PHYSICS: A BRIEF SUMMARY

MIGUEL A. LERMA

1. Introduction

This is a brief introduction to Physics intended for "the impatient". Its purpose is to give a brief summary of a number of core theories in Physics. Usually it takes several years for a Physics student to learn these theories, but for some practical purposes all you need to know can be told in the time it takes to read a booklet like this one.

This work is conceived as a dynamic document, that will be posted on the web and modified periodically to expand some sections, correct possible mistakes, and include further subjects of interest. Look for it at

`http://www.math.northwestern.edu/`
`~mlerma/courses/e11-99s/physics.pdf`

Please, send me your suggestions—email address at the end.

2. Mechanics

2.1. **Newton's Laws.** Ordinary Mechanics is ruled by Newton's laws. The motion of a particle is described by

$$\mathbf{F} = m\mathbf{a}, \tag{2.1}$$

where $\mathbf{F}$ is the applied force, m is the mass of the particle, and $\mathbf{a} = d\mathbf{v}/dt = d^2\mathbf{r}/dt^2$ is the particle's acceleration, with $\mathbf{v}$ being its velocity and $\mathbf{r}$ is position vector.

In coordinates equation (2.1) looks like this:

$$F_i = m\frac{d^2x_i}{dt^2} \qquad (i = 1, 2, 3). \tag{2.2}$$

2.2. **Euler-Lagrange equations.** Newton's law as described above is easy to use in Cartesian coordinates for mechanical problems without constrains, but it can be generalized in a way that makes it easier to apply to more general situations.

In one dimension Newton's law is

$$m\ddot{x} - F(x, t) = 0, \tag{2.3}$$

Date: February 24, 2021.

where the dot denotes time derivative. If the force derives from a potential $V(x,t)$, then $F(x,t) = -\partial V(x,t)/\partial x$. On the other hand, by using the kinetic energy $T(\dot{x}) = m\dot{x}^2/2$, and the momentum $p = m\dot{x} = \partial T/\partial \dot{x}$ we see that $m\ddot{x} = dp/dt$, hence

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{x}} + \frac{\partial V}{\partial x} = 0. \tag{2.4}$$

Now we introduce the *Lagrangian* function, $L(x,\dot{x}) = T(\dot{x}) - V(x)$, and the equation becomes:

$$\frac{d}{dt}\frac{\partial L(x,\dot{x})}{\partial \dot{x}} - \frac{\partial L(x,\dot{x})}{\partial x} = 0. \tag{2.5}$$

Its generalization to any number of (non necessarily Cartesian) coordinates $q_1, q_2, \ldots, q_n$ is the *Euler-Lagrange* equation:

$$\frac{d}{dt}\frac{\partial L(q_k,\dot{q}_k,t)}{\partial \dot{q}_k} - \frac{\partial L(q_k,\dot{q}_k,t)}{\partial q_k} = 0 \qquad (k = 1,2,\ldots,n). \tag{2.6}$$

It turns out that not only mechanical systems but also many other physical systems can be described by an equation like (2.6) with a suitable Lagrangian L. The choice of Lagrangian is dictated by physical experience, although some authors (such as Landau) have tried to derive it from general principles.

2.3. **Hamilton's Principle.** The *action* of a physical system with a given Lagrangian $L(q_k,\dot{q}_k,t)$ between times t_1 and t_2 is defined by the integral

$$S(q_k(t)) = \int_{t_0}^{t_1} L(q_k(t),\dot{q}_k(t),t)\,dt. \tag{2.7}$$

That integral depends on the path $\mathbf{q}(t)$ followed by the system between t_0 and t_1. Equation (2.6) turns out to be equivalent to the fact that the action (2.7) is a critical point (usually a minimum) in the space of paths with fixed endpoints $q_k(t_0)$ and $q_k(t_1)$:

$$\delta S = \delta \int_{t_0}^{t_1} L(q_k,\dot{q}_k,t)\,dt = 0. \tag{2.8}$$

The symbol δS (variation of S) represents the first order approximation of the change of S after a small perturbation $\delta q_k(t)$ of the path $q_k(t)$.

2.4. **Calculus of Variations.** Here we derive the Euler-Lagrange equations (2.6) from Hamilton's principle (2.8).

Consider a variation produced by perturbing the correct path $q_k(t)$ by $\alpha\eta_k(t)$, where $\eta_k(t)$ is an arbitrary differentiable function that vanishes at t_0 and t_1, and α is a real parameter:

$$S(q_k(t) + \alpha\eta_k(t)) = \int_{t_0}^{t_1} L(q_k + \alpha\eta_k, \dot{q}_k + \alpha\dot{\eta}_k, t)\, dt. \tag{2.9}$$

Differentiating respect to α at $\alpha = 0$, and taking into account that $\alpha = 0$ is a critical point we get:

$$\int_{t_0}^{t_1} \sum_k \left(\eta_k \frac{\partial L}{\partial q_k} + \dot{\eta}_k \frac{\partial L}{\partial \dot{q}_k} \right) dt = 0. \tag{2.10}$$

Integrating by parts and using that $\eta(t_0) = \eta(t_1) = 0$:

$$\int_{t_0}^{t_1} \sum_k \eta_k \left(\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} \right) dt = 0. \tag{2.11}$$

Since $\eta_k(t)$ is arbitrary we obtain the Euler-Lagrange equations:

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = 0. \tag{2.12}$$

These considerations are very general and can be applied not only to mechanical systems, but also to the dynamics of fields—see section (9.2).

2.5. **Hamilton's equations.** Hamilton's equations are similar to the Euler-Lagrange equations, but instead of using the generalized coordinates q_k and its derivatives $\dot{q}_k$ we use q_k and the generalized momenta $p_k = \partial L/\partial \dot{q}_k$. Note that for a particle of mass m, $p = m\dot{x}$ is the ordinary momentum.

If we define the *Hamiltonian* function

$$H(q_k, p_k, t) = \sum_j \dot{q}_j p_j - L(q_k, \dot{q}_k, t), \tag{2.13}$$

then the Lagrange-Euler equations can be replaced by Hamilton's equations:

$$\dot{q}_k = \frac{\partial H}{\partial p_k} \tag{2.14}$$

$$\dot{p}_k = -\frac{\partial H}{\partial q_k} \tag{2.15}$$

$$\frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t} \tag{2.16}$$

which have a symmetric look and consists of first order equations only.

For a single particle in a potential V, we have that $H = T + V$, i.e., H is the total energy of the particle. In general, for a conservative system[1] H represents the total energy of the system.

2.6. **Poisson brackets.** The Poisson bracket of two functions $A(p_k, q_k, t)$ and $B(p_k, q_k, t)$ is defined as

$$\{A, B\} = \sum_k \left(\frac{\partial A}{\partial q_k} \frac{\partial B}{\partial p_k} - \frac{\partial A}{\partial p_k} \frac{\partial B}{\partial q_k} \right) \tag{2.17}$$

The time evolution of a function A can be expressed with Poisson brackets in the following way:

$$\frac{dA}{dt} = \{A, H\} + \frac{\partial A}{\partial t}, \tag{2.18}$$

where H is the Hamiltonian.

Two variables r and s satisfying $\{r, s\} = 1$ are called *canonically conjugate*. In particular p_k and q_k are canonically conjugate.

3. Statistical Physics

3.1. **Thermodynamics.** Given a physical system, say a gas, its thermodynamic state is defined by a number of variables such as its temperature T, volume V, pressure p, energy U, etc. The evolution of the state obeys the following laws:

(1) **First Law: conservation of energy.** In any system:

$$dU = dQ + dW, \tag{3.1}$$

where U is the energy of the system, Q is the heat that flows into the system, and W is the work done on the system. For a gas with volume V and pressure p, $dW = -p\,dV$, so the first law for a gas is $dU = dQ - p\,dV$.

(2) **Second Law: increase of entropy.** In any closed system:

$$dS \equiv \frac{dQ}{T} \geq 0, \tag{3.2}$$

where S is the entropy of the system, Q is the heat produced in the system, and T is the absolute temperature of the system. A process is called *reversible* if $dS = 0$, otherwise ($dS > 0$) it is called *irreversible*. All real processes are irreversible.

(3) **Third Law: inaccessibility of the absolute zero.** The entropy of a system at $T = 0$ is constant: $dS = 0$.

[1] A system is called *conservative* if all its forces verify $\oint \mathbf{F}\, ds = 0$ along any closed path. This is equivalent to $\mathbf{F} = -\nabla V$ for some potential V.

3.2. **Equilibrium Statistical Mechanics.** For one mole of an ideal gas:

$$pV = RT, \tag{3.3}$$

where p is the pressure, V is the volume and $R = 8.3143\,\mathrm{J/mol\,K}$ is a constant. In (3.3) only "macroscopic" variables occur, but nothing is said about the physical state of the individual molecules that compose the gas. In order to connect the behavior of the gas to the statistical properties of its molecules we make the following assumptions:

(1) The thermodynamic properties of a macroscopic system are averages of the physical properties of the particles that compose the system.
(2) To each macroscopic state of the system it corresponds a number of possible microscopic states $s_1, s_2, s_3, \ldots, s_j, \ldots$, each with a probability P_j and an energy E_j.
(3) All configurations of the system with the same energy have the same probability of occurring.

So, for instance, the thermodynamic internal energy U of the system is the average energy $U = \sum_j P_j\, E_j$.

Now we find P_j for a volume of gas in contact with a large heat reservoir at a given temperature T. If the total energy is E and the energy of the gas in state s_j is E_j, the energy of the reservoir will be $E - E_j$. Let $\Omega(E - E_j)$ be the number of distinct states of the reservoir with energy $E - E_j$, so that $P_j \sim \Omega(E - E_j)$. If E_j is very small compared to E, we have: $\ln P_j = \ln \Omega(E - E_j) \approx \ln \Omega(E) - \beta E_j$, where $\beta = (\partial\Omega/\partial E)_E$ is the so called *Boltzmann factor*. From here we get the *Boltzmann equation*:

$$P_j = \frac{e^{-\beta E_j}}{Z}, \tag{3.4}$$

where

$$Z = \sum_j e^{-\beta E_j} \tag{3.5}$$

is the *partition function*.

The partition function allows us to find several thermodynamic variables, such as the internal energy:

$$U = \sum_j E_j P_j = -\frac{\partial \ln Z}{\partial \beta} \tag{3.6}$$

and the pressure

$$p = \sum_j p_j P_j = \frac{1}{\beta}\frac{\partial \ln Z}{\partial V}, \tag{3.7}$$

where $p_j = -(\partial E_j/\partial V)$ is the pressure associated with the j-th state.

We have that Z is a function of β and V, so:

$$d\ln Z = \frac{\partial \ln Z}{\partial \beta}\, d\beta + \frac{\partial \ln Z}{\partial V}\, dV = -d(\beta U) + \beta\, T\, dS, \tag{3.8}$$

hence:

$$T\, dS = \frac{1}{\beta}\, d(\ln Z + \beta U). \tag{3.9}$$

This implies $1/\beta = kT$ for some constant k—the *Boltzmann constant* $k = 1.3805 \times 10^{-23}$ J/K—, and

$$kT \ln Z = -F, \tag{3.10}$$

where $F = U - TS$ is the *Helmholtz free energy*. From $S = k \ln Z + U/T$ we also get the *entropy relationship*:

$$S = -k \sum_j P_j \, \ln P_j. \tag{3.11}$$

In systems with low energy fluctuations most states lay in a narrow band of Ω states around energy U, and then the entropy is given by

$$S = k \ln \Omega. \tag{3.12}$$

4. Electromagnetism

4.1. **Maxwell's equations.** The laws of electromagnetism can be summarized by Maxwell's equations:

$$\begin{array}{rcll} \nabla \times \mathbf{E} & = & -\dfrac{1}{c}\dfrac{\partial \mathbf{B}}{\partial t} & \text{(Farady's law)} \\ \nabla \times \mathbf{B} & = & \dfrac{4\pi}{c}\mathbf{J} + \dfrac{1}{c}\dfrac{\partial \mathbf{E}}{\partial t} & \text{(Ampèr's law)} \\ \nabla \cdot \mathbf{E} & = & 4\pi\rho & \text{(Coulomb's law)} \\ \nabla \cdot \mathbf{B} & = & 0 & \text{(no magnetic monopoles)} \end{array} \tag{4.1}$$

where $\mathbf{E}$ represents the electric field, $\mathbf{B}$ is the magnetic field, $\mathbf{J}$ is the current density, ρ is the charge density, and c is the speed of light. From the second and third equations we derive the *continuity equation* expressing the conservation of electric charge:

$$\frac{\partial \rho}{\partial t} + \nabla \mathbf{J} = 0\,. \tag{4.2}$$

4.2. **Potentials. Gauge Transformations.** The description of the electromagnetic field can be simplified by introducing the potentials V and $\mathbf{A}$.

If the magnetic field is static, since $\nabla \times \mathbf{E} = 0$, there is some potential V such that $\mathbf{E} = \nabla V$. Similarly, because $\nabla \cdot \mathbf{B} = 0$, we can write $\mathbf{B} = \nabla \times \mathbf{A}$ for some vector field $\mathbf{A}$. More generally, we can define V and $\mathbf{A}$ with the following differential equations:

$$\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A} \\ \mathbf{E} &= -\left(\nabla V + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\right) \end{aligned} \tag{4.3}$$

This automatically ensures that two of the four Maxwell's equations are satisfied. Next we may rewrite the other two equations in terms of V and $\mathbf{A}$. Note, however, that V and $\mathbf{A}$ are not uniquely defined, since they can be changed to

$$\begin{aligned} \mathbf{A} &\to \mathbf{A} + \nabla M \\ V &\to V - \frac{1}{c}\frac{\partial M}{\partial t} \end{aligned} \tag{4.4}$$

for some scalar field M without altering $\mathbf{E}$ and $\mathbf{B}$. Picking an M is called *setting the gauge*.

Two gauges commonly used are the *Lorenz gauge*, for which $\nabla \cdot \mathbf{A} + \partial V/\partial t = 0$, and the *Coulomb gauge*, for which $\nabla \cdot \mathbf{A} = 0$.

In the Lorenz gauge, Maxwell's equations look like this:

$$\begin{aligned} \nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} &= -\frac{4\pi}{c}\mathbf{J} \\ \nabla^2 V - \frac{1}{c^2}\frac{\partial^2 V}{\partial t^2} &= -4\pi\rho \end{aligned} \tag{4.5}$$

4.3. **Electromagnetic waves.** In regions where there are no charges and no current, equations (4.5) become a pair of wave equations with a propagation velocity equal to c.

In terms of $\mathbf{E}$ and $\mathbf{B}$, the wave equations are the following:

$$\nabla^2 \mathbf{E} = \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} \quad \text{and} \quad \nabla^2 \mathbf{B} = \frac{1}{c^2}\frac{\partial^2 \mathbf{B}}{\partial t^2}. \tag{4.6}$$

The plane-wave solutions to these equations are

$$\mathbf{B} = \mathbf{B}_0 e^{i\mathbf{k}\cdot\mathbf{r} - i\omega t} \quad \text{and} \quad \mathbf{E} = \mathbf{E}_0 e^{i\mathbf{k}\cdot\mathbf{r} - i\omega t}, \tag{4.7}$$

where $\mathbf{k}$ is the wave vector that points in the direction of wave propagation, ω is the wave frequency (units of radians/time), $\mathbf{r}$ is the position vector, and $\mathbf{B}_0$ and $\mathbf{E}_0$ are constants. The wave length is $2\pi/k$.

Note that the solutions shown are complex functions. Of course, only the real part of those functions have a physical meaning, but a complex representation makes mathematics easier.

4.4. **Force on moving charges.** The force experienced by a distribution of charges with density ρ and velocity $\mathbf{v}$ in an electromagnetic field is called *Lorentz force* and is given by the following equation:

$$\mathbf{f} = \rho \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B} \right). \tag{4.8}$$

The actual force F on a volume of charge is obtained by integrating $\mathbf{f}$ on that volume. For a single particle with charge q the force is

$$\mathbf{F} = q \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B} \right). \tag{4.9}$$

The dynamics of a charged particle in an electromagnetic field can be described by the Euler-Lagrange equations with a potential

$$U = q \left(V - \frac{1}{c} \mathbf{A} \cdot \mathbf{v} \right). \tag{4.10}$$

5. Special Relativity

5.1. **Change of frame of reference.** An inertial frame of reference is a Cartesian system of coordinates that is either fixed or moving linearly at a constant speed relative to the fixed stars. Events are identified by their space coordinates (x, y, z), plus their time coordinate t.

Before Einstein, the relation between two frames F and F' such that F' moves with speed v in the direction of the X axis was the *Galilean transformation*:

$$\begin{cases} x' = x - vt \\ t' = t \end{cases} \tag{5.1}$$

where (x, t) are the position and time of a given event in frame F and (x', t') are the position and time of the same event in frame F'. Note that if a body moves in the direction of the X axis at a speed u respect to F, then its speed respect to F' would be $u - v$.

The Galilean transformation is incompatible with Maxwell's equations, since they predict that electromagnetic waves move at speed c regardless of the frame of reference. Various experiments (such as Michelson-Morley's) also supported the principle of invariance of the speed of light, so the Galilean transformation must be replaced with a

different one that leaves c invariant. That is the *Lorentz transformation*:

$$\begin{cases} x' = \dfrac{x - vt}{\sqrt{1 - v^2/c^2}} \\ t' = \dfrac{t - vx/c^2}{\sqrt{1 - v^2/c^2}} \end{cases} \tag{5.2}$$

The Lorentz transformation also leaves the following 2-form invariant:

$$ds^2 = c^2\, dt^2 - dx^2 - dy^2 - dz^2. \tag{5.3}$$

The integral of ds along a space-time path can be considered as the "length" of the path. So, by using coordinates $x_0 = ct$, $x_1 = x$, $x_2 = y$, $x_3 = z$, we get that ds is the length element in a flat 4-dimensional pseudo-Riemannian manifold with metric tensor $\eta_{\mu\nu}$ such that $\eta_{00} = 1$, $\eta_{ii} = -1$ for $i = 1, 2, 3$, and $\eta_{\mu\nu} = 0$ for $\mu \neq \nu$.[2]

The time τ as measured by a clock accompanying a moving body is called *proper time* of the body, and it verifies $c\, d\tau = ds$ along the path followed by the body.

5.2. **Tensor notation.** In general vectors are represented by their coordinates A^μ, with upper indices $\mu = 0, 1, 2, 3$. These are called *countervariant* coordinates of the vector. The scalar product by a given vector A^μ defines a linear form of coordinates $A_\mu = \sum_{\nu=0}^{3} \eta_{\mu\nu}\, A^\nu$, that we write with lower indices. These are also called *covariant* coordinates of the vector. Whenever an index appears repeated in two places of an expression we assume that we must sum respect to all values of that index (Einstein's convention), so for instance we can write $A_\mu = \eta_{\mu\nu}\, A^\nu$ without explicitly writing the summation sign. The length squared of a 4-vector A^μ is $A_\mu\, A^\mu$. The coordinates are represented as 4-vectors, hence they must use upper indices: x^μ. The expression defining ds can be written $ds^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu$.

We can raise or lower indices in the coordinates of a general tensor by contraction with $\eta_{\mu\nu}$ or with its countervariant version $\eta^{\mu\nu}$, which turns out to be equal to $\eta_{\mu\nu}$—in a more general system of coordinates with $ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu$, $g^{\mu\nu}$ is defined so that $g_{\lambda\mu}\, g^{\mu\nu} = \delta^\lambda{}_\nu$, where $\delta^\lambda{}_\nu$ is Kronecker's delta (see below). So, for instance, given a tensor $A_{\mu\nu}$, we represent $A^\mu{}_\nu = \eta^{\mu\xi}\, A_{\xi\nu}$, $A_\mu{}^\nu = \eta^{\nu\xi}\, A_{\mu\xi}$, $A^{\mu\nu} = \eta^{\mu\xi}\, \eta^{\nu\rho}\, A_{\xi\rho}$, $A_{\mu\nu} = \eta_{\mu\xi}\, \eta_{\nu\rho}\, A^{\xi\rho}$, etc.

[2] I am following the *timelike* $(+, -, -, -)$ convention. Some authors use the *spacelike* $(-, +, +, +)$ convention.

If $A_{\mu\nu}$ is a tensor field, its partial derivative respect to x^α is sometimes represented $A_{\mu\nu,\alpha} = \partial_\alpha A_{\mu\nu}$, where $\partial_\alpha = \frac{\partial}{\partial x^\alpha}$. Also $\partial^\mu = \eta^{\mu\nu}\,\partial_\nu$.

Several useful symbols are the following:

(1) Kronecker's delta $\delta^\alpha{}_\beta = 1$ if $\alpha = \beta$, and 0 otherwise.
(2) The completely antisymmetric symbol $\epsilon_{\alpha\beta\gamma\delta}$, such that $\epsilon_{0123} = 1$, and is antisymmetric in its indices.

Given a tensor $A_{\alpha\ldots\omega}$, by enclosing n of its indices between brackets we denote the result of *antisymmetrizing* respect to those indices. i.e.:

$$A_{\alpha\ldots[\mu\ldots\theta]\ldots\omega} = \frac{1}{n!}\sum \sigma(\mu',\ldots,\theta')\, A_{\alpha\ldots[\mu'\ldots\theta']\ldots\omega}, \tag{5.4}$$

where the sum is carried out through all permutations $(\mu',\ldots,\theta')$ of $(\mu,\ldots,\theta)$, and $\sigma(\mu',\ldots,\theta')$ is the signature of the permutation.

5.3. **Covariance.** A formula representing some law of Nature is said to be *covariant* if it has the same form in any frame of reference. Maxwell's equations are not covariant under Galileo's transformation, but they are under Lorentz's transformation. However Newton's second law, which is covariant under Galileo's transformation, is not under Lorentz's transformation, so the dynamics of a particle needs to be reformulated.

5.4. **Dynamics of a particle.** The relativistic version of Newton's second law can be expressed with 4-vectors in the following way:[3]

$$\frac{d\pi^\mu}{d\tau} = \phi^\mu \tag{5.5}$$

where π^μ is the 4-momentum, and ϕ^μ is the *Minkowski* force. The components of π^μ are $\pi^0 = mc\gamma$, and $\pi^i = m\dot{x}^i\gamma$ for $i = 1,2,3$, $\gamma = 1/\sqrt{1 - v^2/c^2}$, $v =$ speed of the particle, $m =$ rest mass of the particle. So the space components of π^μ are the momentum of a particle with mass $m' = m/\sqrt{1 - v^2/c^2}$. The time component of the relativistic momentum is $\pi^0 = T/c$, where $T =$ relativistic kinetic energy. For small speed v we have:

$$T = \frac{mc^2}{\sqrt{1 - v^2/c^2}} \approx mc^2 + \frac{mv^2}{2}. \tag{5.6}$$

The second term is the non-relativistic kinetic energy of the body, and the first term $E = mc^2$ is its rest energy.

[3]This formula fits the dynamics of a charged particle in an electromagnetic field, and is assumed to apply to other types of force too.

The energy of a moving body can also be expressed with the equation:

$$E^2 = p^2c^2 + m^2c^4 \tag{5.7}$$

For a free particle equation (5.5) becomes $d\pi^\mu/d\tau = 0$, which can be derived from a variational principle with the following action:

$$S = -mc \int_{s_0}^{s_1} ds, \tag{5.8}$$

i.e., a particle follows a path of extremal length or *geodesic*.

5.5. **Electromagnetism.** The relativistic version of Maxwell's equations uses the 4-vectors $A^\mu$and j^μ, where $A^0 = V$, its space components are those of the usual vector potential; $j^0 = 4\pi\rho$, and $j^i = \frac{4\pi}{c}J^i$ for $i = 1, 2, 3$. Then Maxwell's equations with the Lorenz gauge $\partial_\mu A^\mu = 0$ become:

$$\partial^\mu \partial_\mu A^\alpha = j^\alpha. \tag{5.9}$$

The Minkowski force on a moving charge is given by $\phi_\mu = \dfrac{q}{c} F_{\mu\nu}\, u^\nu$, where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{5.10}$$

is the *field-strength tensor*, and $u^\mu = dx^\mu/d\tau$.

In terms of $F_{\mu\nu}$, Maxwell's equations are:

$$\begin{aligned} \partial_\nu F^{\mu\nu} &= j^\mu \\ \partial_\theta F_{\mu\nu} + \partial_\mu F_{\nu\theta} + \partial_\nu F_{\theta\mu} &= 0. \end{aligned} \tag{5.11}$$

Equivalently, they can be expressed as ${}^*F_{[\lambda\mu,\nu]} = j^\mu$ and $F_{[\lambda\mu,\nu]} = 0$ respectively, where ${}^*F_{\mu\nu} = \frac{1}{2}F^{\alpha\beta}\,\epsilon_{\alpha\beta\mu\nu}$ is the dual tensor of $F_{\mu\nu}$ and the brackets denote antisymmetrization.

These equations can be derived from a variational principle with the following action:

$$S = -\frac{1}{4}\int F_{\mu\nu}F^{\mu\nu}\, d^3\mathbf{x}\, dt. \tag{5.12}$$

Here the components of the tensor $F_{\mu\nu}$ play the role of generalized coordinates.

On the other hand, the action of a charged particle in an electromagnetic field is

$$S = \frac{q}{c}\int A_\mu\, dx^\mu. \tag{5.13}$$

6. General Relativity

6.1. **Newtonian Gravitation.** Newton's gravitational law states that the force between two masses m_1 and m_2 at a distance r is

$$F = -k\,\frac{m_1 m_2}{r^2}, \tag{6.1}$$

where $k = 6.670 \times 10^{-11}\ \mathrm{N\,m^2\,kg^{-2}}$ is Newton's gravitational constant. The corresponding potential produced by a mass m is $\phi = km/r$, and for a continuous fluid of density ρ it is given by the *Poisson's equation*:

$$\nabla^2\phi = 4\pi k\,\rho. \tag{6.2}$$

Since Poisson's equation is not Lorentz invariant, it is not appropriate for a relativistic theory of gravity.

There are various ways of developing a covariant theory of gravitation in the frame of Special Relativity, but they are inconsistent or lead to predictions that do not match experimental observations. So a completely new approach is needed.

6.2. **The Principle of Equivalence.** In General Relativity it is postulated that a gravitational field is locally indistinguishable from an accelerated frame. In Relativity non inertial frames are represented by non Cartesian coordinates in 4-space. So gravitation is not a "force", but a change in the geometry of space-time. Particles that are not under the action of some other field behave like free particles, and their paths are still geodesics in space-time.

6.3. **Geometry in a curved space-time.** Space-time is represented by a pseudo-Riemannian 4-dimensional manifold. Its geometry is given by the (symmetric) *metric tensor* $g_{\mu\nu}$.[4] The line element is

$$ds^2 = g_{\mu\nu}\,dx^\mu\,dx^\nu. \tag{6.3}$$

The usual derivatives of the components of a tensor respect to the coordinates are not covariant in a general frame, particularly in a curved space-time. This is due to the fact that the basic vectors $\mathbf{e}_\mu = \frac{\partial}{\partial x^\mu}$ are not constant. So, in particular

$$\partial_\nu(A^\mu\,\mathbf{e}_\mu) = (\partial_\nu A^\mu)\,\mathbf{e}_\mu + A^\mu\,\partial_\nu\mathbf{e}_\mu = (\partial_\nu A^\mu + \Gamma^\mu{}_{\alpha\nu}\,A^\alpha)\,\mathbf{e}_\mu\,, \tag{6.4}$$

where $\Gamma^\mu{}_{\alpha\nu}$ are the components of $\partial_\nu\mathbf{e}_\alpha$.[5] The components of $\partial_\nu\mathbf{A}$ are denoted $A^\mu{}_{;\nu} = \partial_\nu A_\mu + \Gamma^\mu{}_{\alpha\nu}A^\alpha$. There are similar formulas for covariant coordinates and for tensors of higher rank. Since $A^\mu{}_{;\nu}$ are

[4] if the space-time is flat then $g_{\mu\nu} = \eta_{\mu\nu}$ in any inertial frame.

[5] It can be proved that $g_{\alpha\beta}\,\Gamma^\beta{}_{\mu\nu} = \Gamma_{\alpha\mu\nu} = \frac{1}{2}(\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\alpha\mu} - \partial_\alpha g_{\mu\nu})$. The symbols $\Gamma^\alpha{}_{\mu\nu}$ and $\Gamma_{\alpha\mu\nu}$ are called *Christoffel symbols*.

the components of a tensor, its covariant derivative respect to x^γ also makes sense, and is denoted $A^\mu{}_{;\nu\gamma}$.

Note that covariant derivatives do not commute in general, so that for instance

$$A_{\alpha;\beta\gamma} - A_{\alpha;\gamma\beta} = R^\mu{}_{\alpha\beta\gamma}\, A_\mu\,, \tag{6.5}$$

where

$$R^\mu{}_{\alpha\beta\gamma} = \partial_\beta \Gamma^\mu{}_{\alpha\gamma} - \partial_\gamma \Gamma^\mu{}_{\alpha\beta} + \Gamma^\nu{}_{\alpha\gamma}\, \Gamma^\mu{}_{\nu\beta} - \Gamma^\nu{}_{\alpha\beta}\, \Gamma^\mu{}_{\nu\gamma} \tag{6.6}$$

is the *Riemann tensor*.[6] Other related tensors are the *Ricci curvature tensor* $R_{\alpha\beta} = R^\mu{}_{\alpha\beta\mu}$, the *scalar curvature* $R = R^\mu{}_\mu$ and the *Einstein curvature tensor* $G^\alpha{}_\beta = R^\alpha{}_\beta - \frac{1}{2}\delta^\alpha{}_\beta\, R$. Two important properties are the *Bianchi* identities:

$$R^\alpha{}_{\beta\,[\lambda\mu;\nu]} = 0, \tag{6.7}$$

where the brackets denote antisymmetrization, and the contracted Bianchi identities:

$$G^{\mu\nu}{}_{;\nu} = 0. \tag{6.8}$$

6.4. **Dynamics in a curved space-time.** Free particles follow geodesics, given by the equation:

$$\ddot{x}^\alpha + \Gamma^\alpha{}_{\mu\nu}\, \dot{x}^\mu\, \dot{x}^\nu = 0, \tag{6.9}$$

where the dot denotes derivative respect to s. This equation can also be expressed as $\dot{x}^\alpha{}_{;\nu}\, \dot{x}^\nu = 0$.

When matter is distributed continuously in space-time (like a fluid), its dynamical properties can be represented by the (symmetric) *energy-momentum tensor* $T_{\mu\nu}$, defined so that $T^\mu{}_\nu\, u^\nu$ gives the density of 4-momentum that flows in the direction of the unit 4-vector u^μ. In particular, for a "dust" of non interacting particles: $T^{\mu\nu} = \rho\, c^2\, \dot{x}^\mu\, \dot{x}^\nu$, where ρ is the proper density of the dust (as measured in a frame respect to which the fluid is locally at rest). For the electromagnetic field: $T^{\mu\nu} = F^{\mu\alpha}\, F^\nu{}_\alpha - \frac{1}{4}\, g^{\mu\nu}\, F_{\alpha\beta}\, F^{\alpha\beta}$, where $F_{\mu\nu} = A_{\nu;\mu} - A_{\mu;\nu}$. In terms of the energy-momentum tensor, the conservation of 4-momentum is given by the equation:

$$T^{\mu\nu}{}_{;\nu} = 0. \tag{6.10}$$

[6] In a flat space-time the Riemann tensor is always zero.

6.5. **Einstein's gravitational equations.** The gravitational equations in a curved space-time must generalize Poisson's equation (6.2). In a weak gravitational field we can write $ds^2 \approx (1+\varphi(\mathbf{x}))\, c^2 dt^2 - d\mathbf{x}^2$, where φ is small. In that field the equations of geodesics (6.9) become approximately $\ddot{\mathbf{x}} \approx -\frac{1}{2}\nabla\varphi$, hence $g_{00} \approx 1 + 2\phi/c^2$, where ϕ is the gravitational potential. Also, for low velocities: $T_{00} = \rho c^2$. So Poisson's equation becomes approximately: $\nabla^2 g_{00} = \frac{8\pi k}{c^4}\, T_{00}$. This suggests that the gravitational equations must be of the form $G_{\mu\nu} = \kappa\, T_{\mu\nu}$, where $G_{\mu\nu}$ is some symmetric 2-rank tensor such that

(1) it is constructed solely from $g_{\mu\nu}$, $g_{\mu\nu;\alpha}$ and $g_{\mu\nu;\alpha\beta}$, and is linear in $g_{\mu\nu;\alpha\beta}$,
(2) it has a vanishing divergence $G^{\mu\nu}{}_{;\nu} = 0$, and
(3) it vanishes when space-time is flat.

It turns out that the only tensor fulfilling these conditions is, up to a multiplicative constant, Einstein's curvature tensor $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}\, g_{\mu\nu}\, R$. After adjusting the constant we get Einstein's gravitational equations:

$$G_{\mu\nu} = \kappa\, T_{\mu\nu}, \tag{6.11}$$

where $\kappa = 8\pi k/c^4$.

Those equations also can be obtained from Hilbert's variational principle, with the following action:

$$S = \frac{c^3}{16\pi k} \int R\, \sqrt{-g}\, d\Omega, \tag{6.12}$$

where R is the scalar curvature, g is the determinant of $g_{\mu\nu}$, and $d\Omega = dx^0\, dx^1\, dx^2\, dx^3$ is the 4-volume element.

7. Quantum Mechanics

7.1. **Introduction.** In spite of its success, classical Mechanics (Newtonian Mechanics extended with Relativity) cannot explain a number of phenomena, such as

(1) Stability of electrons in discrete orbits in an atom.
(2) Spectrum of the black body radiation.
(3) Photoelectric effect.
(4) Diffraction of electrons.
(5) Quantization of angular momentum (Stern-Gerlach).

Everything seemed to suggest that

(1) Energy comes in discrete units (*quanta*).

(2) Light is made up of particles (photons), of energy $E = h\nu$ and momentum $p = h/\lambda$, where ν = frequency, λ = wave length, $h = 6.6256 \times 10^{-34}$ J s, Planck's constant.
(3) Particles have wavelike properties.
(4) Some details of the state of small physical systems cannot be determined exactly and we must resort to a probabilistic description.

7.2. **The wave function.** The state of a single particle can be described with its *wave function*, which is a complex function $\psi(\mathbf{x}, t) \in L^2(\mathbb{R}^3)$, such that $|\psi(\mathbf{x}, t)|^2$ represents the probability density for the particle being at point $\mathbf{x}$. That function is assumed normalized by $\int \psi(\mathbf{x}, t)\, d^3\mathbf{x} = 1$, where the integral extends to the whole space. In Dirac's notation the state of the particle is represented with a *ket* $|\psi\rangle$.

The set of possible wave functions of the particle form a Hilbert space $\mathcal{H}$ with inner product

$$\langle \xi, \psi \rangle = \int \overline{\xi(\mathbf{x}, t)}\, \psi(\mathbf{x}, t)\, d\mathbf{x}, \tag{7.1}$$

where the bar denotes complex conjugation.

The symbol $\langle \xi|$, also called a *bra*, represents an element of the dual Hilbert space consisting of the map $|\psi\rangle \mapsto \langle \xi|\psi\rangle = \langle \xi, \psi\rangle$.

In general, the set of possible states of a physical system is represented by a complex Hilbert space $\mathcal{H}$, usually assumed to be separable.

7.3. **Composite systems.** For composite systems consisting of two or more subsystems with Hilbert spaces $\mathcal{H}^{(1)}$, $\mathcal{H}^{(2)}$, ..., $\mathcal{H}^{(N)}$ respectively, the space of states is the tensor product $\mathcal{H} = \mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)} \otimes \cdots \otimes \mathcal{H}^{(N)}$. A typical example is a system of two particles with wave functions $\psi^{(1)}(\mathbf{x}, t)$ and $\psi^{(2)}(\mathbf{x}, t)$ respectively. The state of the join system is represented by a wave function of the form $\psi(\mathbf{x}_1, \mathbf{x}_2, t) = \psi^{(1)}(\mathbf{x}_1, t)\, \psi^{(2)}(\mathbf{x}_2, t)$.

7.4. **Observables.** Observable quantities such as position, momentum, energy, etc., are represented by *selfadjoint linear operators* that act on the Hilbert space.

A linear operator A is said to be *Hermitian* or *selfadjoint* if $A^\dagger = A$, where $A^\dagger$ is the adjoint operator of A, defined by the relation $\langle \xi, A\psi\rangle = \langle A^\dagger \xi, \psi\rangle$. As a matter of notation: $\langle \psi|A = \langle A\psi|$, and if A is selfadjoint, $\langle \xi|A|\psi\rangle = \langle \xi, A\psi\rangle = \langle A\xi, \psi\rangle$.

For a particle the position is represented by multiplication by $\mathbf{x}$, time is multiplication by t, linear momentum is $-i\hbar\nabla$, where $\hbar = h/2\pi$, energy is $i\hbar\frac{\partial}{\partial t}$.

The possible values of an observable are represented by its eigenvalues, and its eigenvectors represent states for which that observable has a well defined value. For instance, a plane wave $\psi(\mathbf{x},t) = e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$ is an eigenvector for both the momentum and energy operators:[7]

$$\begin{aligned} -i\hbar\nabla\psi(\mathbf{x},t) &= \hbar\mathbf{k}\,\psi(\mathbf{x},t) \\ i\hbar\frac{\partial}{\partial t}\,\psi(\mathbf{x},t) &= \hbar w\,\psi(\mathbf{x},t). \end{aligned} \tag{7.2}$$

So the momentum of a plane wave is $\hbar\mathbf{k}$, and its energy is $\hbar\omega$.

If a given state $|\psi\rangle$ is not an eigenvector for the operator A associated to a given observable, we still can define the *average value* of that observable in that state in the following way:

$$\langle A\rangle_\psi = \langle\psi|A|\psi\rangle. \tag{7.3}$$

The *dispersion* of an observable A in a state $|\psi\rangle$ is

$$\Delta_\psi A = \langle (A - \langle A\rangle_\psi)^2\rangle_\psi^{1/2} = \left(\langle A^2\rangle_\psi - \langle A\rangle_\psi^2\right)^{1/2}. \tag{7.4}$$

7.5. **Uncertainty relations.** Given two observables A and B, their dispersions in any state $|\psi\rangle$ satisfy the following relation:

$$\Delta_\psi A\,\Delta_\psi B \geq \frac{1}{2}|\langle [A,B]\rangle_\psi|, \tag{7.5}$$

where $[A,B] = AB - BA$. Hence, if two observables do not commute, they cannot have well defined values in the same state. In particular, for the position x and momentum $p = -i\hbar\frac{\partial}{\partial x}$ of a particle in one dimension:

$$[x,p] = i\hbar \quad\Longrightarrow\quad \Delta_\psi x\,\Delta_\psi p \geq \hbar/2. \tag{7.6}$$

Similarly for the time t and energy $E = i\hbar\frac{\partial}{\partial t}$:

$$[E,t] = i\hbar \quad\Longrightarrow\quad \Delta_\psi E\,\Delta_\psi t \geq \hbar/2. \tag{7.7}$$

7.6. **Complete sets of commuting observables.** A complete set of commuting observables is a set of observables which all commute with one another and for which there is only one simultaneous eigenstate belonging to any set of eigenvalues.

The existence of a complete set of commuting observables allows to set up an orthogonal basis for the Hilbert space consisting of simultaneous eigenstates for those observables. The elements of the basis can be distinguished by their different eigenvalues.

[7] A plane wave is a special case, since it cannot be normalized in the usual sense.

7.7. **Measurements.** Assume that an observable A has eigenvectors $|\psi_n\rangle$ with different eigenvalues a_n ($n = 1, 2, 3, \dots$). A measurement of A on a physical system initially in a state $|\psi\rangle$ will yield the value a_n with probability $p_n = |\langle\psi_n|\psi\rangle|^2$. After the measurement the state of the system becomes $|\psi_n\rangle$.

Formally the effect of the measurement is an statistical mixture represented by the following *density operator*:

$$\rho = \sum_n p_n \, |\psi_n\rangle\langle\psi_n|, \tag{7.8}$$

where $\sum_n p_n = 1$. The entropy of the mixture is

$$s(\rho) = -\mathrm{Tr}(\rho \ln \rho) = -\sum_n p_n \ln p_n, \tag{7.9}$$

where $\mathrm{Tr(A)} = \sum_n \langle\psi_n|A|\psi_n\rangle$ ($\{|\psi_n\rangle\}$ form a Hilbert basis) represents the *trace* of an operator A. For a pure state $|\psi\rangle$ the entropy is always zero, but for a mixture it is positive in general, which indicates a loss of information.

7.8. **Schrödinger's equation.** The time evolution of a physical system is given by the time dependent Schrödinger's equation:

$$i\hbar \frac{\partial}{\partial t} \, |\psi\rangle = H \, |\psi\rangle, \tag{7.10}$$

where H is the Hamiltonian operator representing the energy of the system. The form of H depends on the system. Often in can be obtained from the classical Hamiltonian $H(q_k, p_K, t)$ by substituting the variables q_k and p_k by their associated quantum operators.

If H does not depend explicitly on time, the wave function can be written as $e^{-\frac{i}{\hbar}Et} \, |\psi\rangle$, where $|\psi\rangle$ verifies the time independent Schrödinger's equation:

$$H|\psi\rangle = E|\psi\rangle. \tag{7.11}$$

For a particle in a potential $V(\mathbf{x})$, the classical Hamiltonian is $H = p^2/2m + V(\mathbf{x})$. By using the quantum momentum $p = -i\hbar\nabla x$ instead we get $H = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})$. The time independent Schrödinger's equation in this case becomes:

$$-\frac{\hbar^2}{2m}\nabla^2|\psi\rangle + V|\psi\rangle = E|\psi\rangle. \tag{7.12}$$

7.9. **Pictures.** Schrödinger's equation (7.10) is valid in the so called *Schrödinger's picture*, in which states are assumed to evolve in time, and operators are time independent.

An alternate point of view is given by *Heisenberg's picture*, in which states are time independent, while the operators now change with time.

Let $U(t) = e^{-iHt/\hbar}$ be the unitary operator that maps $|\psi(0)\rangle$ to $|\psi(t)\rangle$ in Schrödinger's picture:

$$|\psi(t)\rangle = U(t)\,|\psi(0)\rangle. \tag{7.13}$$

In Heisenberg's picture the states remain constant and equal to $|\psi(0)\rangle$, while operators evolve in time. If A_S is an operator in Schrödinger's picture, then the corresponding operator in Heisenberg's picture is

$$A_H(t) = U(t)\,A_S(t)\,U^\dagger(t), \tag{7.14}$$

and it evolves in time according to the equation

$$\frac{dA_H}{dt} = \frac{i}{\hbar}[H, A_H]\,. \tag{7.15}$$

In both pictures the matrix elements $\langle\xi|A|\psi\rangle$ are the same, so they provide physically equivalent descriptions of the system.

7.10. **S-matrix.** In some interactions, such as a collision, the Hamiltonian can be written as $H = H_0 + H_I$, where is H_0 corresponds to a free particle, and H_I, the part that rules the interaction, vanishes asymptotically at large distances. For $t \to \pm\infty$ the state of the system $|\psi(t)\rangle$ approaches asymptotic states $|\psi_{\text{in}}(t)\rangle$ and $|\psi_{\text{out}}(t)\rangle$ respectively, which evolve according to the Schrödinger equation with Hamiltonian H_0. The S-matrix is the operator that relates the asymptotic states:

$$S\,|\psi_{\text{in}}(t)\rangle = |\psi_{\text{out}}(t)\rangle\,. \tag{7.16}$$

7.11. **Canonical quantization.** In general, given a classical system with Hamiltonian $H(q_k, p_k, t)$, in the corresponding quantum system we consider $q_k(t)$ and $p_k(t)$ as operators satisfying the commutation relations

$$[q_k, p_l] = i\hbar\delta_{kl}, \tag{7.17}$$

where δ_{kl} is Kronecker's delta. In Heisenberg's picture the time evolution of an operator $A(t)$ is given by *Heisenberg's equation*:

$$i\hbar\,\frac{dA(t)}{dt} = [A(t), H(t)] + i\hbar\,\frac{\partial A(t)}{\partial t}, \tag{7.18}$$

which is analogous to the classical one (2.18).

7.12. **Momentum representation.** Given a one-dimensional wave function $\psi(x)$, we can write it as

$$\psi(x) = \int \hat{\psi}(p)\, \frac{e^{ipx/\hbar}}{2\pi\hbar}\, dp, \tag{7.19}$$

where

$$\hat{\psi}(p) = \int \psi(x)\, \frac{e^{-ipx/\hbar}}{2\pi\hbar}\, dx \tag{7.20}$$

is its Fourier transform. The function $\hat{\psi}(p)$ is the *momentum* representation of the state; $\psi(x)$ is the *position* representation. In the momentum representation, the momentum operator is multiplication by p, and the position operator is $x = i\hbar\frac{\partial}{\partial p}$.

7.13. **Harmonic oscillator.** A harmonic oscillator is a particle in a potential $V(x) = kx^2/2$. With this potential, equation (7.12) becomes

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}|\psi\rangle + \frac{k}{2}x^2|\psi\rangle = E|\psi\rangle. \tag{7.21}$$

This differential equation may be solved by standard techniques, but it is more convenient to use an operational approach—the "ladder" method. Instead of x and $p = -i\hbar\frac{\partial}{\partial x}$ we use the operators:

$$P = \frac{1}{\sqrt{\hbar m\omega}}\, p \quad \text{and} \quad Q = \sqrt{\frac{m\omega}{\hbar}}\, x\,, \tag{7.22}$$

where $\omega = \sqrt{k/m}$= frequency of the oscillator. These operators verify $[P, Q] = i$. Next we define

$$\mathfrak{a} = \frac{1}{\sqrt{2}}\,(Q + iP)\ , \quad \mathfrak{a}^\dagger = \frac{1}{\sqrt{2}}\,(Q - iP)\ , \quad N = \mathfrak{a}^\dagger\mathfrak{a}\,, \tag{7.23}$$

which verify $[\mathfrak{a}, \mathfrak{a}^\dagger] = 1$, $[N, \mathfrak{a}] = -\mathfrak{a}$, $[N, \mathfrak{a}^\dagger] = \mathfrak{a}^\dagger$. The Hamiltonian becomes:

$$H = \tfrac{1}{2}\hbar\omega\left(P^2 + Q^2\right) = \tfrac{1}{2}\hbar\omega\left(\mathfrak{a}^\dagger\mathfrak{a} + \tfrac{1}{2}\right) = \tfrac{1}{2}\hbar\omega\left(N + \tfrac{1}{2}\right). \tag{7.24}$$

The eigenvalues of N are integers $n = 0, 1, 2, 3\ldots$, and the corresponding eigenvectors, which are also eigenvectors of H with eigenvalues $\frac{1}{2}\hbar\omega(n + \frac{1}{2})$, are represented $|n\rangle$.[8] The operators $\mathfrak{a}^\dagger$ and $\mathfrak{a}$ increase and decrease n respectively: $\mathfrak{a}^\dagger|n\rangle \sim |n+1\rangle$, $\mathfrak{a}|n\rangle \sim |n-1\rangle$, so that $|n\rangle \sim \mathfrak{a}^{\dagger n}|0\rangle$.

An harmonic oscillator can be interpreted as a system of n identical particles of energy $\hbar\omega$ each. The ground state $|0\rangle$ represents a system

[8]Note that the energy of the ground state ($n = 0$) is not zero—this is a consequence of the uncertainty principle.

with no particles, i.e.: vacuum. The operators $\mathfrak{a}$ and $\mathfrak{a}^\dagger$ are called *annihilation* and *creation* operators respectively, since they map a state with n particles to another state with $n-1$ or $n+1$ particles. N is the *number* operator, since it represents the number of particles in the system.

7.14. **Angular momentum.** Angular momentum is defined the same as in classical Mechanics, but using the quantum linear momentum in the definition: $\mathbf{L} = \mathbf{r} \times \mathbf{p} = -i\hbar\, \mathbf{r} \times \nabla$. In order to represent a more general type of angular momentum, including spin, we use letter $\mathbf{J}$ instead of $\mathbf{L}$. Its components J_x, J_y and J_z verify:

$$[J_x, J_y] = iJ_z, \quad [J_y, J_z] = iJ_x, \quad [J_z, J_x] = iJ_y. \tag{7.25}$$

However they commute with $\mathbf{J}^2$, i.e., $[J_k, \mathbf{J}^2] = 0$ for $k = x, y, z$. The eigenvalues and eigenvectors of $\mathbf{J}^2$ and one of its components, say J_z, can be found by an operational method similar to the one used in the harmonic oscillator. With the help of the ladder operators $J_+ = J_x + iJ_y$ and $J_- = J_x - iJ_y$, we get that the eigenvectors of $\mathbf{J}^2$ and J_z are of the form $|J, M\rangle$, where J and M are integers or half-integers representing the total angular momentum and its z-component respectively. The number J is non negative, and M can take the values $M = -J, -J+1, -J+2, \ldots, J-1, J$. The eigenvalues are:

$$\mathbf{J}^2|J, M\rangle = \hbar^2 J(J+1)|J, M\rangle, \quad J_z|J, M\rangle = \hbar M|J, M\rangle. \tag{7.26}$$

The operators $J_\pm$ increase and decrease M respectively: $J_\pm|J, M\rangle \sim |J, M \pm 1\rangle$.

A particular type of angular momentum is the intrinsic angular momentum or *spin* $\mathbf{S}$. Besides the trivial case of a particle of spin 0, the most simplest situation is that of spin $1/2$. In this case the Hilbert space is 2-dimensional, and in the base $\{|\frac{1}{2}, \frac{1}{2}\rangle, |\frac{1}{2}, -\frac{1}{2}\rangle\}$, the operators can be represented as $S_k = \frac{\hbar}{2}\sigma_k$, where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{7.27}$$

are Pauli's matrices.

For a particle with spin the total angular momentum is $\mathbf{J} = \mathbf{L} + \mathbf{S}$, and the Hilbert space of angular momentum is the tensor product $\mathcal{H}^{(J)} = \mathcal{H}^{(L)} \otimes \mathcal{H}^{(S)}$. The eigenvectors of $\mathbf{J}^2$ and J_z are of the form $|L, S, M_L, S_L\rangle = |L, M_L\rangle|S, M_S\rangle$, where $|L, M_L\rangle$ is an eigenvector of $\mathbf{L}^2$ and L_z, and $|S, M_S\rangle$ is an eigenvector of $\mathbf{S}^2$ and S_z.

7.15. **Translations.** If q represents position and $p = -i\hbar\frac{\partial}{\partial q}$ is the momentum operator, a translation $\psi(q) \mapsto \psi(q+a)$ of the wave function can be represented with the operator $T(a) = e^{aip/\hbar}$.[9]

Similarly, from the time dependent Schrödinger equation we get that time translation is given by the operator $U(t) = e^{-itH/\hbar}$, where H is the Hamiltonian. This operator can be applied to a state vector in order to obtain its time evolution: $|\psi(t)\rangle = U(t)|\psi(0)\rangle$.

A similar relation holds for azimuthal angle α and z-component of the angular momentum J_z, so that a rotation about the z-axis is given by the operator $R(\alpha) = e^{-i\alpha J_z/\hbar}$. If the angular momentum consists of the spin $\mathbf{S} = \hbar\,\mathbf{s}$ of a particle, then the rotation operator is $R(\alpha) = e^{-i\alpha s_z}$.[10]

In general, if p and q are two canonically conjugate variables, translations in q are given by an operator of the form $T(a) = e^{iap/\hbar}$. In "infinitesimal form" this is equivalent to $T(dq) = \mathrm{Id} + i\,p\,dq/\hbar$, where Id is the identity. The operator p is said to be the *generator* of translations in q. If we keep the state vectors fixed, then we may assume that observables are transformed in the following way in a translation:

$$A(q+a) = T(a)^\dagger\, A(q)\, T(a). \tag{7.28}$$

In infinitesimal form it becomes:

$$dA(q) = [A(q), p]\,\frac{i}{\hbar}\,dq. \tag{7.29}$$

7.16. **Symmetry and conservation laws.** Consider a physical system that is symmetric respect to translations $q \mapsto q + a$. In particular that implies that its Hamiltonian is invariant respect to infinitesimal translations $q \mapsto q + dq$, hence

$$[H, p] = 0, \tag{7.30}$$

where p is the generator of translations. But this implies that in Heisenberg's picture $dp/dt = 0$, i.e.: the observable p is conserved.

In general each kind of symmetry leads to a conservation law (*Noether's Theorem*):

---

[9]By Taylor (assuming that $\psi(q)$ is analytic):

$$\psi(q+a) = \sum_{n=0}^{\infty} \frac{a^n}{n!}\frac{\partial^n}{\partial q^n}\psi(q) = \sum_{n=0}^{\infty} i^n \frac{a^n}{n!}\frac{p^n}{\hbar^n}\psi(q) = e^{aip/\hbar}\psi(q)$$

[10]Note that for an eigenstate of s_z with eigenvalue $1/2$, $R(2\pi) = -1$, hence a rotation of 2π about the z-axis changes the sign of the wave function.

the following symmetry:	implies conservation of:
space	linear momentum
time	energy
rotational	angular momentum

7.17. **Identical particles.** Two particles are identical if they cannot be distinguished by observables.

Let P be the operator that permutes the particles: $P\,|\psi_1\rangle|\psi_2\rangle = |\psi_2\rangle|\psi_1\rangle$. Its eigenvectors are *symmetrized* states

$$|\psi_S\rangle = \frac{1}{\sqrt{2}}\left(|\psi_1\rangle|\psi_2\rangle + |\psi_2\rangle|\psi_1\rangle\right),$$

with eigenvalue $+1$, and *antisymmetrized* states

$$|\psi_A\rangle = \frac{1}{\sqrt{2}}\left(|\psi_1\rangle|\psi_2\rangle - |\psi_2\rangle|\psi_1\rangle\right),$$

with eigenvalue -1. If the particles are identical then P must commute with any observable of the system, in particular with the Hamiltonian $[P, H] = 0$. This implies that its eigenvalues must be constant in time, so that a symmetric (resp. antisymmetric) state must remain always symmetric (resp. antisymmetric).

A combination of various theoretical considerations and experimental results shows that, in fact, every system of two or more identical particles can be only in either symmetric or antisymmetric states. Furthermore, the *Spin-Statistics Theorem* states that if the spin of the particles is integral then the states are symmetric, and if it is half-integral then the states are antisymmetric. Particles with integral spin are called *bosons*, and those with half-integral spin are called *fermions*. For instance, photons (spin 1) and pions (spin 0) are bosons, while proton, neutrons and electrons (spin 1/2) are fermions.

A consequence of the spin-statistics theorem is *Pauli's exclusion principle*: two fermions cannot be in the same state, because otherwise they would be in a state simultaneously symmetric and antisymmetric, which can only be zero.

7.18. **Feynman's formulation of QM.** In Classical Mechanics a particle going from a point $(\mathbf{x}_0, t_0)$ to another $(\mathbf{x}, t)$ follows a path $\mathbf{y}(t')$ that minimizes the action $S = \int_{t_0}^{t} L(\mathbf{y}, \dot{\mathbf{y}}, t')\, dt'$. In Quantum Mechanics other paths are also possible, but they contribute to the probability of the particle going from $(\mathbf{x}_0, t_0)$ to $(\mathbf{x}, t)$ with an amplitude proportional to $\exp\{\frac{i}{\hbar}S(\mathbf{y}(t'))\}$. The probability is the integral

$$G(\mathbf{x}_0, t_0; \mathbf{x}, t) = \int \exp\{\tfrac{i}{\hbar}S(\mathbf{y}(t'))\}\, \mathcal{D}(\mathbf{y}(t')) \tag{7.31}$$

for all possible paths $\mathbf{y}(t')$ connecting $(\mathbf{x}_0, t_0)$ and $(\mathbf{x}, t)$. The integral in (7.31) is defined as a limit for $N \to \infty$ of an integral of the form

$$\int\int\cdots\int \exp\{\tfrac{i}{\hbar}S(\mathbf{y}_N)\}\,\frac{d^3\mathbf{x}_1\,d^3\mathbf{x}_2\ldots d^3\mathbf{x}_{N-1}}{\theta}\,, \tag{7.32}$$

where $y_N(t')$ is a polygonal path through the points $(\mathbf{x}_k, t_k)$, $k = 0, 1, \ldots, N$, $(\mathbf{x}_N, t_N) = (\mathbf{x}, t)$, $t_{k+1} - t_k = (t_N - t_0)/N$, and θ is a normalizing factor chosen so that the result can be interpreted as a probability.

Schrödinger's equation and the other principles of Quantum Mechanics can be derived from this formulation. Also, for macroscopic systems the contribution to the integral of all paths tend to cancel out except for the one that corresponds to the minimum action—which yields the classical Hamilton's principle.

8. Relativistic Quantum Theory

(Note: in the following we use *natural units* in which $c = \hbar = 1$.)

8.1. **Klein-Gordon equation.** Schrödinger's equation (7.12) is not Lorentz invariant, so it is incompatible with (special) relativity. A first attempt to replace it with a relativistic equation consists of replacing the non-relativistic Hamiltonian $H = p^2/2m + V(\mathbf{x})$ used in its derivation with a relativistic one. For a spin-zero particle of mass m and charge e in an electromagnetic field $A^\mu = (\varphi, \mathbf{A})$, we have

$$H = e\varphi + \sqrt{(\mathbf{p} - e\mathbf{A})^2 + m^2}\,. \tag{8.1}$$

Eliminating the square root we get

$$(H - e\varphi)^2 - (\mathbf{p} - e\mathbf{A})^2 = m^2\,, \tag{8.2}$$

and replacing the Hamiltonian and momentum with the corresponding quantum operators $H = i\partial/\partial t$, $\mathbf{p} = -i\nabla$, we get the *Klein-Gordon equation*:

$$\left[\left(i\frac{\partial}{\partial t} - e\varphi\right)^2 - \left(\frac{1}{i}\nabla - e\mathbf{A}\right)^2\right]|\psi\rangle = m^2|\psi\rangle\,, \tag{8.3}$$

which can be written

$$[(\partial_\mu + ieA_\mu)(\partial^\mu + ieA^\mu) + m^2]\,|\psi\rangle = 0\,, \tag{8.4}$$

obviously Lorentz invariant.

For a null field the equation becomes

$$(\Box + m^2)|\psi\rangle = 0\,, \tag{8.5}$$

where $\Box = \dfrac{\partial^2}{\partial t^2} - \nabla^2$ is the D'Alembert operator.

A difficulty with this equation is that its solutions may have a probability density that is not defined positive. In the context of quantum field theory this difficulty is overcome by reinterpreting the probability density as a charge density instead.

8.2. **Dirac equation.** The Dirac equation is a generalization of the Klein-Gordon equation for particles of spin 1/2 (such as an electron).

We start by replacing the non-relativistic Hamiltonian with a relativistic one, as we did for the Klein-Gordon equation, but instead of eliminating the square root by squaring, we "calculate" the square root $\sqrt{H^2 - p^2} = \left(-\frac{\partial}{\partial t} + \nabla^2\right)^{1/2}$ by finding appropriate A, B, C, D verifying

$$\left(A\frac{\partial}{\partial x} + B\frac{\partial}{\partial y} + C\frac{\partial}{\partial z} + D\frac{\partial}{\partial t}\right)^2 = \nabla^2 - \frac{\partial}{\partial t}\,. \tag{8.6}$$

This can be accomplished using 4×4 matrices for A, B, C, D. At the end the Dirac equation (in a null field) becomes

$$i\frac{\partial}{\partial t}|\psi\rangle = (-i\boldsymbol{\alpha}\cdot\nabla + \beta m)\,|\psi\rangle\,. \tag{8.7}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$, and $\beta, \alpha_1, \alpha_2, \alpha_3$ are the Dirac matrices:

(8.8)

$$\beta = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix},\ \alpha_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -\sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix},$$

$$\alpha_2 = \begin{pmatrix} 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} = \begin{pmatrix} -\sigma_y & 0 \\ 0 & \sigma_y \end{pmatrix},\ \alpha_3 = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} -\sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix},$$

where $\sigma_x, \sigma_y, \sigma_z$ are Pauli's matrices.

For a particle of mass m and charge e in an electromagnetic field $A^\mu = (\varphi, \mathbf{A})$, the equation becomes:

$$\left[\left(i\frac{\partial}{\partial t} - e\varphi\right) - \boldsymbol{\alpha}\cdot(-i\nabla - e\mathbf{A}) - \beta m\right]|\psi\rangle = 0\,. \tag{8.9}$$

The Lorentz invariance of the Dirac equation is easier to see by rewriting it like this:

$$[\gamma^\mu(i\partial_\mu - eA_\mu) - m]|\psi\rangle = 0\,, \tag{8.10}$$

where $\gamma^0 = \beta$, $\gamma^k = \beta\alpha_k$, $k = 1, 2, 3$ are the γ-Dirac matrices:

(8.11)
$$\gamma^0 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix}, \gamma^1 = \begin{pmatrix} 0 & \sigma_x \\ -\sigma_x & 0 \end{pmatrix}, \gamma^2 = \begin{pmatrix} 0 & \sigma_y \\ -\sigma_y & 0 \end{pmatrix}, \gamma^3 = \begin{pmatrix} 0 & \sigma_z \\ -\sigma_z & 0 \end{pmatrix}.$$

The solutions to the Dirac equations have four components:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}, \quad \psi_+ = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \quad \psi_- = \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix}. \tag{8.12}$$

In the non-relativistic limit (for a particle of positive energy) the two lower components ψ_- become negligible, and the two upper components ψ_+ obey the non-relativistic Schrödinger equation for a spin-1/2 particle, with each of ψ_1, ψ_2 representing a component of spin.

An important aspect of the Dirac equation is the existence of solutions with negative energy, which can be illustrated as follows. In a null field consider a solution of the form $\psi(\mathbf{x}) = u(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}$, where $u(\mathbf{p})$ is a 4-component *spinor* independent from $\mathbf{x}$. The latter is determined by the equation:

$$Hu(\mathbf{p}) = Eu(\mathbf{p}), \tag{8.13}$$

where $H = \boldsymbol{\alpha}\cdot\mathbf{p} + \beta m$, and $E = \pm\sqrt{p^2 + m^2}$. The solutions with $E < 0$ correspond to negative energy (for these "negative energy" solutions the two upper components ψ_+ are "small", and the two lower components ψ_- are "large").

The existence of negative energy solutions is problematic, because particles could jump to negative energy states by emitting energy, and actual particles do not behave that way. In order to solve the problem Dirac postulated that all negative energy states are occupied (Dirac's sea), and the Pauli exclusion principle prevents positive energy particles from falling to those states. However, a negative energy particle could absorb energy and jump to a positive state leaving behind a "hole", which would behave as a particle with opposite charge. The hole left by an electron jumping from negative energy state to positive energy was identified with a new particle, the *positron*.

A refinement of these ideas leads to quantum field theory.

9. Quantum Field Theory

9.1. **Introduction.** Quantum field theory is developed under the assumption that fields, and not particles, are fundamental entities in nature. So, in the same way the photon is the particle associated to

the electromagnetic field, the electron will be only a manifestation of some "electron field". There are various reasons to take fields and not particles as fundamental:

(1) Locality. There is no "action at a distance", every action between two separate points in space propagates through a field.
(2) The combination of quantum mechanics and special relativity implies that the number of particles is not constant. This is a consequence of the uncertainty principle together with the relativistic mas/energy relation: a particle trapped in a small box may have an uncertainty in its moment/energy larger than the mass of the particle, and at that scale (the *Compton wavelength*) pairs particle-antiparticle may appear spontaneously.
(3) Particle identity: the exchange of two identical particles (such as two electrons or two photons) leave their joint state unchanged (apart from a possible minus sign). They behave as made of the same "thing"—in fact they are manifestations of the same field.

9.2. **Field dynamics.** A field is a quantity defined at every point of space and time (**x**,t). In classical particle mechanics a finite number of generalized coordinates $q_k(t)$ is used. In field theory those generalized coordinates are replaced with the values of the components of the field $\phi_k(\mathbf{x}, t)$ at each point of space and time. The number of degrees of freedom is now infinite.

The dynamics of the field is governed by a Lagrangian that is a function of $\phi(\mathbf{x}, t)$, $\dot{\phi}(\mathbf{x}, t)$, and $\nabla\phi(\mathbf{x}, t)$, and can be written as follows:

$$L(t) = \int \mathcal{L}(\phi_k, \partial_\mu \phi_k)\, d^3\mathbf{x}\,, \tag{9.1}$$

where $\mathcal{L}$ is called *Lagrangian density*.

The action is

$$S = \int_{t_1}^{t_2} L(t)\, dt = \int \mathcal{L}\, d^3\mathbf{x}\, dt\,. \tag{9.2}$$

The equations of motion are determined by the principle of least action $\delta S = 0$, which yields the following Euler-Lagrange equations:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_k)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_k} = 0\,. \tag{9.3}$$

An example is equation (5.12), where a variational principle is used to derive Maxwell's equations using Lagrangian density $\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$.

*(***** This section is under construction. *****)*

E-mail address: mlerma@math.northwestern.edu