

P528 Notes #0: Notation and Background

David Morrissey

January 2, 2017

1 Notation

Most of this course will deal with highly relativistic systems. For this reason, we will use notation and units that are geared to this situation [1, 2, 3].

1.1 Natural Units

We will almost always express quantities in so-called *natural units*, defined by

$$\hbar = c = 1 , \tag{1}$$

where \hbar is the usual quantum mechanics thing and c is the speed of light. Since \hbar has units of energy times time, $\hbar = 1$ implies that we are measuring time in units of inverse energy. Similarly, $c = 1$ means we are measuring distance in units of time and therefore in units of inverse energy as well. This simplifies dimensional analysis since now all dimensionful quantities can be expressed in units of energy. For example,

$$[E] = [P] = [M] = +1, \quad [L] = [T] = -1, \quad \left[\frac{d}{dx} \right] = +1, \tag{2}$$

where the square brackets denotes the energy dimension of the quantity (in natural units, of course). The specific unit we will use for energy is the electron Volt (eV), corresponding to the energy acquired by an electron passing through a potential difference of one Volt. We will also use keV = 10^3 eV, MeV = 10^6 eV, GeV = 10^9 eV, and TeV = 10^{12} eV.

To put a result back into regular units, just add powers of \hbar and c ($\sim L/T$) until you get what you want. In doing so, it's handy to remember a few things:

$$\hbar c = 1 \simeq 0.197 \text{ GeV} \cdot \text{fm} , \tag{3}$$

$$c = 1 \simeq 3.0 \times 10^{10} \text{ cm} , \tag{4}$$

$$m_p \simeq 0.938 \text{ GeV} \simeq 1.67 \times 10^{-27} \text{ kg} , \tag{5}$$

where $1 \text{ fm} = 10^{-13} \text{ cm}$. Despite our use of natural units, scattering cross sections (which have units of area) will often be expressed in barns (b), with $1 \text{ b} = 10^{-24} \text{ cm}^2$. Sometimes it is also convenient to express temperatures in natural units by setting $k_B = 1$. This implies $300 \text{ K} \simeq (1/40) \text{ eV}$ (or about room temperature). Other useful mnemonics are $m_p \simeq 1 \text{ g}/N_A$ (where $N_A \simeq 6.02 \times 10^{23}$ is Avogadro's number) and $1 \text{ yr} \simeq \pi \times 10^7 \text{ s}$.

1.2 Index Notation – Vectors, Matrices, and More

Index notation will be used a lot in this class, and you'll need to be comfortable with it. As a first example, let us apply it to vectors and matrices.

An n -component vector v can be written as an $n \times 1$ matrix,

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}. \quad (6)$$

Clearly, we can express each of the components of v by the numbers v_i , $i = 1, 2, \dots, n$. Equivalently, specifying the components v_i completely fixes the vector.

The same thing can be done for matrices of any size. For instance, we can write the elements of the $n \times n$ matrix M as M_{ij} :

$$M = \begin{pmatrix} M_{11} & M_{12} & \dots & M_{1n} \\ M_{21} & M_{22} & \dots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \dots & M_{nn} \end{pmatrix}. \quad (7)$$

Each entry in the matrix is written as M_{ij} and is called a *matrix element*. Specifying the matrix elements fixes the matrix.

Index notation can also be used to write the products of vectors and matrices. The dot product of a pair of vectors can be thought of as the matrix product of the transpose of the first with the second:

$$u \cdot v = u^t v = (u_1, u_2, \dots, u_n) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \sum_i u_i v_i. \quad (8)$$

From the last equality, written in index notation, it is obvious that $u \cdot v = v \cdot u = v^t u$. Note as well that the label we use for the index that is summed over (i in this case) does not matter: $\sum_i u_i v_i = \sum_j u_j v_j$. For this reason, indices that are summed over are often called *dummy indices*.

The product of an $n \times n$ matrix M with a column vector v is itself a column vector (Mv). The components of this vector are given by

$$(Mv)_i = \sum_j M_{ij} v_j. \quad (9)$$

Note here that j is a dummy index (that we can rename), while i is a *fixed index* that must match up on both sides of the equation. Here, i labels the elements of the vector (Mv).

Note: be careful not to use a fixed index to label a dummy index because you will get horribly confused and mistaken! For example

$$(Mv)_i = \sum_k M_{ik}v_k \quad (10)$$

$$\neq \sum_i M_{ii}v_i . \quad (11)$$

The product of two matrices M and N is a matrix (MN) with elements

$$(MN)_{ij} = \sum_k M_{ik}N_{kj} = \sum_\ell M_{i\ell}N_{\ell j} . \quad (12)$$

Here, the i and j indices are fixed and must match up on both sides of the equation, while the k index that is summed over is a dummy index (that we can also call ℓ or anything else).

Index notation is also useful for expressing various matrix operations:

$$(M^t)_{ij} = M_{ji} , \quad (13)$$

$$(M^*)_{ij} = M_{ij}^* , \quad (14)$$

$$(M^\dagger)_{ij} = M_{ji}^* , \quad (15)$$

$$tr(M) = \sum_i M_{ii} . \quad (16)$$

For the special case of 3-vectors, we have a cross product operation that takes a pair of vectors and makes another. It can be written in terms of indices using the *antisymmetric tensor* ϵ_{ijk} with

$$\epsilon_{123} = +1 = \epsilon_{231} = \epsilon_{312} , \quad (17)$$

$$\epsilon_{132} = -1 = \epsilon_{213} = \epsilon_{321} , \quad (18)$$

and all other entries equal to zero. This name comes from the fact that you get a factor of -1 whenever you exchange a pair of indices in ϵ . The cross product is then given by

$$(a \times b)_i = \sum_{j,k} \epsilon_{ijk} a_j b_k . \quad (19)$$

If you are not convinced by this, work out the components explicitly.

1.3 Index Notation - Relativistic

We will mostly study systems that are invariant under special relativity, meaning that the underlying equations of the system take the same form after applying Lorentz transformations (boosts and rotations). Discussing such systems is much easier if we use an index

notation appropriate to the underlying mathematical structure. Instead of writing (t, x, y, z) for a specific point in space and time, we will use

$$x^\mu = (t, x, y, z), \quad \mu = 0, 1, 2, 3. \quad (20)$$

We call this a (position) 4-vector, and it is useful because the components of x^μ transform linearly into each other under Lorentz transformations:

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu_\nu x^\nu, \quad (21)$$

where Λ^μ_ν is a transformation matrix corresponding to some combination of boosts and rotations. Below, we will discuss the conditions that Λ^μ_ν must satisfy to count as a Lorentz transformation.

Any four-component object that transforms according to Eq. (21) is called a 4-vector. A second important example is the momentum 4-vector:

$$p^\mu = (E, p^x, p^y, p^z), \quad (22)$$

where $\vec{p} = (p^x, p^y, p^z)$ is the spatial (3-)momentum of the system and E is the energy. Recall that for a relativistic particle of mass m , we have $E = \sqrt{m^2 + \vec{p}^2}$.

For any 4-vector $a^\mu = (a^0, a^1, a^2, a^3)$ with an upper index, we define a corresponding 4-vector with a lower index by

$$a_\mu = \eta_{\mu\nu} a^\nu, \quad (23)$$

where the two-index object $\eta_{\mu\nu}$ has components

$$\eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (24)$$

In writing Eq. (23), we have also used Einstein's summation convention, where we implicitly sum over repeated indices.¹ In full gory detail, Eq. (23) is equal to

$$a_\mu = \eta_{\mu\nu} a^\nu \quad (25)$$

$$= \sum_{\nu=0}^4 \eta_{\mu\nu} a^\nu \quad (26)$$

$$= \begin{cases} a^0 & ; \quad \mu = 0 \\ -a^i & ; \quad \mu = i = 1, 2, 3 \end{cases} \quad (27)$$

It is really important to keep upper and lower indices distinct because it will turn out that they refer to different transformation properties under Lorentz. It is also conventional to use Greek letters for the indices of 4-vectors and Roman letters for the components of 3-vectors.

¹ Apparently Einstein considered this his greatest contribution to physics.

In the same way that we lower indices, we can also use the inverse of $\eta_{\mu\nu}$, called $\eta^{\mu\nu}$, to raise indices. It is not hard to see that viewed as a matrix, the components of both are the same: $\eta_{\mu\nu} = \eta^{\mu\nu}$. By definition of the inverse, we have

$$\eta^{\mu\kappa}\eta_{\kappa\nu} = \delta_{\nu}^{\mu} \quad , \quad \eta_{\mu\kappa}\eta^{\kappa\nu} = \delta_{\mu}^{\nu} \quad (28)$$

where $\delta_{\nu}^{\mu} = \delta_{\nu}^{\mu}$ is the 4×4 unit matrix. Raising an index with $\eta^{\mu\nu}$ is consistent with our previous definition of lowering in that if we lower an index and then raise it again, we get back what we started with:

$$a^{\mu} = \eta^{\mu\nu}a_{\nu} = \eta^{\mu\nu}(\eta_{\nu\kappa}a^{\kappa}) = \delta_{\kappa}^{\mu}a^{\kappa} = a^{\mu} . \quad (29)$$

These manipulations might seem trivial, but they actually have a lot of content.

A 4×4 matrix Λ_{ν}^{μ} must satisfy certain conditions for it to be a Lorentz transformation. To describe these conditions, it helps to define a 4-vector dot product by

$$a \cdot b = \eta_{\mu\nu}a^{\mu}b^{\nu} \quad (30)$$

$$= a_{\nu}b^{\nu} = a^{\mu}b_{\mu} \quad (31)$$

$$= a^0b^0 - \vec{a} \cdot \vec{b} . \quad (32)$$

The defining property of a Lorentz transformation is that it leaves the dot product of any two 4-vectors invariant. To be precise, if

$$a^{\mu} \rightarrow a'^{\mu} = \Lambda_{\nu}^{\mu}a^{\nu} \quad \text{and} \quad b^{\mu} \rightarrow b'^{\mu} = \Lambda_{\nu}^{\mu}b^{\nu} , \quad (33)$$

the matrix Λ_{ν}^{μ} must be such that

$$a' \cdot b' = a \cdot b . \quad (34)$$

This implies the conditions

$$\eta_{\mu\nu} \Lambda_{\lambda}^{\mu} \Lambda_{\kappa}^{\nu} = \eta_{\lambda\kappa} . \quad (35)$$

Equivalently, $\eta_{\nu\lambda} \eta^{\mu\kappa} \Lambda_{\kappa}^{\lambda} := \Lambda_{\nu}^{\mu} = (\Lambda^{-1})_{\nu}^{\mu}$.

Going back to our definition of the momentum 4-vector, taking its dot product with itself gives

$$p^2 := p \cdot p = E^2 - \vec{p}^2 . \quad (36)$$

For a relativistic particle, this is just equal to m^2 , the square of the particle (“rest”) mass which does not change under Lorentz transformations.

Using the chain rule, it is simple to show that a derivative with respect to x^{μ} transforms like a 4-vector with a lower index. As a result, we will use the simplified notation

$$\partial_{\mu} := \frac{\partial}{\partial x^{\mu}} \quad (37)$$

$$\partial^2 := \eta^{\mu\nu} \partial_{\mu} \partial_{\nu} = \partial_t^2 - \vec{\nabla}^2 \quad (38)$$

This implies

$$\partial_\mu x^\nu = \delta_\mu^\nu . \quad (39)$$

Note that we treat t and \vec{x} as independent variables. Thinking ahead to quantum mechanics, we will later make the identification $P^\mu = (P^0, \vec{P}) \sim i(\partial_t, -\vec{\nabla}) = i\partial^\mu$ as operators in a position-space basis.

2 Relativistic Kinematics and Scattering

The most important observables related to the experimental study of the SM are scattering cross sections and related quantities. We review here some of the main points you should be familiar with for computing them.

2.1 Relativistic Kinematics

Energy and momentum conservation carry over to relativistic systems. In particular, the sum of the 4-momenta of all the particles in the initial state of a system must be equal to the sum of the 4-momenta of all the particles in the final state. This is straightforward to apply, and it leads to important constraints [4]. We will simply illustrate it with a couple of examples.

e.g. 1. Inelastic scattering $aa \rightarrow bb$ in the CM Frame

Consider the process $aa \rightarrow bb$, where the a particles are massless and the b particles have mass m . Let us label the initial 4-momenta by p_1 and p_2 and the final 4-momenta by p_3 and p_4 . In the centre-of-mass (CM) frame, the net 3-momentum of the initial state vanishes by definition. Choosing the z -axis to align with the collision axis of the initial a particles, we have

$$p_1 = (p, 0, 0, p) , \quad p_2 = (p, 0, 0, -p) , \quad (40)$$

for some initial 3-momentum magnitude p . Note as well that $E = p$ for these particles because they are massless. Momentum conservation implies that the sum of the 3-momenta of the outgoing particles must also vanish, and thus the final-state 4-vectors must take the form

$$p_3 = (E_3, \vec{k}) , \quad p_4 = (E_4, -\vec{k}) . \quad (41)$$

The usual relation between the relativistic energy and 3-momentum of a particle gives $E_3 = \sqrt{m^2 + \vec{k}^2} = E_4$. At this point, we have already balanced the 3-momenta of the initial and final states, but we still have to make sure that the energies match up. This gives the relation

$$p_1^0 + p_2^0 = E_3 + E_4 , \quad (42)$$

which after simplifying and solving for $k = |\vec{k}|$ gives

$$k = p \sqrt{1 - \left(\frac{m}{p}\right)^2} . \quad (43)$$

This solution only makes sense for $p \geq m$, since at least this much energy is needed to create a pair of b particles of mass m . Note as well that energy and momentum conservation fix the magnitudes and energies of the outgoing particles, but they do not fix their directions other than forcing them to be back-to-back.

e.g. 2. Particle Decays

Suppose particle A of mass M decays to a pair of B particles of mass m , $A \rightarrow BB$. In the rest frame of A , the initial 4-momentum is

$$p = (M, \vec{0}) . \quad (44)$$

Conservation of 3-momentum then implies that the two outgoing B particles must be back-to-back in this frame so that their net 3-momentum vanishes. Choosing the z -axis to align with the direction of one them, we can write the 4-momenta of the final-state particles as

$$k_1 = (E_1, 0, 0, k) , \quad k_2 = (E_2, 0, 0, -k) , \quad (45)$$

with $E_1 = \sqrt{m^2 + k^2} = E_2$. Applying total energy conservation then gives $M = E_1 + E_2$, implying

$$k = \left(\frac{M}{2}\right) \sqrt{1 - \left(\frac{2m}{M}\right)^2} . \quad (46)$$

This only makes sense for $M \geq 2m$.

2.2 Cross Sections and Decay Rates

To compute a scattering cross section or particle decay rate, one follows the standard quantum mechanics procedure of computing the transition matrix element \mathcal{M} for the process, and then summing it over all possible final states. For particle scattering or decays, the final state is typically a set of well-separated particles, and it is usually and excellent approximation to treat them as 4-momentum eigenstates. The sum over final states then reduces to a set of integrals over the space of outgoing momenta constrained by overall 4-momentum conservation.

For $2 \rightarrow n$ scattering with two initial particles colliding to make a final state with n particles, let us label the initial 4-momenta by p_1 and p_2 and the final 4-momenta by p_3 - p_{n+2} . The formula for the scattering cross-section is

$$\sigma = \frac{S}{v} \frac{1}{4E_1 E_2} \int \frac{d^3 p_3}{(2\pi)^3 2E_3} \cdots \int \frac{d^3 p_{n+2}}{(2\pi)^3 2E_{n+2}} (2\pi)^4 \delta^{(4)}(k_1 + k_2 - \sum_{i=3}^{n+2} p_i) |\mathcal{M}|^2 , \quad (47)$$

where $v = \sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2} / E_1 E_2$ is the magnitude of the relative velocity of the incident particles, and S is a combinatoric factor equal to one times $1/k!$ for every set of k identical particles in the final state. Derivations of this result can be found in the textbooks by Peskin & Schroeder [1] and Srednicki [2].

The result of Eq. (47) has a lot going on, but its physical content is very simple. First, $|\mathcal{M}|^2$ is the probability density for a single initial state $(p_1 + p_2)$ to scatter into the specific final state $(p_3 + \dots + p_{n+2})$. The delta function enforces overall 4-momentum conservation. The scattering probability density is then summed over all distinct final states with a relativistic normalization. Collectively, this set of final states is often called the *phase space*. The prefactor before the integrations is a normalization to convert the result for a single initial state to the scattering probability rate per unit incident flux (= number of incident particles per unit area per unit time). At the end of the day, the cross section has units of area. The factor of S accounts for sets of indistinguishable particles.

e.g. 3. Cross Section for $aa \rightarrow bb$ in the CM Frame

Let us return to the process $aa \rightarrow bb$ in the CM frame, where a is massless and b has mass m . We computed the kinematics of the final state above in *e.g.* 1. Assume further that the matrix element for the process is $\mathcal{M}(p_1, p_2, p_3, p_4)$. Applying the formula above, we obtain

$$\sigma = \frac{(1/2!)}{2} \frac{1}{4p^2} \int \frac{d^3 p_3}{(2\pi)^3 2E_3} \int \frac{d^3 p_4}{(2\pi)^3 2E_4} (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) |\mathcal{M}|^2, \quad (48)$$

where we have used $v = 2$, $S = 1/2!$, and $E_1 = E_2 = p$. To simplify this, we use three of the four delta functions to eliminate the integral over \vec{p}_4 , forcing $\vec{p}_3 = -\vec{p}_4$ and $E_3 = E_4 = \sqrt{m^2 + k^2}$. This leads to

$$\begin{aligned} \sigma &= \frac{1}{64} \frac{1}{(2\pi)^2} \int d\Omega_3 \int_0^\infty d\|\vec{p}_3\| \frac{1}{E_3^2} \delta(2p - 2E_3) |\mathcal{M}|^2 \\ &= \frac{1}{128} \frac{1}{(2\pi)^2} \frac{k}{E_3} \int d\Omega_3 |\mathcal{M}|^2, \end{aligned} \quad (49)$$

where Ω_3 is the solid angle corresponding to the direction of \vec{p}_3 and k is the magnitude of \vec{p}_3 consistent with 4-momentum conservation found above in Eq. (43). To get to the last line, we have used the delta-function trick of Eq. (58) below.

The quantity of interest for particle decays is the average decay rate Γ . Given an initial sample of N_0 particles at time $t = 0$, the number of particles after time t is

$$N(t) = N_0 e^{-\Gamma t}. \quad (50)$$

The lifetime τ of a particle species is defined to be

$$\tau = 1/\Gamma. \quad (51)$$

Sometimes you will also hear of half-lives, given by $\tau_{1/2} = \tau \ln 2$. In natural units, the decay rate has units of mass.

When a particle has more than one distinct decay mode, we also speak of partial decay rates Γ_f , corresponding to the relative probability of decaying in that way. The total decay rate is the sum of the partial rates of all the individual decay channels,

$$\Gamma = \sum_f \Gamma_f = \Gamma \sum_f BR_f , \quad (52)$$

where $BR_f = \Gamma_f/\Gamma$ is the *branching ratio* to the final state f .

The formula for the partial decay rate of an unstable particle of mass M at rest to decay to a final state containing n particles ($M \rightarrow 1 + 2 + \dots + n$) is

$$\Gamma(M \rightarrow n) = \frac{S}{2M} \int \frac{d^3 p_1}{2E_1(2\pi)^3} \cdots \int \frac{d^3 p_n}{2E_n(2\pi)^3} (2\pi)^4 \delta^{(4)}(p_M - \sum_{i=1}^n p_i) |\mathcal{M}|^2 , \quad (53)$$

where $|\mathcal{M}|^2$ is the corresponding $M \rightarrow n$ amplitude defined in the same way as for scattering, and S is the symmetry factor.

A Appendix: Useful Math

The mathematics to be used in this course will consist mostly of vector calculus, together with a bit of group theory. We will get to the group theory later on, so let us just mention a few useful results in vector calculus [5].

We will write n -dimensional integrals over spacetime (one time and $(n - 1)$ spatial dimensions) as

$$\int d^n x = \int dt \int dx^1 \int dx^2 \dots \int dx^{n-1} \quad (54)$$

The divergence theorem generalizes to this case: for any n -dimensional volume V with boundary ∂V , we have

$$\int_V d^n x \partial_\mu f^\mu = \int_{\partial V} dA n_\mu f^\mu , \quad (55)$$

where n_μ is the outwardly pointing normal unit vector to the bounding surface ∂V . In most cases of interest to us, the surface integral over ∂V will be zero for one reason or another.

Multi-dimensional delta functions will appear frequently:

$$\int d^n x \delta^{(n)}(x - x') f(x) = f(x') . \quad (56)$$

Equivalently $\delta^{(n)}(x - x') = \prod_{\mu=0}^{n-1} \delta(x^\mu - x'^\mu)$. We also have the Fourier transform relation

$$\int d^n x e^{ik \cdot x} = (2\pi)^n \delta^{(n)}(k), \quad \int d^n k e^{ik \cdot x} = (2\pi)^n \delta^{(n)}(x) . \quad (57)$$

Finally, recall that

$$\int dx \delta(f(x)) g(x) = \sum_a g(x_a) / |df/dx|_{x_a} , \quad (58)$$

where x_a refers to any value of x such that $f(x_a) = 0$.

Sometimes it will be useful to use contour integration in the complex plane. The key result is Cauchy's theorem, which is that for any integral around a closed counterclockwise contour C in the complex plane,

$$f(z_0) = \frac{1}{2\pi i} \oint_C dz \frac{f(z)}{(z - z_0)} , \quad (59)$$

provided $f(z)$ has no poles in the region bounded by the contour. If the contour runs clockwise, the formula picks up an overall minus sign. Differentiating both sides with respect to z_0 gives

$$\frac{d^n}{dz^n} f(z_0) = \frac{n!}{2\pi i} \oint_C dz \frac{f(z)}{(z - z_0)^{n+1}} . \quad (60)$$

This is really useful because it also applies to any deformation of the contour C that does not intersect the pole. Sometimes the stuff in the integrand (with the pole factored off) is called the *residue* of the pole. Cauchy's theorem is therefore often stated as

$$\oint_C dz f(z) = 2\pi i \times (\text{sum of residues}) . \quad (61)$$

e.g. 4. $I = \int_0^\infty dx 1/(1 + x^2)$

We can do this integral easily using Cauchy's theorem. See Fig. 1 for an illustration of the contours C_1 and C_2 .

$$\begin{aligned} I &= \frac{1}{2} \int_{-\infty}^{\infty} dx \frac{1}{1 + x^2} \\ &= \frac{1}{2} \lim_{R \rightarrow \infty} \int_{-R}^{+R} dx \frac{1}{1 + x^2} = \lim_{R \rightarrow \infty} \int_{C_1} dz \frac{1}{1 + z^2} \\ &= \frac{1}{2} \lim_{R \rightarrow \infty} \left[\oint_{C=C_1+C_2} dz \frac{1}{(z+i)(z-i)} - \int_{C_2} dz \frac{1}{1 + z^2} \right] \\ &= \frac{1}{2} \times 2\pi i \times \frac{1}{2i} - 0 \\ &= \frac{\pi}{2} . \end{aligned}$$

In the last line, we have used the fact that the contour integration along C_2 can be parametrized by $z = Re^{i\theta}$ and goes like

$$\lim_{R \rightarrow \infty} \int_{C_2} dz \frac{1}{1 + z^2} = \lim_{R \rightarrow \infty} \int d\theta iRe^{i\theta} \frac{1}{1 + R^2 e^{2i\theta}} \rightarrow 0 .$$

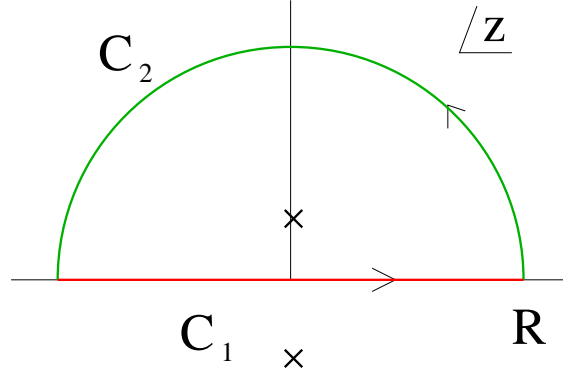


Figure 1: Contours for *e.g.* 4.

B Appendix: Useful Physics

Having established notation, we turn next to a quick review of the essential physics background for this course. If you are not familiar with this material, you should come speak to me as soon as possible.

B.1 Electromagnetism

Maxwell's equations [6]:

$$\vec{\nabla} \cdot \vec{E} = \rho \quad (62)$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (63)$$

$$\vec{\nabla} \cdot \vec{B} = 0 \quad (64)$$

$$\vec{\nabla} \times \vec{B} = \frac{\partial \vec{E}}{\partial t} + \vec{j} \quad (65)$$

where ρ is the charge density and \vec{j} is the charge current. In writing the equations in this way, we have implicitly also set $\epsilon_0 = \mu_0 = 1$ (consistent with $c = 1/\sqrt{\epsilon_0\mu_0}$).

The electric and magnetic fields can be written in terms of electric and magnetic potentials:

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}, \quad \vec{B} = \vec{\nabla} \times \vec{A}. \quad (66)$$

Down the road, we will combine the electric and magnetic potentials into a single 4-vector

$$A^\mu = (\phi, \vec{A}). \quad (67)$$

With these conventions, the electric potential due to a stationary point charge Q is

$$\phi = \frac{Q}{4\pi r}. \quad (68)$$

B.2 Classical Mechanics

In the Lagrangian formulation of classical mechanics, a physical system is described by a set of generalized coordinates $q_i(t)$. The time evolution of the system is determined by the action [7]:

$$S = \int_a^b dt L(q_i, \dot{q}_i) . \quad (69)$$

In many cases, the Lagrangian L is equal to $L = T - V$. The equations of motion that describe the time evolution of the coordinates $q_i(t)$ are derived from the principle of stationary action, namely that the physical configuration is the one that satisfies $\delta S = 0$ for *any* infinitesimal variation $q_i(t) \rightarrow q_i(t) + \delta q_i(t)$ (with $\delta q_i = 0$ for $t = a, b$). Applying this condition to the action gives the equations of motion:

$$\partial_t \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0 , \quad (70)$$

for each value of i . This gives one second-order equation for each coordinate $q_i(t)$. The solution is unique once initial values for $q_i(t)$ and $\dot{q}_i(t)$ are specified.

A second way to formulate classical mechanics is in terms of a Hamiltonian H that depends on the coordinates $q_i(t)$ and $p_i(t)$ [7]. The Hamiltonian formulation can be derived from the Lagrangian, and vice versa. Given a Lagrangian $L(q_i, \dot{q}_i)$, the generalized momenta are defined to be

$$p_i = \frac{\partial L}{\partial \dot{q}_i} . \quad (71)$$

Using these equations, it is possible to write each \dot{q}_i as a function of the p_i and q_i . The Hamiltonian $H(q_i, p_i)$ is then defined to be

$$H = \sum_i \dot{q}_i p_i - L , \quad (72)$$

where all the \dot{q}_i s on the right-hand side are to be viewed as functions of the q_i and p_i . The equations of motion in this formulation can be written in terms of Poisson brackets, defined according to

$$\{f, g\}_{PB} = \sum_j \left(\frac{\partial f}{\partial q_j} \frac{\partial g}{\partial p_j} - \frac{\partial f}{\partial p_j} \frac{\partial g}{\partial q_j} \right) . \quad (73)$$

With this definition, the equations of motion are

$$\dot{q}_i = \{q_i, H\}_{PB}, \quad \dot{p}_i = \{p_i, H\}_{PB} . \quad (74)$$

Two useful facts to keep in mind are

$$\{q_i, p_j\}_{PB} = \delta_{ij} , \quad (75)$$

and for any function $f(q, p, t)$

$$\frac{df}{dt} = \{f, H\}_{PB} + \frac{\partial f}{\partial t} . \quad (76)$$

B.3 Quantum Mechanics

The rules of quantum mechanics (QM) for relativistic systems are exactly the same as what you have learned in previous classes. However, the procedure we will follow to formulate a consistent quantum theory of relativistic particles will probably be new to you. In most QM courses, you are typically given a system with a Hamiltonian and asked to solve for the corresponding energy eigenvalues and eigenstates. Instead, the major challenge to be addressed in this course will be figuring out how to build the system itself and to deduce what its Hamiltonian should be. So, even though the way we approach things might seem new and unusual, keep in mind that the basic underlying rules are the same as ever.

Recall that a quantum mechanical system consists of a set of states (comprising a Hilbert space) together with the set of operators acting on them [8]. Starting from a classical system defined in terms of q_i , p_j , and H , we can construct a corresponding quantum system by elevating q_i and p_j to operators on the Hilbert space satisfying

$$[q_i, p_j] = i \delta_{ij} . \quad (77)$$

Notice the similarity to the Poisson brackets. The Hamiltonian H is just a function of the q_i and p_j , and is now also a well-defined operator on the Hilbert space of states (up to possible ambiguities in the order of the q_i and p_j).

Quantum states can be viewed as vectors in a (possibly infinite-dimensional) *vector space*, and observables correspond to operators on the space. The vector spaces that arise in quantum mechanics also have an *inner product* defined on them. A key result in linear algebra that we will use a lot is that the eigenstates of a Hermitian operator form a complete basis for the vector space. For example, suppose $\{|n\rangle\}$ are the eigenstates of the Hamiltonian operator H , with $H|n\rangle = E_n|n\rangle$. Completeness means that any state can be expanded in terms of them:

$$|\psi\rangle = \sum_n a_n |n\rangle , \quad (78)$$

where the coefficients a_n are just numbers. We can also choose to arrange the energy eigenstates to be orthonormal, $\langle m|n\rangle = \delta_{mn}$. In this case, we have

$$a_n = \langle n|\psi\rangle . \quad (79)$$

A convenient way to express completeness and orthonormality (which applies to any set of orthonormal basis states) is the *resolution of the identity*:

$$\mathbb{I} = \sum_n |n\rangle\langle n| , \quad (80)$$

where \mathbb{I} is the identity operator. Acting the identity, written in this form, on $|\psi\rangle$ then returns Eq. (78) with the coefficients as in Eq. (79).

Time evolution in quantum mechanics is determined by the Hamiltonian. There are two popular ways to describe this. The first is the Schrödinger picture in which the operators are time-independent while the state of the system evolves according to

$$|\alpha(t)\rangle = e^{-iH(t-t_0)}|\alpha(t_0)\rangle , \quad (81)$$

where t_0 is some reference time. Equivalently, states satisfy the Schrödinger equation,

$$\frac{d}{dt}|\alpha(t)\rangle = -iH|\alpha(t)\rangle . \quad (82)$$

The second standard way to describe time evolution in QM is the Heisenberg picture. Here, the states are time independent while the operators depend on time. Specifically, we have $|\alpha(t)\rangle = |\alpha(t_0)\rangle$ together with

$$\mathcal{O}(t) = e^{iH(t-t_0)}\mathcal{O}(t_0)e^{-iH(t-t_0)} . \quad (83)$$

This relation is equivalent to the commutation relation

$$\frac{d}{dt}\mathcal{O} = -i[\mathcal{O}, H] , \quad (84)$$

which should remind you of the classical Poisson bracket result.

C Appendix: Classical Fields

Most of the topics of this course will be based on quantum field theory (QFT). To organize the QFT describing the Standard Model (SM), we will make particular use of symmetries. Since the treatment of symmetries in QFTs is very similar to that in classical field theories, we review some of the key features of classical field theories here. A more detailed treatment of these topics can be found in Ref. [9]

C.1 Actions and Equations of Motion

Classical continuous systems can be described in terms of an action, just like we do for discrete (point particle) systems. In particular, the equations of motion for a continuous system can be derived from the *principle of least action* based on a Lagrangian or Hamiltonian.

Suppose our system is described by the field $\phi(t, \vec{x})$ defined in four $(3 + 1)$ spacetime dimensions. To apply Lagrangian (or Hamiltonian) mechanics to it, it is just a matter of thinking about it as an infinite set of generalized time-dependent coordinates labelled by \vec{x} :

$$\phi(t, \vec{x}) = q_{\vec{x}}(t) . \quad (85)$$

The time evolution of the field is specified by the action, which for a theory that maintains locality and causality must take the form

$$S[\phi] = \int d^4x \mathcal{L}(\phi, \partial_\mu\phi) , \quad (86)$$

where \mathcal{L} is called the *Lagrangian density* and $d^4x = dt d^3x$. Applying the principle of least action, we obtain the classical equation of motion,

$$0 = \frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right] . \quad (87)$$

If there are multiple fields, there is an equation of motion of this form for each field.

The basic action for a relativistic real scalar field ϕ is

$$S = \int d^4x \left[\frac{1}{2} \eta^{\mu\nu} (\partial_\mu \phi) (\partial_\nu \phi) - V(\phi) \right] , \quad (88)$$

where d^4x runs over all spacetime, and ϕ and $\partial_\mu \phi$ are assumed to vanish at spacetime infinity. To cut down on writing, we will often use the standard abbreviation $\eta^{\mu\nu} (\partial_\mu \phi) (\partial_\nu \phi) = (\partial\phi)^2$. Applying the general expression for the equation of motion of Eq. (87) here, we find

$$(\partial_t^2 - \vec{\nabla}^2)\phi = -\frac{\partial V}{\partial \phi} . \quad (89)$$

For general $V(\phi)$, this is a complicated non-linear partial differential equation.

In the specific case of $V(\phi) = (m^2/2)\phi^2$, the equation of motion linearizes and can be solved exactly. It becomes

$$(-\partial^2 - m^2)\phi = 0 . \quad (90)$$

This is called the Klein-Gordon equation. A solution is

$$\phi(x) = a(k) e^{-ik \cdot x} \quad (91)$$

with $k^0 = \pm \sqrt{\vec{k}^2 + m^2}$. This looks like the energy E_k of a relativistic point particle with mass m and 3-momentum \vec{k} . The most general solution is obtained by taking linear combinations of these specific solutions and making sure the result is real (since $\phi(x)$ itself is real):

$$\phi(x) = \int \widetilde{dk} \left[a(\vec{k}) e^{-ikx} + a^*(\vec{k}) e^{ikx} \right] , \quad (92)$$

with $k^0 = E_k = +\sqrt{\vec{k}^2 + m^2}$ and

$$\widetilde{dk} = \frac{d^3k}{2E_k (2\pi)^3} . \quad (93)$$

This funny combination has been chosen to obtain an integration measure that is Lorentz invariant. While the basic d^3k measure is not, the following combination is:

$$\int d^4k \Theta(k^0) \delta(k^2 - m^2) = \int d^3k \frac{1}{2E_k} , \quad E_k = \sqrt{\vec{k}^2 + m^2} . \quad (94)$$

Another relativistic action that we will study down the road is that for electromagnetism. Let us combine the scalar and vector potentials into a single 4-vector:

$$A^\mu = (\phi, \vec{A}) . \quad (95)$$

In terms of them, we define the field strength tensor to be

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu . \quad (96)$$

Note that it is antisymmetric: $F_{\mu\nu} = -F_{\nu\mu}$. The action for electromagnetism is

$$S = \int d^4x \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right) . \quad (97)$$

This is invariant under Lorentz transformations because all the indices are contracted. The equations of motion derived from this action turn out to be Maxwell's Equations! (More precisely they are Maxwell's equations with no sources, $\rho = \vec{j} = 0$.)

A complication that arises in electromagnetism is that the basic variables we have used to describe the theory, A^μ , are not unique. In particular, transforming A^μ by

$$A_\mu \rightarrow A_\mu + \partial_\mu \alpha , \quad (98)$$

leads to the same electric and magnetic fields for any function $\alpha(x)$. Fortunately, such transformations do not change $F_{\mu\nu}$, and therefore we obtain the same equations of motion. The physical interpretation is that the A^μ are redundant variables, with different values able to describe the same physical configuration. This implies that all physical observables should be unaffected by the *gauge transformation* of Eq. (98).

C.2 Symmetries and Noether's Theorem

Symmetries play a key role in physics, and they have interesting implications for field theories (both classical and quantum). A symmetry is a transformation of the system that leaves the physics the same. For classical theories, "the same physics" means that the equations of motion should have the same form both before and after the transformation. We will consider both discrete and continuous symmetries, but the continuous case will turn out to be more interesting.

e.g. 5. A discrete symmetry of our simple scalar theory.

Recall that the Lagrangian (density) was $\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2$. This is clearly unaffected by $\phi \rightarrow \phi' = -\phi$, and the equations of motion for ϕ' are identical to those for ϕ . On the other hand, this would not be a symmetry if V contained a term proportional to ϕ^3 .

A necessary and sufficient condition for a transformation to be a symmetry is that the action written in terms of the transformed coordinates take the same form as the original

theory. There are two ways to think about this. The first is the *active picture*, where the configuration of the system is changed. For an action that is a functional of a set of fields $\{\phi_i\}$, this means that we explicitly change the fields

$$\phi_i \rightarrow \phi'_i = f_i(\phi) . \quad (99)$$

In turn, this changes the action according to

$$S[\phi] \rightarrow S[\phi'] := S'[\phi] , \quad (100)$$

where $S'[\phi]$ is some new functional of the original (untransformed) fields ϕ_i . In this active picture, a transformation is a symmetry if and only if

$$S'[\phi] = S[\phi] . \quad (101)$$

A standard example of an active transformation is performing a rotation on the system.

The second way to think of symmetries is in terms of the *passive picture*. Here, the idea is that we don't actually modify the system, but rather that we just use different variables to describe it. Suppose $\phi_i = f_i(\phi')$, for some new set of fields $\{\phi'_j\}$. Rewriting the action in terms of these new field variables, we find

$$S[\phi] = S[f(\phi')] = S'[\phi'] . \quad (102)$$

Such a transformation is a symmetry of the system if and only if

$$S'[\phi'] = S[\phi'] . \quad (103)$$

The standard example of a passive transformation is choosing a rotated set of axes to describe the system. Relative to the new axes, the coordinates that describe the system are different even though the system has not changed at all. You should convince yourself that both pictures are equivalent.

e.g. 6. A continuous symmetry for two fields.

Consider a theory with two real fields ϕ_1 and ϕ_2 :

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} [(\partial\phi_1)^2 + (\partial\phi_2)^2] - \frac{1}{2}m^2 (\phi_1^2 + \phi_2^2) \\ &= \frac{1}{2}(\partial\phi)^t(\partial\phi) - \frac{1}{2}m^2\phi^t\phi , \end{aligned}$$

where $\phi = (\phi_1, \phi_2)^t$. This is clearly invariant under the discrete transformation $\phi \rightarrow -\phi$. It is also unchanged under transformations of the form

$$\phi \rightarrow \phi' = \mathcal{O}\phi , \quad (104)$$

where \mathcal{O} is any 2×2 orthogonal matrix – satisfying $\mathcal{O}^t\mathcal{O} = \mathbb{I}$. These are called continuous transformations because they can be parametrized by a continuous parameter. In

particular, up to a few signs, we can write any orthogonal matrix \mathcal{O} in terms of the parameter α :

$$\mathcal{O} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}. \quad (105)$$

Note as well that this is the active picture, and we have just argued that $S[\phi] \rightarrow S[\mathcal{O}\phi] = S'[\phi] = S[\phi]$. Alternatively, in the passive picture we would write

$$\phi = \mathcal{O}\phi'', \quad (106)$$

for some new field coordinates ϕ'' . Now, $S[\phi] = S[\mathcal{O}\phi'']$ by assumption, but the fact that this is a symmetry corresponds to $S[\mathcal{O}\phi''] := S''[\phi''] = S[\phi'']$.

Continuous symmetries are especially interesting because they imply conservation laws. This relationship is called *Noether's theorem*. Consider a continuous (active) transformation described by the parameter α :

$$\phi_i(x) \rightarrow \phi'_i(x) = \phi_i(x) + \alpha \Delta \phi_i(x), \quad (107)$$

where we will treat $|\alpha| \ll 1$ and work to linear order. This will change the action by changing the Lagrangian, which we assume can be written as a Lagrangian density. For the transformation to be a symmetry we must have $S[\phi'] := S'[\phi] = S[\phi]$. Equivalently, the Lagrangian density should obey

$$\mathcal{L}(\phi') := \mathcal{L}'(\phi) = \mathcal{L}(\phi) + \alpha \partial_\mu K^\mu. \quad (108)$$

Note that the Lagrangian density can change by a total divergence; this will only produce a vanishing surface term in the action, and will therefore yield the same equations of motion. Plugging the form of Eq. (107) into Eq. (108) and applying the equations of motion, one finds [1, 9]

$$0 = \partial_\mu j^\mu, \quad \text{where} \quad j^\mu = \sum_i \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \Delta \phi_i - K^\mu. \quad (109)$$

This result is called *Noether's Theorem*.

The 4-vector j^μ obtained from Noether's theorem is said to be a conserved current. The reason for this is that if we define the corresponding *conserved charge* (not necessarily electric charge!) by

$$Q = \int d^3x j^0, \quad (110)$$

we find that

$$\partial_t Q = \int d^3x \partial_t j^0 = \int d^3x \vec{\nabla} \cdot \vec{j} = 0. \quad (111)$$

Note that we get zero because everything vanishes on the boundary, by assumption. The physical interpretation of $j^\mu = (j^0, \vec{j})$ is that j^0 is a charge density and \vec{j} is a current density.

***e.g.* 7. More of our two-field example**

Applying this general result to the theory of *e.g.* 6, we find that $\mathcal{L}(\phi') = \mathcal{L}(\phi)$, and thus $K^\mu = 0$. For small rotation angles α , we also have

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \rightarrow \begin{pmatrix} \phi'_1 \\ \phi'_2 \end{pmatrix} = \begin{pmatrix} 1 & -\alpha \\ \alpha & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \phi_1 + \alpha\Delta\phi_1 \\ \phi_2 + \alpha\Delta\phi_2 \end{pmatrix}. \quad (112)$$

Thus, $\Delta\phi_1 = -\phi_2$ and $\Delta\phi_2 = \phi_1$. The conserved current is therefore

$$j^\mu = -(\partial^\mu\phi_1)\phi_2 + \phi_1(\partial^\mu\phi_2). \quad (113)$$

It is straightforward to check that this current is indeed conserved.

A particularly important set of continuous transformations are the spacetime translations,

$$x^\lambda \rightarrow x^\lambda - a^\lambda. \quad (114)$$

For nearly all the theories we will study in this course, these translations will be symmetries of the system. For now, let's look specifically at our simple scalar theory. It is easiest to think of the translations as an active shift of the system: $\phi(x) \rightarrow \phi'(x) = \phi(x + a)$ (with the integrals and derivatives in the action unchanged, possible because we integrate over all spacetime). For infinitesimal a^λ , we have

$$\phi'(x) = \phi(x) + a^\lambda\partial_\lambda\phi. \quad (115)$$

Applying this to the Lagrangian, we find

$$\mathcal{L}(\phi') = \mathcal{L}(\phi) + a^\lambda\partial_\mu(\delta_\lambda^\mu\mathcal{L}). \quad (116)$$

Thus, spacetime translations are a symmetry of our theory with $K_\lambda^\mu = \delta_\lambda^\mu\mathcal{L}$. Applying our general result, the corresponding conserved currents are

$$j_\lambda^\mu = \partial^\mu\phi\partial_\lambda\phi - \delta_\lambda^\mu\mathcal{L}. \quad (117)$$

At this point, let us emphasize that we have just considered four different symmetries at once; each value of $\lambda = 0, 1, 2, 3$ corresponds to a different transformation.² In contrast, μ labels the spacetime index that always arises on the current. However, since $\eta_{\mu\nu}j_\lambda^\nu = j_{\mu\lambda} = j_{\lambda\mu}$ in this case, we can afford to be a bit careless with the indices.

For the specific case of time translations, we should take $\lambda = 0$. The corresponding charge is

$$\int d^3x j_0^0 = \int d^3x \left[\frac{1}{2}(\partial_t\phi)^2 + \frac{1}{2}(\vec{\nabla}\phi)^2 + V(\phi) \right]. \quad (118)$$

²This is why we used λ instead of ν .

This is just the Hamiltonian H of the system. Thus, invariance under time translations corresponds to energy conservation, $\dot{H} = 0$. Similarly, for spatial translations the related charge is

$$\int d^3x j_i^0 = \int d^3x (\partial_t \phi) \partial_i \phi, \quad (119)$$

corresponding to a conserved spatial momentum P_i . Given the physical interpretation of j_λ^μ , it is given a special symbol

$$j^{\mu\nu} = T^{\mu\nu}, \quad (120)$$

and is called the *energy-momentum tensor*. The corresponding charges are usually combined into a single conserved 4-vector, $P^\mu = \int d^3x j^{0\mu} = (H, \vec{P})$.

References

- [1] M. E. Peskin and D. V. Schroeder, “An Introduction To Quantum Field Theory,” *Reading, USA: Addison-Wesley (1995) 842 p*
- [2] M. Srednicki, “Quantum field theory,” *Cambridge, UK: Univ. Pr. (2007) 641 p*
- [3] L. H. Ryder, “Quantum Field Theory,” *Cambridge, UK: Univ. Pr. (1985) 443p*
- [4] D. Griffiths, “Introduction to elementary particles,” Weinheim, Germany: Wiley-VCH (2008) 454 p
- [5] J. Mathews, R. L. Walker, “Mathematical Methods of Physics”, *Redwood City, USA: Addison Wesley (1970)*
- [6] D. J. Griffiths, “Introduction to Electrodynamics”, *Saddle River, USA: Prentice Hall (1999)*
- [7] H. Goldstein, “Classical Mechanics”, *Cambridge USA: Addison-Wesley (1950)*
- [8] D. J. Griffiths, “Introduction to Quantum Mechanics”, *Saddle River, USA: Prentice Hall (2004)*
- [9] D. E. Morrissey, “PHYS-526 Notes #1”,
<http://trshare.triumf.ca/~dmorri/Teaching/PHYS526-2013/notes-01.pdf>