ is called the displacement current and it is generally very small compared to $\vec{J}$.

In analogy with Faraday's law (C.70) which states that a changing magnetic field induces an electric field the Ampère–Maxwell's law (C.76) states that a changing electric field induces a magnetic field. Maxwell's equations consist of Gauss's law (C.68), Faraday's law (C.70), Ampère–Maxwell's law (C.76) and equation (C.69). Together with the Lorentz force law they summarize classical electrodynamics. The continuity equation can be derived by applying the divergence to Ampère–Maxwell's law (C.76).

C.5 Electromagnetic energy and Poynting's theorem

The work done by the Lorentz force on a charge dq is

$$\begin{aligned} dW &= \vec{F} \cdot \vec{dl} \\ &= dq(\vec{E} + \vec{v} \times \vec{B}) \cdot \vec{v} dt \\ &= \rho dV \vec{E}\vec{v} dt \\ &= \vec{E}\vec{J} dV dt. \end{aligned} \tag{C.77}$$

The work per unit time done on all charges inside a volume V is

$$\frac{dW}{dt} = \int_V \vec{E}\vec{J} dV. \tag{C.78}$$

By using Ampère–Maxwell's and Faraday's laws we compute

$$\begin{aligned}\vec{E}\vec{J} &= \frac{1}{\mu_0}\vec{E}(\vec{\nabla}\times\vec{B}) - \epsilon_0\vec{E}\frac{\partial\vec{E}}{\partial t}\\ &= \frac{1}{\mu_0}\vec{B}(\vec{\nabla}\times\vec{E}) - \frac{1}{\mu_0}\vec{\nabla}(\vec{E}\times\vec{B}) - \frac{\partial}{\partial t}\left(\frac{1}{2}\epsilon_0\vec{E}^2\right)\\ &= -\frac{1}{\mu_0}\vec{\nabla}(\vec{E}\times\vec{B}) - \frac{\partial}{\partial t}\left(\frac{1}{2}\epsilon_0\vec{E}^2 + \frac{1}{2\mu_0}\vec{B}^2\right).\end{aligned} \tag{C.79}$$

Hence, by using the divergence theorem we get

$$\frac{dW}{dt} = -\frac{dU_{\rm em}}{dt} - \oint_A \vec{S}\cdot\vec{dA}. \tag{C.80}$$

The total energy stored in the electromagnetic field is $U_{\rm em}$ and it is given by

$$U_{\rm em} = \int_V \left(\frac{1}{2}\epsilon_0\vec{E}^2 + \frac{1}{2\mu_0}\vec{B}^2\right)dV. \tag{C.81}$$

The vector $\vec{S}$ is called the Poynting vector and it is defined by

$$\vec{S} = \frac{1}{\mu_0}(\vec{E}\times\vec{B}). \tag{C.82}$$

This is the energy per unit time per unit area transported by the field. Thus the Poynting vector expresses the flow of energy. The work done on the charges increases their mechanical energy $U_{\rm mech}$, i.e.

$$\frac{dW}{dt} = \frac{dU_{\rm mech}}{dt}. \tag{C.83}$$

Thus we get Poynting's equation

$$\frac{d}{dt}(U_{\rm em} + U_{\rm mech}) + \oint_A \vec{S}\cdot\vec{dA} = 0. \tag{C.84}$$

The rate of change of the total energy (mechanical energy of the charges + electromagnetic energy stored in the field) within a volume V is equal to the energy per unit time transported by the field across the surface A which encloses the volume V. This is Poynting's theorem which expresses conservation of energy. Let $u_{\rm em}$ be the energy density of the electromagnetic field and $u_{\rm mech}$ be the energy density of the charges. In other words

$$u_{\rm em} = \frac{dU_{\rm em}}{dV} = \frac{1}{2}\epsilon_0\vec{E}^2 + \frac{1}{2\mu_0}\vec{B}^2. \tag{C.85}$$

$$u_{\rm mech} = \frac{dU_{\rm mech}}{dV}. \tag{C.86}$$

Poynting's equation becomes

$$\frac{\partial}{\partial t}(u_{\text{em}} + u_{\text{mech}}) + \vec{\nabla}\vec{S} = 0. \tag{C.87}$$

C.6 Electromagnetic waves

Maxwell's equations in a vacuum read

$$\vec{\nabla}\vec{E} = 0. \tag{C.88}$$

$$\vec{\nabla}\vec{B} = 0. \tag{C.89}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}. \tag{C.90}$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}. \tag{C.91}$$

We compute

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = \vec{\nabla}(\vec{\nabla}\vec{E}) - \vec{\nabla}^2 \vec{E} = -\vec{\nabla}^2 \vec{E}. \tag{C.92}$$

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{B}) = \vec{\nabla}(\vec{\nabla}\vec{B}) - \vec{\nabla}^2 \vec{B} = -\vec{\nabla}^2 \vec{B}. \tag{C.93}$$

On the other hand

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = \vec{\nabla} \times \left(-\frac{\partial \vec{B}}{\partial t}\right) = -\frac{\partial}{\partial t}(\vec{\nabla} \times \vec{B}) = -\mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2}. \tag{C.94}$$

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{B}) = \vec{\nabla} \times \left(\mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}\right) = \mu_0 \epsilon_0 \frac{\partial}{\partial t}(\vec{\nabla} \times \vec{E}) = -\mu_0 \epsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2}. \tag{C.95}$$

Thus we get the equations

$$\left(\vec{\nabla}^2 - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2}\right)\vec{E} = 0. \tag{C.96}$$

$$\left(\vec{\nabla}^2 - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2}\right)\vec{B} = 0. \tag{C.97}$$

These are three-dimensional wave equations since they are of the form

$$\left(\vec{\nabla}^2 - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right)f = 0. \tag{C.98}$$

Thus, there exist electromagnetic waves in the vacuum propagating with a speed equal to

$$v = \frac{1}{\sqrt{\mu_0 \epsilon_0}} = 3 \times 10^8 \text{ m s}^{-1}. \tag{C.99}$$

This is precisely the speed of light.

An interesting set of solutions to the wave equations (C.96) and (C.97) is given by the set of monochromatic plane waves. A monochromatic plane wave with a frequency ω and propagating in the direction $\vec{k}$ is given by

$$\vec{E}(\vec{r}, t) = \vec{E}_0 e^{i(\vec{k}\vec{r} - \omega t)}, \quad \vec{B}(\vec{r}, t) = \vec{B}_0 e^{i(\vec{k}\vec{r} - \omega t)}. \tag{C.100}$$

These fields satisfy equations (C.96) and (C.97) provided

$$k = \frac{\omega}{c}. \tag{C.101}$$

The Maxwell's equations $\vec{\nabla}\vec{E} = \vec{\nabla}\vec{B} = 0$ lead to the constraints

$$\vec{k}\vec{E} = \vec{k}\vec{B} = 0. \tag{C.102}$$

The electric and magnetic fields are perpendicular to the directions of the propagation of the waves. We say that the electromagnetic wave is transverse. The electric and magnetic fields are themselves perpendicular to each other. Indeed we derive from the Maxwell's equation $\vec{\nabla} \times \vec{E} = -\partial \vec{B}/\partial t$ the constraint

$$\vec{B}_0 = \frac{1}{c}\hat{k} \times \vec{E}_0. \tag{C.103}$$

C.7 Potential and fields

Given any vector $\vec{X}$ we have the identity $\vec{\nabla}(\vec{\nabla} \times \vec{X}) = 0$. Therefore, Maxwell's equation $\vec{\nabla}\vec{B} = 0$ means that we can write $\vec{B}$ as

$$\vec{B} = \vec{\nabla} \times \vec{A}. \tag{C.104}$$

The vector $\vec{A}$ is called the vector potential. Putting this equation in Faraday's law yields

$$\vec{\nabla} \times \vec{E} = -\frac{\partial}{\partial t}(\vec{\nabla} \times \vec{A}). \tag{C.105}$$

This can be put into the form

$$\vec{\nabla} \times \left(\vec{E} + \frac{\partial \vec{A}}{\partial t}\right) = 0. \tag{C.106}$$

Given any function V we have the identity $\vec{\nabla} \times (\vec{\nabla} V) = 0$. Hence we can parameterize the electric field as

$$\vec{E} = -\vec{\nabla} V - \frac{\partial \vec{A}}{\partial t}. \tag{C.107}$$

The function V is called the scalar potential. With the introduction of V and $\vec{A}$ we have solved Maxwell's equations (C.69), (C.70). In terms of V and $\vec{A}$ Gauss's equation (C.68) becomes

$$\vec{\nabla}^2 V + \frac{\partial}{\partial t}\vec{\nabla}\vec{A} = -\frac{\rho}{\epsilon_0}. \tag{C.108}$$

In terms of V an $\vec{A}$ Ampère–Maxwell's equation (C.76) becomes (using also the identity $\vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla}(\vec{\nabla}\vec{A}) - \vec{\nabla}^2\vec{A}$)

$$\vec{\nabla}^2\vec{A} - \mu_0\epsilon_0\frac{\partial^2\vec{A}}{\partial t^2} - \vec{\nabla}\left(\vec{\nabla}\vec{A} + \mu_0\epsilon_0\frac{\partial V}{\partial t}\right) = -\mu_0\vec{J}. \tag{C.109}$$

The task now is to solve equations (C.108) and (C.109).

We have a gauge freedom in choosing $\vec{A}$ and V. Let us choose a new vector potential $\vec{A}'$ and a new scalar potential V' such that

$$\begin{aligned} \vec{A}' &= \vec{A} + \vec{\alpha} \\ V' &= V + \beta. \end{aligned} \tag{C.110}$$

Let us require that $\vec{B} = \vec{\nabla} \times \vec{A} = \vec{\nabla} \times \vec{A}'$. Then one must have

$$\vec{\nabla} \times \vec{\alpha} = 0. \tag{C.111}$$

In other words

$$\vec{\alpha} = \vec{\nabla}\lambda. \tag{C.112}$$

We also require $\vec{E} = -\vec{\nabla} V - \partial\vec{A}/\partial t = -\vec{\nabla} V' - \partial\vec{A}'/\partial t$. Thus we must have

$$\vec{\nabla}\beta + \frac{\partial\vec{\alpha}}{\partial t} = 0. \tag{C.113}$$

In other words

$$\vec{\nabla}\left(\beta + \frac{\partial\lambda}{\partial t}\right) = 0. \tag{C.114}$$

Hence $\beta + \partial\lambda/\partial t = f(t)$ for some function f of time. The function $f(t)$ can be absorbed in λ without changing the vector $\vec{\alpha}$. In other words we can set $f(t) = 0$ without loss of generality. Thus we get

$$\beta = -\frac{\partial\lambda}{\partial t} \tag{C.115}$$

We get therefore the gauge transformations

$$\vec{A}' = \vec{A} + \vec{\nabla}\lambda$$
$$V' = V - \frac{\partial \lambda}{\partial t}. \tag{C.116}$$

The set of potentials V and $\vec{A}$ and the set of potentials V' and $\vec{A}'$ give the same physical fields $\vec{E}$ and $\vec{B}$. In order to simplify equations (C.108) and (C.109) we can therefore choose the function λ appropriately. This is called a gauge choice.

The Coulomb gauge consists of choosing λ in such a way that the vector potential $\vec{A}$ satisfies

$$\vec{\nabla}\vec{A} = 0. \tag{C.117}$$

Equation (C.108) becomes

$$\vec{\nabla}^2 V = -\frac{\rho}{\epsilon_0}. \tag{C.118}$$

This is Poisson's equation. As will soon be clear, the solution is not causal. This is the first disadvantage of the Coulomb gauge. The second disadvantage is the fact that equation (C.109) becomes complicated in this gauge. It reads

$$\vec{\nabla}^2\vec{A} - \mu_0\epsilon_0\frac{\partial^2\vec{A}}{\partial t^2} = -\mu_0\vec{J} + \mu_0\epsilon_0\vec{\nabla}\frac{\partial V}{\partial t}. \tag{C.119}$$

The Lorentz gauge consists of choosing λ in such a way that the vector potential $\vec{A}$ satisfies

$$\vec{\nabla}\vec{A} = -\mu_0\epsilon_0\frac{\partial V}{\partial t}. \tag{C.120}$$

Equations (C.108) and (C.109) become

$$\left(\vec{\nabla}^2 - \mu_0\epsilon_0\frac{\partial^2}{\partial t^2}\right)V = -\frac{\rho}{\epsilon_0}. \tag{C.121}$$

$$\left(\vec{\nabla}^2 - \mu_0\epsilon_0\frac{\partial^2}{\partial t^2}\right)\vec{A} = -\mu_0\vec{J}. \tag{C.122}$$

The operator

$$\vec{\nabla}^2 - \mu_0\epsilon_0\frac{\partial^2}{\partial t^2} \tag{C.123}$$

is the d'Alembertian which in some sense is a generalization of the Laplacian. Thus, in the Lorentz gauge V and $\vec{A}$ solve the inhomogeneous wave equation with a source term.

For static fields we get the Poisson's equations

$$\vec{\nabla}^2 V = -\frac{\rho}{\epsilon_0}. \tag{C.124}$$

$$\vec{\nabla}^2\vec{A} = -\mu_0\vec{J}. \tag{C.125}$$

The solutions V and $\vec{A}$ for charge and current densities ρ and $\vec{J}$ which go to zero at infinity are given by (with $\vec{R} = \vec{r} - \vec{r}\,' = R\hat{u}$)

$$V(\vec{r}\,) = \frac{1}{4\pi\epsilon_0}\int_V dV'\frac{\rho(\vec{r}\,')}{R}. \tag{C.126}$$

$$\vec{A}(\vec{r}\,) = \frac{\mu_0}{4\pi}\int_V dV'\frac{\vec{J}(\vec{r}\,')}{R}. \tag{C.127}$$

The proof relies on the two identities

$$\vec{\nabla}\left(\frac{1}{r}\right) = -\frac{\hat{r}}{r^2}, \quad \vec{\nabla}\left(\frac{\hat{r}}{r^2}\right) = 4\pi\delta^3(\vec{r}\,). \tag{C.128}$$

For non-static fields the situation is more involved. The electromagnetic effect of the infinitesimal charge and infinitesimal current which exist at time t at the source point $\vec{r}\,'$ will reach the field point $\vec{r}$ only after a time R/c. This means that the scalar and vector potentials at time t will be affected by the charge and current densities at the field point $\vec{r}$ which existed at an earlier time t_r known as the retarded time. The retarded time is given by

$$t_r = t - \frac{R}{c}. \tag{C.129}$$

The solutions V and $\vec{A}$ for charge and current densities $\rho(\vec{r}, t)$ and $\vec{J}(\vec{r}, t)$ which go to zero at spatial infinity will read

$$V(\vec{r}, t) = \frac{1}{4\pi\epsilon_0}\int_V dV'\frac{\rho(\vec{r}\,', t_r)}{R}. \tag{C.130}$$

$$\vec{A}(\vec{r}, t) = \frac{\mu_0}{4\pi}\int_V dV'\frac{\vec{J}(\vec{r}\,', t_r)}{R}. \tag{C.131}$$

These are called the retarded potentials. In order to show this we write

$$V(\vec{r}, t) = \frac{1}{4\pi\epsilon_0}\int dV'\frac{\rho(\vec{r}\,', t_r)}{R}. \tag{C.132}$$

Then

$$\begin{aligned}\vec{\nabla}V(\vec{r}, t) &= \frac{1}{4\pi\epsilon_0}\int dV'\vec{\nabla}\frac{\rho(\vec{r}\,', t_r)}{R} \\ &= \frac{1}{4\pi\epsilon_0}\int dV'\left[\vec{\nabla}\rho\cdot\frac{1}{R} + \rho\vec{\nabla}\left(\frac{1}{R}\right)\right].\end{aligned} \tag{C.133}$$

We use

$$\vec{\nabla}\rho = \dot{\rho}\vec{\nabla}_R t_r = -\frac{\dot{\rho}}{c}\hat{R}. \tag{C.134}$$

Thus

$$\vec{\nabla} V(\vec{r},\, t) = \frac{1}{4\pi\epsilon_0}\int dV'\left[-\frac{1}{c}\dot{\rho}\frac{\hat{R}}{R} - \rho\frac{\hat{R}}{R^2}\right]. \tag{C.135}$$

Taking the divergence again we get

$$\begin{aligned}\vec{\nabla} V(\vec{r},\, t) = \frac{1}{4\pi\epsilon_0}\int dV'\Bigg[&-\frac{1}{c}\vec{\nabla}\dot{\rho}.\frac{\hat{R}}{R} - \frac{1}{c}\dot{\rho}\cdot\vec{\nabla}\left(\frac{\hat{R}}{R}\right)\\ &-\vec{\nabla}\rho\cdot\frac{\hat{R}}{R^2} - \rho\cdot\vec{\nabla}\left(\frac{\hat{R}}{R^2}\right)\Bigg].\end{aligned} \tag{C.136}$$

We use

$$\vec{\nabla}\left(\frac{\hat{R}}{R}\right) = \frac{1}{R^2}. \tag{C.137}$$

$$\vec{\nabla}\dot{\rho} = \ddot{\rho}\vec{\nabla}_R t_r = -\frac{\ddot{\rho}}{c}\hat{R}. \tag{C.138}$$

We get

$$\begin{aligned}\vec{\nabla} V(\vec{r},\, t) &= \frac{1}{4\pi\epsilon_0}\int dV'\left[\frac{1}{c^2}\frac{\ddot{\rho}}{R} - \frac{1}{c}\frac{\dot{\rho}}{R^2} + \frac{1}{c}\frac{\dot{\rho}}{R^2} - 4\pi\rho\delta^3(\vec{R})\right]\\ &= \frac{1}{4\pi\epsilon_0 c^2}\int dV'\frac{\ddot{\rho}}{R} - \frac{1}{\epsilon_0}\rho\\ &= \frac{1}{c^2}\frac{\partial^2 V}{\partial t^2} - \frac{1}{\epsilon_0}\rho.\end{aligned} \tag{C.139}$$

The proof for the vector potential is identical. Next we need to check that the retarded potentials satisfy the Lorentz condition. We have

$$\vec{\nabla}\vec{A}(\vec{r},\, t) = \frac{\mu_0}{4\pi}\int_V dV'\vec{\nabla}\frac{\vec{J}(\vec{r}\,',\, t_r)}{R}. \tag{C.140}$$

We use the identities

$$\vec{\nabla}\left(\frac{\vec{J}}{R}\right) + \vec{\nabla}'\left(\frac{\vec{J}}{R}\right) = \frac{1}{R}\vec{\nabla}(\vec{J}) + \frac{1}{R}\vec{\nabla}'(\vec{J}). \tag{C.141}$$

$$\vec{\nabla}(\vec{J}) = -\frac{\dot{\vec{J}}}{c}\hat{R}. \tag{C.142}$$

$$\begin{aligned}\vec{\nabla}'(\vec{J}) &= \frac{\dot{\vec{J}}}{c}\hat{R} + \vec{\nabla}'\vec{J} \\ &= \frac{\dot{\vec{J}}}{c}\hat{R} - \dot{\rho}.\end{aligned} \tag{C.143}$$

Hence

$$\begin{aligned}\vec{\nabla}\vec{A}(\vec{r},\, t) &= \frac{\mu_0}{4\pi}\int_V dV'\left[-\vec{\nabla}'\frac{\vec{J}(\vec{r}',\, t_r)}{R} - \frac{\dot{\rho}}{R}\right] \\ &= -\frac{\mu_0}{4\pi}\int_V dV'\frac{\dot{\rho}}{R} \\ &= -\mu_0\epsilon_0\frac{\partial V}{\partial t}.\end{aligned} \tag{C.144}$$