B2: Symmetry and Relativity

J Tseng

October 22, 2019

Contents

1 Introduction **2**

1.1 Books . . . 2

1.2 Postulates . . . 2

1.3 *Vector transformation . . . 3

1.4 Symmetry and relativity . . . 4

1.5 Units . . . 5

1.5.1 Other conventions . . . 5

1.6 4-vector basics . . . 5

1.7 *Invariants . . . 7

1.8 Spacetime diagrams . . . 8

1.8.1 *Proper time . . . 8

1.9 Basic 4-vectors . . . 9

1.9.1 Position . . . 9

1.9.2 Velocity . . . 9

1.9.3 Acceleration . . . 10

1.9.4 Momentum . . . 10

1.9.5 Force . . . 11

2 Applications of 4-momentum **12**

2.1 *Conservation of energy-momentum . . . 12

2.2 *Annihilation, decay, and formation . . . 12

2.2.1 *Center of momentum frame . . . 12

2.2.2 Decay at rest . . . 13

2.2.3 In-flight decay . . . 14

2.3 *Collisions . . . 15

2.3.1 Absorption . . . 15
2.3.2 Particle formation . . . 16
2.3.3 *Compton scattering . . . 18

3 Force **20**
3.1 Pure force . . . 20
3.2 Transformation of force . . . 21
3.3 *Force and simple motion problems . . . 22
3.3.1 Motion under pure force . . . 22
3.3.2 Linear motion . . . 23
3.3.3 Hyperbolic motion . . . 24
3.3.4 Constant external force . . . 25
3.3.5 Circular motion . . . 26

4 Lagrangians **28**
4.1 *Equations of particle motion from the Lagrangian . . . 28
4.1.1 Central force problem . . . 31

5 Further kinematics **34**
5.1 *Doppler effect . . . 34
5.2 *Aberration . . . 36
5.3 Headlight effect . . . 36
5.4 Generators . . . 37
5.5 Thomas precession . . . 40

6 Scalars, Vectors, and Tensors **43**
6.1 Generalized Lorentz transformation . . . 43
6.1.1 Index notation . . . 43
6.1.2 Summation convention . . . 45
6.1.3 Metric . . . 45
6.2 Lorentz transformation matrices . . . 46
6.3 Tensors . . . 48
6.4 *4-gradient . . . 49

7 Groups **51**

7.1 Example: permutation group . . . 52
7.2 Rotations . . . 53
7.3 Representations . . . 54
7.3.1 Orthogonal matrices . . . 54
7.3.2 Spinor representation . . . 59
7.3.3 Spinor representation space . . . 60
7.3.4 Spinor representation with matrices . . . 62
7.3.5 Higher-spin representations . . . 63

8 Lorentz group **65**

8.1 Commutators . . . 66
8.2 Fundamental representations . . . 67
8.2.1 Direct product representations . . . 72
8.3 Space inversion . . . 73

9 Poincaré group **75**

9.1 Casimir operators . . . 77
9.2 Representation space of the Poincaré group . . . 78
9.2.1 Supersymmetry and spacetime . . . 79
9.3 Physics tensors . . . 80
9.3.1 3D tensors . . . 80
9.3.2 *Transformation of electromagnetic fields . . . 80
9.3.3 *The Maxwell field tensor . . . 82

10 Classical fields **84**

10.1 The field viewpoint . . . 84
10.2 Continuous systems . . . 85
10.3 Lagrangian density . . . 86

11 Relativistic field equations **90**

11.1 *Classical Klein-Gordon equation . . . 90
11.1.1 Complex-valued fields . . . 92
11.2 Dirac equation . . . 93

11.3 Weyl equation 96

12 Electromagnetism **97**

12.1 Revision: Maxwell's equations and potentials 97

12.2 *Electromagnetic potential as a 4-vector 97

12.3 *Gauge invariance 100

12.4 Lagrangian for em fields, equations of motion 101

12.4.1 Use of invariants 102

12.4.2 Motion in an electromagnetic field 104

13 Radiation **106**

13.1 Conservation of energy 106

13.2 Plane waves in vacuum 107

14 Fields with sources **110**

14.1 *Fields of a uniformly moving charge 110

14.2 *Retarded potentials 111

14.3 Arbitrarily moving charge 115

15 Accelerated charge **118**

15.1 Slowly oscillating dipole 118

15.2 *Field of an accelerated charge (details) 120

15.3 *Half-wave electric dipole antenna 123

15.3.1 *Radiated power 124

15.3.2 Energy loss in accelerators 127

16 Energy-momentum tensor **128**

16.1 Fluid examples 129

16.2 *Energy-momentum tensor of the EM field 130

16.3 *Applications with simple geometries 131

16.3.1 *Parallel-plate capacitor 131

16.3.2 *Long straight solenoid 131

16.3.3 *Plane waves 132

17 Noether's theorem **133**

17.1 Discrete systems . . . 133
17.1.1 Action invariance . . . 134
17.1.2 On-shell variation . . . 134
17.1.3 Noether's Theorem . . . 135
17.1.4 Examples . . . 136
17.2 Noether's Theorem for classical fields . . . 140
17.2.1 Translations . . . 141
17.2.2 Complex fields . . . 142
17.2.3 Maxwell's equations . . . 144
17.3 Local gauge invariance . . . 146

Chapter 1

Introduction

These notes are in the process of construction. Comments, clarifications, and especially corrections, are welcome by the author.

1.1 Books

The main text for the course is AM Steane, *Relativity Made Relatively Easy*, Oxford University Press, 2012.

However, I've also drawn on quite a few other sources, including

- JD Jackson, *Classical Electrodynamics*, Wiley, 1975.
- H Goldstein, *Classical Mechanics*, Addison Wesley, 1980.
- WK Tung, *Group Theory in Physics*, World Scientific, 1985.
- lots of Oxford lecture notes, especially from AM Steane, CWP Palmer, S Balbus, and J Binney.

In some cases, there are more recent editions of these textbooks.

1.2 Postulates

Postulates:

1. *Principle of Relativity*
   The motions of bodies included in a given space are the same among themselves, whether that space is at rest or moves uniformly (forward) in a straight line.
   - The laws of physics take the same mathematical form in all inertial frames of reference.

2. *Light speed postulate*
   There is a finite maximum speed for signals.

   - There is an inertial frame in which the speed of light in vacuum is independent of the motion of the source.

Additional postulates (or assumptions):

1. Flat space, sometimes called "Euclidean".
2. Internal interactions in an isolated system cannot change the system's total momentum.

   - or translational symmetry, for reasons which will be developed later.

1.3 *Vector transformation

(Newtonian and relativistic)

We'll use "frame" in the sense of a coordinate system in space and time. Since we usually talk of "points" as having spatial coordinates, we'll call events with space and time coordinates "events". I will omit discussion of these concepts in terms of light signals and bouncing off mirrors, as these discussions can obscure the essential simplicity of coordinate systems.

We also draw a distinction between events and when they're observed: it may take time for you to observe an event (in other words, observe its consequences), but once you do, you may be able to deduce the event's spacetime coordinates. So, for instance, we lost contact with the Cassini mission as it plunged into Saturn's atmosphere at 4.55am PDT on 15 September, but we knew that the last signal which could reach us was actually transmitted around 3.32am PDT. By the time we received its last transmission, it had already been vaporized for almost 1.5 hours.

Standard configuration: we have two frames, S and S', with all spatial axes parallel. S' then moves with a constant speed in the x direction with respect to S.

Galilean (Newtonian) transformation between inertial frames in standard configuration:

$$t' = t \tag{1.1}$$
$$x' = x - vt \tag{1.2}$$
$$y' = y \tag{1.3}$$
$$z' = z \tag{1.4}$$

where v is the constant speed.

Lorentz transformation for standard configuration:

$$t' = \gamma\left(t - \frac{vx}{c^2}\right) \tag{1.5}$$
$$x' = \gamma(x - vt) \tag{1.6}$$
$$y' = y \tag{1.7}$$
$$z' = z \tag{1.8}$$

where $\gamma = (1 - \beta^2)^{-1/2}$.

There are a number of derivations of the Lorentz transformation, and you can find some in basic texts in Special Relativity or your CP1 notes. The physics hasn't changed.

1.4 Symmetry and relativity

The idea of symmetry: transformations which do not change the physics. This is basically the first postulate of Special Relativity: equations of motion remain unchanged by the Lorentz transformation.

Consider first one of the first equations of motion in physics, the Newtonian equation for a free particle:

$$0 = \mathbf{f} = m\frac{d^2\mathbf{x}}{dt^2} \tag{1.9}$$

If we plug in the Galilean transformation,

$$m\frac{d^2x'}{dt'^2} = m\frac{d^2}{dt^2}(x - vt) = m\frac{d}{dt}\left(\frac{dx}{dt} - v\right) = m\frac{d^2x}{dt^2} = 0 \tag{1.10}$$

so we see the equation of motion is unchanged by the transformation.

This isn't the only symmetry of Newtonian physics. For instance, consider a rotation in the xy plane:

$$\begin{aligned} t' &= t && (1.11)\\ x' &= x\cos\theta - y\sin\theta && (1.12)\\ y' &= x\sin\theta + y\cos\theta && (1.13)\\ z' &= z && (1.14)\\ & && (1.15) \end{aligned}$$

This also leaves Newton's law invariant, at least in vector form. It also does something more: it leaves the lengths of all displacements the same. So we can rotate the world (or the experiment), and all equations which involve vector displacements are left unchanged. (Note that the same can't be said of Galilean transformations.)

The Lorentz transformation, however, does change the form of the Newton equation of motion. Since we are led to believe that the Lorentz transformation is the proper transformation of space and time coordinates, then it must be that Newton's Law is the one which must be modified: it must be a low-velocity approximation of a better equation of motion.

In this course, one of the things we'll be doing is exploring the implications of the requirement that physics is unchanged by the Lorentz transformation. On the one hand, we require that all physics, ultimately, be invariant under a Lorentz transformation. On the other hand, we also want to find out *everything* which is invariant under such transformations, to see if we can then observe them in Nature. This may also apply to other possible symmetries.

In other words, and more generally, we'll look for what physics tells us about symmetry, and what symmetry tells us about physics.

1.5 Units

First, let's rephrase the Lorentz transformation by moving our c's around:

$$ct' = \gamma(ct - \beta x) \tag{1.16}$$
$$x' = \gamma(x - \beta ct) \tag{1.17}$$
$$y' = y \tag{1.18}$$
$$z' = z \tag{1.19}$$

This makes the equations look more alike, if you take ct as the time-like coordinate. And of course ct now has units of length, as the others already do.

1.5.1 Other conventions

Different textbooks will take different approaches to making these equations look symmetric, so it's worth a word of warning.

- Natural units: just assign $c = 1$. This basically says that length and time are really the same unit, which in a sense is true. And indeed this is how most particle physicists (for instance) work, even to the extent that we'll talk about the "lifetime" of a bottom quark meson to be "about 450 microns". And indeed you can also assign $\hbar = 1$. If you've followed all your units through correctly, you can just reintroduce however many c's and $\hbar$'s as you need to the final result to get the units you need, and it works perfectly. But it can be confusing to students, so for the lectures I'll try to avoid it. However, if I lapse into it (because it's how I normally work), well, apologies in advance.
- Many older textbooks also make the time component imaginary. This makes the coordinates *really* symmetric: the Lorentz boost now really looks like a Euclidean rotation. Most modern textbooks consider this a step too far, because it still remains that time is special—whatever units you use for it. We'll use another way to pick out the special behavior of time. Importantly, it's the way which will extend naturally to general relativity.

1.6 4-vector basics

Now let's return to the Lorentz transformation itself. The form looks remarkably like a rotation: the component being transformed is always multipled by the same factor γ, while the component being mixed in multiplied by another factor $-\beta\gamma$. To see this more clearly, let's reformulate the transformations in terms of vectors and matrices. I believe you will have run into this in CP1, even though it only gets on the syllabus for this paper.

We construct a 4-dimensional vector using the following convention:

$$\mathsf{X} = (ct, \mathbf{x}) \tag{1.20}$$

where you can see that we tend to list the time component first (the "zeroth" component—this makes it more natural for programming in most modern programming languages, by the way). Since you are probably also used to thinking of vectors as single-width column matrices,

$$\mathbf{X} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{1.21}$$

A spatial rotation then looks as follows:

$$\mathbf{X}' = \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \mathbf{RX} \tag{1.22}$$

and a Lorentz transformation

$$\mathbf{X}' = \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \mathbf{LX} \tag{1.23}$$

which can be written in a more suggestive form:

$$\mathbf{X}' = \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \mathbf{LX} \tag{1.24}$$

You can confirm that the trigonometric identities for $\cosh\eta$ and $\sinh\eta$ reproduce that of γ and $\beta\gamma$. Now this looks a lot more like a rotation, but instead of mixing two space coordinates, it mixes time and a space coordinate. In fact, it's like a rotation through an imaginary angle, which reflects that time still has a special place different from space.

The parameter η is called "rapidity", and it has another nice property, which is that addition of parallel velocities is easy. There are two ways to see this. First, you can multiply out two Lorentz transformations for two rapidities η and η'. (I'll do this with 2×2 matrices, since it leaves the other components unchanged.)

$$\begin{pmatrix} \cosh\eta' & -\sinh\eta' \\ -\sinh\eta' & \cosh\eta' \end{pmatrix} \begin{pmatrix} \cosh\eta & -\sinh\eta \\ -\sinh\eta & \cosh\eta \end{pmatrix} \tag{1.25}$$

$$= \begin{pmatrix} \cosh\eta'\cosh\eta + \sinh\eta'\sinh\eta & -\cosh\eta'\sinh\eta - \sinh\eta'\cosh\eta \\ -\cosh\eta'\sinh\eta - \sinh\eta'\cosh\eta & \cosh\eta'\cosh\eta + \sinh\eta'\sinh\eta \end{pmatrix} \tag{1.26}$$

$$= \begin{pmatrix} \cosh(\eta' + \eta) & -\sinh(\eta' + \eta) \\ -\sinh(\eta' + \eta) & \cosh(\eta' + \eta) \end{pmatrix} \tag{1.27}$$

The other way is simpler: notice that $\beta = \tanh\eta$. The addition formula is then just the hyperbolic tangent addition formula:

$$\tanh\eta'' = \tanh(\eta + \eta') = \frac{\tanh\eta + \tanh\eta'}{1 + \tanh\eta\tanh\eta'} \tag{1.28}$$

1.7 *Invariants

The dot product between the two 4-vectors is defined with a slight wrinkle which handles the special behavior of time:

$$\mathsf{A} \cdot \mathsf{B} = \mathbf{A}^T \mathbf{g} \mathbf{B} \tag{1.29}$$

where

$$\mathbf{g} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.30}$$

so the norm or "length" of a 4-vector is

$$\mathsf{X} \cdot \mathsf{X} = -(ct)^2 + \mathbf{x} \cdot \mathbf{x} \tag{1.31}$$

which should be familiar to you as the invariant separation between space-time events. The matrix $\mathbf{g}$ is called the "metric".

But to see whether this is true, let's do an explicit calculation of the dot product, just concentrating on the 2×2 again. Remember, though, we need to apply the Lorentz transform Λ to both vectors, since we want to put the whole system into another frame, not just one part.

$$(\Lambda \mathsf{A}) \cdot (\Lambda \mathsf{B}) = (\mathbf{A}^T \Lambda^T) \mathbf{g} (\Lambda \mathbf{B}) = \mathbf{A}^T (\Lambda^T \mathbf{g} \Lambda) \mathbf{B} \tag{1.32}$$

The parenthesis is then

$$\begin{aligned} \Lambda^T g \Lambda &= \begin{pmatrix} \cosh\eta & -\sinh\eta \\ -\sinh\eta & \cosh\eta \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cosh\eta & -\sinh\eta \\ -\sinh\eta & \cosh\eta \end{pmatrix} && (1.33) \\ &= \begin{pmatrix} -\cosh\eta & -\sinh\eta \\ \sinh\eta & \cosh\eta \end{pmatrix} \begin{pmatrix} \cosh\eta & -\sinh\eta \\ -\sinh\eta & \cosh\eta \end{pmatrix} && (1.34) \\ &= \begin{pmatrix} -\cosh^2\eta + \sinh^2\eta & 0 \\ 0 & -\sinh^2\eta + \cosh^2\eta \end{pmatrix} && (1.35) \\ &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} && (1.36) \end{aligned}$$

which is simply $\mathbf{g}$ again, so we come to

$$(\Lambda \mathsf{A}) \cdot (\Lambda \mathsf{B}) = (\mathbf{A}^T \Lambda^T) \mathbf{g} (\Lambda \mathbf{B}) = \mathbf{A}^T \mathbf{g} \mathbf{B} \tag{1.37}$$

Thus the dot product is invariant. A corollary is that the norm of the interval is invariant, as we expected.

The example we've worked with is fairly specific to the standard configuration, but it's straightforward to generalize once we have it in matrix form. In fact, we can define the Lorentz transformation generically in terms of the metric, in that it's any matrix Λ which satisfies the equation

$$\Lambda^T \mathbf{g} \Lambda = \mathbf{g} \tag{1.38}$$

It then becomes a pre-requisite that any equation which purports to be consistent with Special Relativity ("covariant", but perhaps more accurately "form-invariant" or simply "invariant") can be written in terms of quantities that transform using such a Λ.

1.8 Spacetime diagrams

A spacetime diagram shows time as another axis. Since four dimensions can be hard to visualize, we often make illustrations with just x and t.

First, let's look at the features related to the interval $s^2 = -c^2\Delta t^2 + \Delta x^2$, relative to the origin:

- the null intervals: $s^2 = 0$. This is how light travels.
- time-like intervals: $s^2 < 0$, so the time difference is larger than the space distance. These intervals can be causally connected, in the sense that there's enough time for the past end to affect the future end.
- space-like intervals: $s^2 > 0$. These intervals cannot be causally connected; there's not enough time for a signal to reach one from the other.

Lines of simultaneity are parallel to the x axis (think of measuring the length of rod: it's a difference on the x axis, where the two measuring events have to be at the same t in the frame). Their intersection with the t axis determines their ordering in time.

A worldline is a trajectory of a physical body. Each segment on the worldline is causally connected with the previous segment.

A Lorentz transformation ("boost") pushes the x and t axes closer to the null line. This shouldn't be surprising, since the ultimate boost to light speed should put you on the null line. We can see that the time-like, space-like, and null categories are invariants, *i.e.*, a time-like interval stays time-like in any frame.

1.8.1 *Proper time

Lines of simultaneity remain parallel to the transformed x' axis. This means that the order of causally connected events is also an invariant.

This enables us to parameterize the worldline in terms of some monotonically increasing function. The most convenient is the proper time, which is the time experienced by the body in its own rest frame.

Another way of looking at it is that proper time measures a body's path along its own worldline, since a body at rest has $\Delta\tau^2 = -s^2 = \Delta t^2$. The negative sign is an artefact of the metric we've chosen.

The relationship between proper time and that of any frame can be found by considering the Lorentz transformation, starting from the rest frame:

$$ct = \gamma(c\tau - \beta x) \tag{1.39}$$

resulting in (choosing $x = 0$ for the body in its own rest frame)

$$\frac{dt}{d\tau} = \gamma \tag{1.40}$$

This is actually the familiar time dilation. When written as

$$dt = \gamma d\tau \tag{1.41}$$

we see, for instance, why we see cosmic ray muons live (and travel) much longer than the one microsecond they should live in their own rest frame.

1.9 Basic 4-vectors

Let's look at a few basic 4-vectors.

1.9.1 Position

Position and displacement:

$$\mathsf{X} = (ct, \mathbf{x}) \tag{1.42}$$

The invariant is the usual invariant interval (possibly with a minus sign), which is also the body's proper time

$$\mathsf{X} \cdot \mathsf{X} = -c^2t^2 + \mathbf{x} \cdot \mathbf{x} = -c^2\tau^2 \tag{1.43}$$

The easiest way to see this is to evaluate the dot product in the body's rest frame.

1.9.2 Velocity

Velocity:

$$\mathsf{U} = \frac{d\mathsf{X}}{d\tau} = \left(c\frac{dt}{d\tau}, \frac{d\mathbf{x}}{dt}\frac{dt}{d\tau} \right) = \gamma(c, \mathbf{u}) \tag{1.44}$$

The invariant is the same for all 4-velocities:

$$\begin{aligned} \mathsf{U} \cdot \mathsf{U} &= \gamma^2(-c^2 + \mathbf{u} \cdot \mathbf{u}) && (1.45)\\ &= -c^2\gamma^2(1 - \beta^2) && (1.46)\\ &= -c^2 && (1.47) \end{aligned}$$

One should note, of course, that adding two 4-velocities doesn't give you another 4-velocity. The first clue is that the invariant of the sum is no longer $-c^2$. On the other hand, there is a useful formula for velocity addition which comes from considering the inner product of two 4-velocities:

$$\mathsf{U} \cdot \mathsf{V} = \gamma_u\gamma_v(-c^2 + \mathbf{u} \cdot \mathbf{v}) \tag{1.48}$$

This is true in any given frame. In the rest frame of one of the bodies, however, the 4-velocity is simply $(c, 0)$, in which case the invariant is simply $-\gamma_w c^2$, where γ_w is related to the speed of the other body. So, if a body is moving with a Lorentz factor of γ_u in a frame which is moving with Lorentz factor of γ_v in the "lab" frame, then its Lorentz factor in the lab frame is

$$\gamma_w = \gamma_u\gamma_v\left(1 - \frac{\mathbf{u} \cdot \mathbf{v}}{c^2}\right) \tag{1.49}$$

1.9.3 Acceleration

Acceleration:

$$\begin{aligned} \mathsf{A} &= \frac{d\mathsf{U}}{d\tau} = \gamma \frac{d\mathsf{U}}{dt} && (1.50) \\ &= \gamma(\dot{\gamma}c, \dot{\gamma}\mathbf{u} + \gamma\mathbf{a}) && (1.51) \end{aligned}$$

where the dot signifies a full derivative with respect to t (not τ, the proper time).

The invariant is found by evaluating the dot product in the body's rest frame. This is easier to see using the definition in terms of proper time:

$$\begin{aligned} \mathsf{A} &= \left(\frac{dc}{d\tau}, \frac{d\mathbf{u}}{d\tau}\right) && (1.52) \\ &= (0, \mathbf{a}_0) && (1.53) \end{aligned}$$

where $\mathbf{a}_0$ is the "proper acceleration", *i.e.*, the acceleration as experienced by the body in its (instantaneous) rest frame. So the invariant is

$$\mathsf{A} \cdot \mathsf{A} = a_0^2 \tag{1.54}$$

1.9.4 Momentum

Momentum:

$$\mathsf{P} = m\mathsf{U} = (\gamma mc, \gamma m\mathbf{u}) = (E/c, \mathbf{p}) \tag{1.55}$$

You will have seen the transformation of energy and momentum from CP1.

$$\begin{aligned} E'/c &= \gamma(E/c - \beta p_x) && (1.56) \\ p'_x &= \gamma(p_x - \beta E/c) && (1.57) \\ p'_y &= p_y && (1.58) \\ p'_z &= p_z && (1.59) \end{aligned}$$

In this case, the norm is

$$\mathsf{P} \cdot \mathsf{P} = -(E/c)^2 + \mathbf{p} \cdot \mathbf{p} = -(mc)^2 \tag{1.60}$$

which is proportional to the square of the invariant (rest) mass of the system.

These relationships encapsulate a lot of the formulas you picked up in CP1 relating energy, momentum, mass, and the factors γ and β. Indeed, when you tried working out collision problems without 4-vectors, you probably had the experience of throwing a lot of these formulas at the problem and coming out with really enlightening equations which amounted to $1 = 1$. This is because all those equations were really just different aspects of the relationships given here.

1.9.5 Force

Force:

$$\mathsf{F} = \frac{d\mathsf{P}}{d\tau} = \left(\frac{1}{c}\frac{dE}{d\tau}, \frac{d\mathbf{p}}{d\tau}\right) = \left(\frac{\gamma}{c}\frac{dE}{dt}, \gamma\mathbf{f}\right) \tag{1.61}$$

where

$$\mathbf{f} = \frac{d\mathbf{p}}{dt} \tag{1.62}$$

is the familiar 3-force.

We'll look at this more closely later.

Chapter 2

Applications of 4-momentum

Now, what can we do with 4-vectors?

This chapter should seem a bit like revision, since you'll have covered conservation of momentum and collisions back in CP1. But we'll do them with 4-momentum in order to gain familiarity and demonstrate some of usual techniques and questions.

2.1 *Conservation of energy-momentum

Energy and 3-momentum form a 4-vector:

$$\mathsf{P} = (E/c, \mathbf{p}) \tag{2.1}$$

In elastic collisions, the 4-vector is a conserved quantity.

By this I mean that each component is conserved separately, but one should start thinking of 4-vectors themselves as quantities. (As we'll see later on, the 4-vectors are some of the building blocks of theories which are valid from the point of view of Special Relativity. Numbers are only valid building blocks if they're scalars, not components or parts of other objects like vectors.)

Fundamentally, collisions are always elastic; they are inelastic when we choose to ignore some forms of energy in the final state.

2.2 *Annihilation, decay, and formation

Here we'll consider some examples of using 4-vectors to look at particle interactions.

2.2.1 *Center of momentum frame

A typical pattern in these problems is that you'll select an invariant, and then attempt to evaluate it in some frame. One of the most convenient, of course, is the center of momentum

frame, in which the total 3-momentum of the system is zero. It is also this frame which experiences the passage of proper time.

2.2.2 Decay at rest

Consider the decay of a single particle with 4-momentum P to two particles which 4-momenta P_1 and P_2. The parent particle has mass M, and the daughters m_1 and m_2.

4-momentum conservation gives us

$$\mathsf{P} = \mathsf{P}_1 + \mathsf{P}_2 \tag{2.2}$$

First, I'll do this using a method I don't usually recommend, but is good enough in this case: choose a convenient frame, then write out the components. It's good enough here because in the rest frame,

$$\mathsf{P} = (Mc, 0) \tag{2.3}$$

which means that the other two 4-momenta must be

$$\begin{aligned} \mathsf{P}_1 &= (E_1/c, \mathbf{p}) && (2.4)\\ \mathsf{P}_2 &= (E_2/c, -\mathbf{p}) && (2.5) \end{aligned}$$

with

$$\begin{aligned} E_1^2/c^2 &= m_1^2c^2 + p^2 && (2.6)\\ E_2^2/c^2 &= m_2^2c^2 + p^2 && (2.7) \end{aligned}$$

where $p = |\mathbf{p}|$. So just considering the 0th (energy) component, we can solve for E_1

$$\begin{aligned} Mc &= E_1/c + E_2/c && (2.8)\\ Mc - E_1/c &= E_2/c && (2.9)\\ M^2c^2 + m_1^2c^2 + p^2 - 2ME_1 &= m_2^2c^2 + p^2 && (2.10)\\ E_1 &= \frac{M^2c^4 + m_1^2c^4 - m_2^2c^4}{2Mc^2} && (2.11) \end{aligned}$$

where I've reintroduced a common factor of c^2 in order to put all the terms in units of energy, since most elementary particle masses are quoted in units of energy such as MeV. You can also see why it's a lot less tedious just to drop the c's by setting $c = 1$.

In this connection it's worth looking at the difference between an atom and its constituents. A hydrogen atom at rest has 4-momentum $(Mc, 0)$. But it also consists of a proton (mass m_p) and an electron (mass m_e). If the two particles are infinitely far from one another, then the 4-vectors in the rest frame are simply

$$(m_pc, 0) + (m_ec, 0) = ((m_p + m_e)c, 0) \tag{2.12}$$

In other words, the mass of the total separated system is $m_p + m_e$.

To get from one system to the other, one must do some work to move the electron amd proton far apart. The amount of work amounts to $R_\infty = 13.6$ eV. So even though it's not

a huge effect—the mass of a proton is about 938 MeV and the electron is 0.511 MeV—it can still be said that the hydrogen atom has 13.6 eV less mass than a proton and electron together. The binding energy is a mass deficit. This becomes a lot more noticeable in nuclear physics, when the binding energies are on the order of MeV.

2.2.3 In-flight decay

The more general circumstance is a decay in flight. Here we go back to the 4-momentum conservation again, but in some generic lab frame where the parent particle is not at rest.

$$\mathsf{P} = \mathsf{P}_1 + \mathsf{P}_2 \tag{2.13}$$

This is true in any frame. But remember that the 4-momenta are part of a linear space, so we can perform most of the usual "arithmetic" on them. In this case we isolate P_2 on one side:

$$\mathsf{P} - \mathsf{P}_1 = \mathsf{P}_2 \tag{2.14}$$

and then we square

$$(\mathsf{P} - \mathsf{P}_1) \cdot (\mathsf{P} - \mathsf{P}_1) = \mathsf{P}_2 \cdot \mathsf{P}_2 \tag{2.15}$$

The right hand side is just the invariant of the 4-momentum, which is simply $-m_2^2c^2$. This is true in any frame.

The left hand side, in the meantime, becomes

$$\begin{aligned} (\mathsf{P} - \mathsf{P}_1) \cdot (\mathsf{P} - \mathsf{P}_1) &= \mathsf{P} \cdot P + \mathsf{P}_1 \cdot \mathsf{P}_1 - 2\mathsf{P} \cdot \mathsf{P}_1 & (2.16)\\ &= -M^2c^2 - m_1^2c^2 - 2\mathsf{P} \cdot \mathsf{P}_1 & (2.17) \end{aligned}$$

so combined we have

$$\mathsf{P} \cdot \mathsf{P}_1 = -\frac{1}{2}(M^2c^2 + m_1^2c^2 - m_2^2c^2) \tag{2.18}$$

This equation is still true in any frame: it involves invariants.

We can recover the decay at rest formula by choosing the rest frame of the parent particle and plugging in:

$$\mathsf{P} \cdot \mathsf{P}_1 = -ME_1 \tag{2.19}$$

because the parent's 3-momentum is zero in the rest frame, and the formula comes out directly.

The other usual circumstance is that we actually don't know the mass of the parent particle. The usual excuse is that we haven't discovered the particle yet. Instead, we observe the momenta of the daughter particles, which presumably have been discovered before. Let's say the parent decays into a number of daughters:

$$\mathsf{P} = \sum_i \mathsf{P}_i \tag{2.20}$$

If we square both sides, we get

$$M^2c^2 = -\left(\sum_i \mathsf{P}_i\right)^2 \tag{2.21}$$

The left side is an invariant, while the right side can be evaluated by observations in the lab frame.

(One could note here that the new particle would show up as a resonance peak, the shape of which may be familiar from time-dependent perturbation theory.)

A further consequence of 4-momentum conservation is that it doesn't matter if the decay happens all at once, or through several stages with intermediate particles. This is because the 4-momentum of each intermediate particle is simply the sum of their daughters, so in the end you still end up with the sum of observed 4-momenta. However, the intermediate states do leave their trace: some of the 4-momenta will sum up such that their invariant masses are close to that of the intermediate particle (this is quantum mechanics, after all). This can be used to reduce backgrounds in searching for new particles, if we are expecting those intermediate states.

2.3 *Collisions

Now let's look at particles hitting one another.

2.3.1 Absorption

The first type of experiment, which was the easier one to set up, was to send a beam of particles into a target, and then pick up what came out the other side. We'll start, however, with just exciting the target.

The 4-momentum conservation equation is

$$\mathsf{P}_i + \mathsf{P} = \mathsf{P}_f \tag{2.22}$$

Since there is non-zero linear momentum in this setup, we can't have the final particle with mass M_f at rest. Some of the initial energy therefore has to go into keeping it moving along. The question is then, how much initial energy is needed to start from a target with M and get a final particle of mass M_f?

We pretty much follow the same logic as for the decay in flight: we start from a 4-momentum conservation equation and square both sides. In this case, we get

$$(\mathsf{P}_i + \mathsf{P})^2 = \mathsf{P}_f^2 = -M_f^2 c^2 \tag{2.23}$$

while the left side can also be written as

$$(\mathsf{P}_i + \mathsf{P})^2 = \mathsf{P}_i^2 + \mathsf{P}^2 + 2\mathsf{P}_i \cdot \mathsf{P} = -m_i^2 c^2 - M^2 c^2 - 2(E_i M - \mathbf{p}_i \cdot \mathbf{p}) \tag{2.24}$$

so solving for the initial beam energy E_i,

$$E_i = \frac{M_f^2 c^4 - M^2 c^4 - m_i^2 c^4}{2Mc^2} \tag{2.25}$$

Now let's compare this with our decay at rest formula. Think of the question of whether an atom in an excited state can decay and emit a photon which excites a neighboring identical atom. Let's say the excited atom has mass M^*, and the ground state M. The photon (γ) has zero mass. On the emission side, we use the decay at rest formula:

$$E_\gamma = \frac{M^{*2}c^4 + m_\gamma^2 c^4 - M^2c^4}{2M^*c^2} = \frac{M^{*2}c^4 - M^2c^4}{2M^*c^2} \tag{2.26}$$

To excite the neighboring atom, which is initially at rest, we need a photon with energy

$$E_i = \frac{M^{*2}c^4 - M^2c^4 - m_\gamma^2 c^4}{2Mc^2} = \frac{M^{*2}c^4 - M^2c^4}{2Mc^2} \tag{2.27}$$

Since $M < M^*$, we see that $E_i > E_\gamma$. The emitted photon has lost a little bit of energy to the recoil of the initial atom, and it needs even a bit more energy to enable the recoil of the target atom. So in general one excited atom can't transmit its excitation to another.

One should note, however, that there is another way, which is called the "Mössbauer effect", or sometimes "recoil-less" emission and absorption. It isn't found among isolated particles. The reason is that the recoil is taken up by the environment, such as a crystal lattice in which the atom is embedded. Since the macroscopic material is much more massive than an individual atom, the actual recoil experienced by the individual atom in that case is technically non-zero, but negligible.

2.3.2 Particle formation

In the more general case, the final state consists of a number of known particles which we try to observe. Since this final state is specified, and presumably consists of particles with known properties, we can ask the question, how much energy will it take to create this final state?

We know from the absorption case that some energy will be "lost" to recoil. At the same time, we know that there's one frame in which it's easy to identify the lowest-energy configuration of the final state: it's the frame in which all the produced particles are at rest. This is because, in a final state in some frame, the total energy is

$$E = \sum_i (m_i^2 c^4 + p^2 c^2)^{1/2} \tag{2.28}$$

so we can see that if any particle moves even just a little bit, it only adds to the total energy of the system. Therefore, the lowest energy configuration is when they're all the rest. The invariant is then easy to calculate, since all the particles are at rest. It's just the sum of the masses of the final state particles:

$$\left(\sum_j \mathsf{P}_j\right)^2 = -\left(\sum_j m_j\right)^2 \tag{2.29}$$

The 4-momentum equation is

$$\begin{aligned} \mathsf{P}_{in} + \mathsf{P} &= \sum_j \mathsf{P}_j && (2.30)\\ (\mathsf{P}_{in} + \mathsf{P})^2 &= \left(\sum_j \mathsf{P}_j\right)^2 && (2.31)\\ M^2c^2 + m^2c^2 + 2(E_{in}M) &= \left(\sum_j m_j c\right)^2 && (2.32)\\ E_{in} &= \frac{1}{2Mc^2}\left[\left(\sum_j m_j c^2\right)^2 - M^2c^4 - m^2c^4\right] && (2.33) \end{aligned}$$

Note that the required energy increases as the intended mass squared. So for a Higgs with mass 125.09±0.24 GeV (PDG 2017) to be created from the collision of a proton (938.27 MeV) hitting another proton—and this is assuming that the protons are entirely consumed, which actually can't happen because of other symmetries in particle physics—you would need to have a proton beam with energy 8.3375 TeV. (Even at design energy, the LHC beams are 7 TeV.)

On the other hand, if you can accelerate both protons, you get a very different relationship.

$$\begin{aligned} \mathsf{P}_a + \mathsf{P}_b &= \sum_j \mathsf{P}_j && (2.34)\\ (\mathsf{P}_a + \mathsf{P}_b)^2 &= \left(\sum_j m_j c\right)^2 && (2.35) \end{aligned}$$

In this case, the 4-momenta of the initial states in the lab frame are

$$\begin{aligned} \mathsf{P}_a &= (E, \mathbf{p}) && (2.36)\\ \mathsf{P}_b &= (E, -\mathbf{p}) && (2.37) \end{aligned}$$

and thus

$$\begin{aligned} 4E^2 &= \left(\sum_j m_j c^2\right)^2 && (2.38)\\ E &= \frac{1}{2}\left(\sum_j m_j c^2\right) && (2.39) \end{aligned}$$

This is a lot less energy. Each beam just needs half the mass of the Higgs, so it's about 62.5 GeV.

By comparison, the LHC beam energy was 4 TeV when the Higgs was discovered, and is now about 7 TeV. It's still less than would have been needed in a fixed target experiment, though of course it's a lot harder to guide two high-energy beams into eachother. You can ask

your B4 lecturer or your tutors why all the additional energy is needed. Also why it wasn't done with an electron-positron collider, in spite of the fact that electrons and positrons *are* fundamental particles and thus would have been completely consumed in the collision.

But that's really another course.

2.3.3 *Compton scattering

If we specialize back to a $2 \to 2$ process, and just have the initial particles "bounce" off one another, we have

$$\begin{aligned} \mathsf{P}_1' + \mathsf{P}_2' &= \mathsf{P}_1 + \mathsf{P}_2 && (2.40)\\ \mathsf{P}_2' &= \mathsf{P}_1 + \mathsf{P}_2 - \mathsf{P}_1' && (2.41)\\ (\mathsf{P}_2')^2 &= (\mathsf{P}_1 + \mathsf{P}_2 - \mathsf{P}_1')^2 && (2.42)\\ -m_2^2c^4 &= -m_1^2c^4 - m_2^2c^4 - m_1^2c^4 + 2\mathsf{P}_1\cdot\mathsf{P}_2 - 2\mathsf{P}_1\cdot\mathsf{P}_1' - 2\mathsf{P}_2\cdot\mathsf{P}_1' && (2.43) \end{aligned}$$

where we've used the "isolate and square" trick again to ignore the second (target) mass's final trajectory. We can always solve for it later if we want to.

We now choose a convenient frame. The lab frame, with a stationary target, has a nice zero for the initial momentum. This means

$$\begin{aligned} 2\mathsf{P}_1\cdot\mathsf{P}_2 &= 2E_1m_2 && (2.44)\\ 2\mathsf{P}_2\cdot\mathsf{P}_1' &= 2E_1'm_2 && (2.45)\\ 2\mathsf{P}_1\cdot\mathsf{P}_1' &= 2(\mathbf{p}_1\cdot\mathbf{p}_1' - E_1E_1') && (2.46)\\ &= 2(p_1p_1'\cos\theta - E_1E_1') && (2.47) \end{aligned}$$

where θ is the angle between the initial and final trajectories of the incoming particle. Combining it all, we get

$$0 = m_1^2c^4 + (E_1 - E_1')m_2c^2 - (E_1E_1' - c^2p_1p_1'\cos\theta) \tag{2.48}$$

A special case is Compton scattering, in which the incoming particle is a photon, and the target is an atomic electron, which is considered to be more or less at rest. Since the photon has zero mass,

$$\begin{aligned} m_1c^2 &= 0 && (2.49)\\ m_2c^2 &= m_ec^2 && (2.50)\\ E_1 &= |\mathbf{p}_1|c = hc/\lambda && (2.51)\\ E_1' &= |\mathbf{p}_1'|c = hc/\lambda' && (2.52) \end{aligned}$$

where we've also used the relationship between the photon energy and its wavelength. We plug these things in and get

$$\begin{aligned} 0 &= (E_1 - E_1')m_ec^2 - E_1E_1'(1-\cos\theta) && (2.53)\\ \frac{E_1 - E_1'}{E_1E_1'} &= \frac{1-\cos\theta}{m_ec^2} && (2.54) \end{aligned}$$

But since

$$\frac{E_1 - E_1'}{E_1 E_1'} = \frac{\lambda\lambda'}{hc}\left(\frac{1}{\lambda} - \frac{1}{\lambda'}\right) = \frac{\lambda\lambda'}{hc}\left(\frac{\lambda' - \lambda}{\lambda\lambda'}\right) = \frac{\lambda' - \lambda}{hc} \tag{2.55}$$

we have the usual formula

$$\lambda' - \lambda = \frac{h}{m_e c}(1 - \cos\theta) \tag{2.56}$$

Chapter 3

Force

A reminder of the 4-force:

$$\mathsf{F} = \frac{d\mathsf{P}}{d\tau} = \left(\frac{1}{c}\frac{dE}{d\tau}, \frac{d\mathbf{p}}{d\tau}\right) = \left(\frac{\gamma}{c}\frac{dE}{dt}, \gamma\mathbf{f}\right) \tag{3.1}$$

where

$$\mathbf{f} = \frac{d\mathbf{p}}{dt} \tag{3.2}$$

is the familiar 3-force.

3.1 Pure force

Consider a particle with 4-vector velocity $\mathsf{U} = \gamma(c, \mathbf{u})$ subject to 4-force F. We form the scalar product

$$\mathsf{U} \cdot \mathsf{F} = \gamma^2 \left(\mathbf{u} \cdot \mathbf{f} - \frac{dE}{dt}\right) \tag{3.3}$$

which is invariant. Since it is, we can calculate the value in a convenient frame, which in this case is the rest frame of the particle. In this case, $\mathbf{u} = 0$, $\gamma = 1$, $dt = d\tau$, $\mathbf{p} = 0$ and $E = mc^2$, so

$$\mathsf{U} \cdot \mathsf{F} = -c^2 \frac{dm}{dt} \tag{3.4}$$

where m is the rest mass.

If the force doesn't change the rest mass, then all the work is done changing kinetic energies, and we have

$$0 = \mathsf{U} \cdot \mathsf{F} = \gamma^2 \left(\mathbf{u} \cdot \mathbf{f} - \frac{dE}{dt}\right) \tag{3.5}$$

$$\frac{dE}{dt} = \mathbf{u} \cdot \mathbf{f} \tag{3.6}$$

which is the usual classical result.

3.2 Transformation of force

(Steane 4)

Again, consider a particle with 4-vector velocity $\mathsf{U} = \gamma(c, \mathbf{u})$ in frame S, and subject to 4-force F. Let S' be a frame moving with velocity $\mathbf{v}$ with respect to S.

We apply the Lorentz transformation on the force 4-vector, for which we split the spatial part into $f_\parallel$ parallel to $\mathbf{v}$, and $\mathbf{f}_\perp$ perpendicular to it:

$$\begin{aligned}
\mathsf{F}'^0 &= \gamma_v(\mathsf{F}^0 - \beta_v \mathsf{F}_\parallel) && (3.7)\\
&= \gamma_v \left(\frac{\gamma_u}{c}\frac{dE}{dt} - \frac{\mathbf{v}}{c}\gamma_u f_\parallel \right) && (3.8)\\
\frac{\gamma'_u}{c}\frac{dE'}{dt'} &= \gamma_v\gamma_u \left(\frac{1}{c}\frac{dE}{dt} - \beta_v f_\parallel \right) && (3.9)\\
\gamma'_u f'_\parallel &= \gamma_v \left(\gamma_u f_\parallel - \beta_v \frac{\gamma_u}{c}\frac{dE}{dt} \right) && (3.10)\\
&= \gamma_v\gamma_u \left(f_\parallel - \frac{\beta_v}{c}\frac{dE}{dt} \right) && (3.11)\\
\gamma'_u \mathbf{f}'_\perp &= \gamma_u \mathbf{f}_\perp && (3.12)
\end{aligned}$$

To change the left sides into more convenient expressions, we use the following expression from the addition of velocities:

$$\gamma_w = \gamma_u\gamma_v(1 - \mathbf{u}\cdot\mathbf{v}/c^2) \tag{3.13}$$

which gives us the transformed forces themselves:

$$\begin{aligned}
\mathbf{f}'_\perp &= \frac{\mathbf{f}_\perp}{\gamma_v(1 - \mathbf{u}\cdot\mathbf{v}/c^2)} && (3.14)\\
f'_\parallel &= \frac{f_\parallel - \left(\frac{\mathbf{v}}{c^2}\frac{dE}{dt}\right)}{1 - \mathbf{u}\cdot\mathbf{v}/c^2} && (3.15)
\end{aligned}$$

The last equation, in the special case of a pure force, simplifies to

$$f'_\parallel = \frac{f_\parallel - \mathbf{v}(\mathbf{f}\cdot\mathbf{u})/c^2}{1 - \mathbf{u}\cdot\mathbf{v}/c^2} \tag{3.16}$$

We make the following observations:

- $\mathbf{f}$ is not invariant between frames.
- $\mathbf{f}$ which is independent of its subject's velocity in one frame is actually dependent on it in another.

We can also see that for $\mathbf{u} = 0$, $\mathbf{f}$ is the force acting in the rest frame. In another frame, however, the transverse force is

$$\mathbf{f}'_\perp = \mathbf{f}_\perp/\gamma_v \tag{3.17}$$

which is reduced. This means that there are internal tensions, and so, for instance, the breaking strength of extended objects is smaller when they move (cf. Trenton-Noble experiment).

3.3 *Force and simple motion problems

3.3.1 Motion under pure force

(Steane 4.2)

Let's investigate the motion of a particle under a given force. We still have

$$\mathbf{f} = \frac{d\mathbf{p}}{dt} \tag{3.18}$$

but $\mathbf{p}$ now has to be the relativistic version,

$$\mathbf{p} = \gamma_u m\mathbf{u} \tag{3.19}$$

For a pure force, we have

$$\begin{aligned}
\frac{dm}{dt} &= 0 && (3.20)\\
\mathbf{f} &= \frac{d}{dt}(\gamma_u m\mathbf{u}) && (3.21)\\
&= \gamma_u m\mathbf{a} + m\frac{d\gamma_u}{dt}\mathbf{u} && (3.22)\\
&= \gamma_u m\mathbf{a} + m\left(\frac{1}{mc^2}\mathbf{f}\cdot\mathbf{u}\right)\mathbf{u} && (3.23)\\
&= \gamma_u m\mathbf{a} + \frac{1}{c^2}(\mathbf{f}\cdot\mathbf{u})\mathbf{u} && (3.24)
\end{aligned}$$

where $\mathbf{u}$ is the velocity of the particle. The first term is as we'd expect. The second term is not so intuitive, since it means the change in the velocity isn't necessarily in the direction of the force. In fact, it's only the case in two special cases:

1. if the speed doesn't change ($d\gamma/dt = 0$), such as we might see in circular motion; and
2. if the force is along the direction of motion $\hat{\mathbf{u}}$.

Since we often apply $\mathbf{f}$ to a particle with a known $\mathbf{u}$, it's convenient to resolve the motion

into components parallel and perpendicular to $\mathbf{u}$.

$$\begin{aligned}
f_\parallel &= \gamma m a_\parallel + \frac{1}{c^2} f_\parallel u^2 && (3.25)\\
f_\parallel \left(1 - \frac{u^2}{c^2}\right) &= \gamma m a_\parallel && (3.26)\\
f_\parallel &= \gamma^3 m a_\parallel && (3.27)\\
\mathbf{f}_\perp &= \mathbf{f} - f_\parallel \hat{\mathbf{u}} && (3.28)\\
&= \gamma m \mathbf{a} + \frac{1}{c^2}(f_\parallel u)\mathbf{u} - f_\parallel \hat{\mathbf{u}} && (3.29)\\
&= \gamma m \mathbf{a} + \frac{u^2}{c^2} f_\parallel \hat{\mathbf{u}} - f_\parallel \hat{\mathbf{u}} && (3.30)\\
&= \gamma m \mathbf{a} - \frac{f_\parallel}{\gamma^2}\hat{\mathbf{u}} && (3.31)\\
&= \gamma m (\mathbf{a} - a_\parallel \hat{\mathbf{u}}) && (3.32)\\
&= \gamma m \mathbf{a}_\perp && (3.33)
\end{aligned}$$

Note that since $\gamma \geq 1$, we need more $f_\parallel$ to increase $a_\parallel$ than in the perpendicular case. So there is greater resistance to inertial changes in $\mathbf{u}$ direction than transverse to it.

3.3.2 Linear motion

We examine the motion of a particle under some acceleration as observed by the particle itself. This is the case where the force is parallel to the motion of the particle.

For this, we have to think of a sequence of "instantaneous rest frames" $\{A\}$ which happen to have the same velocity as the particle at a given time t in the laboratory frame. We need to label the frames A by some function which increases monotonically in t; we take the particle's proper time τ as this parameter.

In each frame, the particle is initially at rest, but then picks up velocity $dv = a_0 d\tau$.

Now, we have two frames we want to relate: the instantaneous rest frame, and the laboratory frame. So let's choose an invariant (of course!).

$$\mathsf{A} \cdot \mathsf{A} = a_0^2 \tag{3.34}$$

is a pretty convenient relationship between the acceleration in some instantaneous rest frame, and whatever other frame we choose to use. Note that a_0 could be a function of a parameter such as proper time.

Let's evaluate $\mathsf{A} \cdot \mathsf{A}$ in the laboratory frame:

$$\begin{aligned}
\mathsf{A} &= \gamma(\dot{\gamma}c, \dot{\gamma}\mathbf{u} + \gamma\mathbf{a}) && (3.35)\\
\mathsf{A} \cdot \mathsf{A} &= \gamma^2[-\dot{\gamma}^2 c^2 + \dot{\gamma}^2 u^2 + \gamma^2 a^2 + 2\gamma\dot{\gamma}ua] && (3.36)\\
&= -\dot{\gamma}^2 c^2 + \gamma^2[\gamma^2 a^2 + 2\gamma\dot{\gamma}ua] && (3.37)
\end{aligned}$$

where we've used the fact that **u** and **a** are parallel in this case. At this point, it's convenient to change to rapidities:

$$\begin{aligned}
\gamma &= \cosh\eta && (3.38)\\
\dot{\gamma} &= \dot{\eta}\sinh\eta && (3.39)\\
\beta\gamma &= \sinh\eta && (3.40)\\
\beta &= \tanh\eta && (3.41)\\
\dot{\beta} &= \frac{1}{\cosh^2\eta}\dot{\eta} && (3.42)
\end{aligned}$$

So the acceleration now becomes

$$\begin{aligned}
\mathsf{A}\cdot\mathsf{A} &= -c^2\dot{\eta}^2\sinh^2\eta + && (3.43)\\
&\quad \cosh^2\eta\left[c^2\cosh^2\eta\left(\frac{\dot{\eta}}{\cosh^2\eta}\right)^2 + 2c^2\cosh\eta(\dot{\eta}\sinh\eta)\tanh\eta\frac{\dot{\eta}}{\cosh^2\eta}\right] && (3.44)\\
&= -c^2\dot{\eta}^2\sinh^2\eta + c^2\dot{\eta}^2 + 2c^2\dot{\eta}^2\sinh^2\eta && (3.45)\\
&= c^2(\dot{\eta}^2 + \dot{\eta}^2\sinh^2\eta) && (3.46)\\
\frac{a_0^2}{c^2} &= \dot{\eta}^2\cosh^2\eta && (3.47)\\
\frac{a_0}{c} &= \dot{\eta}\cosh\eta && (3.48)\\
&= \frac{d}{dt}\sinh\eta && (3.49)\\
\sinh\eta &= \frac{1}{c}\int a_0(t)dt + C && (3.50)
\end{aligned}$$

3.3.3 Hyperbolic motion

Let's take the special case of a constant acceleration in the particle's rest frame (such as in the case of a rocket). This means a_0 is constant, and we'll take the rocket to start from rest in the lab frame. We then have

$$\begin{aligned}
\beta\gamma = \sinh\eta &= \frac{a_0 t}{c} && (3.51)\\
\frac{\beta}{\sqrt{1-\beta^2}} &= \frac{a_0 t}{c} && (3.52)\\
\beta^2 &= \left(\frac{a_0 t}{c}\right)^2(1-\beta^2) && (3.53)\\
\beta^2\left(1+\left(\frac{a_0 t}{c}\right)^2\right) &= \left(\frac{a_0 t}{c}\right)^2 && (3.54)\\
\beta &= \frac{a_0 t/c}{[1+(a_0 t/c)^2]^{1/2}} && (3.55)\\
u(t) &= \frac{a_0 t}{(1+a_0^2 t^2/c^2)^{1/2}} && (3.56)
\end{aligned}$$

At large t, note that

$$u(t) \to \frac{a_0 t}{a_0 t/c} = c \tag{3.57}$$

in other words, the speed approaches (and doesn't exceed) c, as we'd expect.

To calculate the distance travelled,

$$\frac{dx}{dt} = \beta c = c \tanh \eta \tag{3.58}$$

$$dx = c \tanh \eta dt \tag{3.59}$$

but we also know that

$$\sinh \eta = a_0 t/c$$

so

$$dt = \frac{c}{a_0} \cosh \eta d\eta \tag{3.60}$$

Then we get to integrate

$$dx = \frac{c^2}{a_0} \tanh \eta \cosh \eta d\eta \tag{3.61}$$

$$= \frac{c^2}{a_0} \sinh \eta d\eta \tag{3.62}$$

$$x = \frac{c^2}{a_0} \int \sinh \eta d\eta \tag{3.63}$$

$$= \frac{c^2}{a_0} \cosh \eta + b \tag{3.64}$$

$$x - b = \frac{c^2}{a_0} \cosh \eta = \frac{c^2}{a_0} (1 + \sinh^2 \eta)^{1/2} = \frac{c^2}{a_0} \left(1 + \left(\frac{a_0 t}{c} \right)^2 \right)^{1/2} \tag{3.65}$$

$$(x - b)^2 = \frac{c^4}{a_0^2} \left(1 + \frac{a_0^2 t^2}{c^2} \right) = \frac{c^2}{a_0^2} (c^2 + a_0^2 t^2) \tag{3.66}$$

$$(x - b)^2 - c^2 t^2 = \left(\frac{c^2}{a_0} \right)^2 \tag{3.67}$$

which is a hyperbola. This is in contrast with the Newtonian case, where a constant acceleration (such as uniform gravity on the surface of the Earth) gives a parabola.

3.3.4 Constant external force

Another meaning of "constant force" is a force $\mathbf{f}$ which is constant in time and space in a given frame S. An example would be the force of a uniform electric field $\mathbf{E}$ on a charge.

In this case

$$\frac{d\mathbf{p}}{dt} = \mathbf{f} \tag{3.68}$$

results in

$$\mathbf{p}(t) = \mathbf{p}_0 + \mathbf{f} t \tag{3.69}$$

If we take $\mathbf{p}_0 = 0$, then we have linear motion with $\mathbf{p}$ parallel to $\mathbf{f}$ at all times. Then we have

$$p = \gamma m u = ft = \frac{mu}{(1 - u^2/c^2)^{1/2}} \tag{3.70}$$

$$(1 - u^2/c^2) f^2 t^2 = m^2 u^2 \tag{3.71}$$

$$f^2 t^2 = \left(m^2 + \frac{f^2 t^2}{c^2} \right) u^2 \tag{3.72}$$

$$u(t) = \frac{ft}{(m^2 + f^2 t^2/c^2)^{1/2}} \tag{3.73}$$

which also approaches c as $t \to \infty$.

In fact, this is rather like hyperbolic motion: at any instantaneous rest frame, the force is the same as in the first frame at the start. At rest, $\gamma = 1$, so $f = ma_0$. All the conclusions and observations from hyperbolic motion then apply.

3.3.5 Circular motion

In this case, we have a force from a constant magnetic field

$$\mathbf{f} = q\mathbf{u} \wedge \mathbf{B} \tag{3.74}$$

The general equation of motion is then

$$\mathbf{f} = \gamma m \mathbf{a} + m \frac{d\gamma}{dt} \mathbf{u} = \gamma m \mathbf{a} + \frac{\mathbf{f} \cdot \mathbf{u}}{c^2} \mathbf{u} \tag{3.75}$$

The second term is zero, since the force is already a cross product of $\mathbf{u}$ and $\mathbf{B}$. Then we have a simple acceleration which for $\mathbf{u} \perp \mathbf{B}$,

$$\mathbf{f} = \gamma m \mathbf{a} \tag{3.76}$$

and the magnitude $f = quB$.

Remember that for a circle, it's still true (this is just normal 3D geometry)

$$a = \frac{u^2}{r} \tag{3.77}$$

so

$$r = \frac{u^2}{a} = \frac{u^2 \gamma m}{f} = \frac{\gamma m u^2}{quB} = \frac{\gamma m u}{qB} = \frac{p}{qB} \tag{3.78}$$

This is a simple relationship between the radius of a circular path and the momentum of the particle.

However, let's look at the period:

$$T = \frac{2\pi r}{u} = 2\pi \frac{\gamma m}{qB} \tag{3.79}$$

which introduces a dependence of the period on the speed, in contrast with the Newtonian result, which is independent of speed. This complicates trying to accelerate particles with a synchronized electric field in a synchrotron.

The circular motion result generalizes to helical motion, *i.e.*, linear in one direction, but circular in the transverse plane. For instance, consider a solenoidal magnetic field $\mathbf{B} = B\hat{\mathbf{z}}$.

Since $\mathbf{f} = q\mathbf{u} \wedge \mathbf{B}$, it's still true that $\mathbf{f} \cdot \mathbf{u} = 0$, so

$$\mathbf{f} = \gamma m\mathbf{a} = q\mathbf{u} \wedge \mathbf{B} = qB(\mathbf{u} \wedge \hat{\mathbf{z}}) = qB(u_y\hat{\mathbf{x}} - u_x\hat{\mathbf{y}}) \tag{3.80}$$

so the acceleration remains only in the plane transverse to the $\mathbf{B}$ field. This is a typical situation in a modern particle physics collider detector: you have a solenoidal magnetic field. Particles with some momentum in the z direction travel with a constant speed in that direction, while the curvature of the track is related simply with the transverse part of its momentum. In this way we can reconstruct the total 3-momentum of the particle emerging from the collision.

Unfortunately, this isn't perfect: this only works for charged particles which leave bits of energy in the detectors. And then to get the total 4-momentum, we need to add at least one more piece of information: this could be energy from a calorimeter (though this usually doesn't have very good resolution when compared with tracking detectors), speed from the time of flight going through the tracking volume, or mass-dependent energy loss in the detector. The latter two methods only work for relatively small momentum ranges, though. Instead, we often just "guess" the particle identity; what we depend on is that we can get so much data that the additional correlation that comes from real resonances peeks up above a smooth background level of random combinations.

Chapter 4

Lagrangians

You first ran into Lagrangians in CP1 as a way to come up with equations of motion. The non-relativistic Lagrangian is simply the difference between the kinetic and potential energies:

$$L(q_i, \dot{q}_i, t) = T - V \tag{4.1}$$

where T and V are evaluated for the different generalized coordinates q_i and their time derivatives $\dot{q}_i$. We then used Hamilton's Principle, which is that the classical path is the one for which the action integral

$$S[q(t)] = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t) dt \tag{4.2}$$

is stationary with respect to changes in the path. This resulted in equations of motion of the form

$$0 = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} \tag{4.3}$$

Note that action isn't a property of a particle. Instead, it's a functional (function of functions) with a specific job, which is to be stationary for classical paths. So there's nothing wrong with the idea of an action integral in Special Relativity, nor with finding a stationary path.

The question is whether we can write a Lagrangian which is consistent with Special Relativity. $T - V$ is not manifestly form-invariant: energy is a component of a vector, and so is a frame-dependent quantitiy. The Lagrangian also gives a special place to time (though S then integrates it out).

4.1 *Equations of particle motion from the Lagrangian

(Following Jackson 12.1)

We want an action which is invariant, so that the results derived from it will be invariant. Let's also change the integral above to use an invariant differential element, the proper time τ:

$$S[q(t)] = \int L dt = \int L\gamma d\tau \tag{4.4}$$

We see then that $L\gamma$ must be invariant. For a free-particle Lagrangian, the only invariants we have available are scalars and $\mathsf{U}\cdot\mathsf{U} = -c^2$. So we have as one possible Lagrangian

$$L = -mc^2/\gamma = -mc^2\left(1 - \frac{\dot{x}^2}{c^2}\right)^{1/2} \tag{4.5}$$

so that the action is

$$S[x(t)] = -mc^2\int\left(1 - \frac{\dot{x}^2}{c^2}\right)^{1/2} dt \tag{4.6}$$

And indeed one does get the relativistic equations of motion. The momentum is

$$\frac{\partial L}{\partial \dot{x}} = -\frac{1}{2}mc^2\left(1 - \frac{\dot{x}^2}{c^2}\right)^{-1/2}\left(\frac{-2\dot{x}}{c^2}\right) = \gamma m\dot{x} = p \tag{4.7}$$

which is the usual relativistic form, which because there is no dependence on x itself,

$$\frac{d}{dt}(\gamma m v) = 0 \tag{4.8}$$

The Lagrangian isn't manifestly "form-invariant" however. To do this, we need to replace it with things which transform properly. One possible replacement

$$L = -mc(-\mathsf{U}\cdot\mathsf{U})^{1/2} \tag{4.9}$$

with the action

$$S[\mathsf{X}(\tau)] = -mc\int(-\mathsf{U}\cdot\mathsf{U})^{1/2}d\tau \tag{4.10}$$

which is now a function only of Lorentz scalars and proper invariant intervals. (Note that this is a different action, not the old one transformed, so we're letting the γ of the old action drop out.) Along the classical path this is just the same as before, since $\mathsf{U}\cdot\mathsf{U} = -c^2$. So we have to vary U along the path, keeping in mind that in the end the constraint holds.

There's some subtlety to the limits in the integral, because simultaneity is lost over space-like distances; different paths may have different proper lengths, and therefore proper time. But we can define a function $s(\tau)$ which increases monotonically along with τ and does begin and end at the same value. Along this path,

$$\left(-\frac{d\mathsf{X}}{ds}\cdot\frac{d\mathsf{X}}{ds}\right)^{1/2} ds = \left(-\frac{d\mathsf{X}}{d\tau}\cdot\frac{d\mathsf{X}}{d\tau}\right)^{1/2} d\tau = (-\mathsf{U}\cdot\mathsf{U})^{1/2}d\tau \tag{4.11}$$

But, as I said, it's a subtle point: in the end, for the classical path, you get the right proper length.

First, rewrite the action integral in terms of s:

$$S[\mathsf{X}(s)] = -mc\int_{s_1}^{s_2}\left(-\frac{d\mathsf{X}}{ds}\cdot\frac{d\mathsf{X}}{ds}\right)^{1/2} ds = -mc\int_{s_1}^{s_2}(-\dot{\mathsf{X}}\cdot\dot{\mathsf{X}})^{1/2}ds \tag{4.12}$$

where the dot indicates a derivative with respect to s. The Lagrangian is

$$L = -mc(-\dot{\mathsf{X}} \cdot \dot{\mathsf{X}})^{1/2} = -mc(c^2\dot{t}^2 - \dot{x}^2 - \dot{y}^2 - \dot{z}^2)^{1/2} \tag{4.13}$$

We can now do the usual variation. We evaluate the derivative with respect to one of the space components:

$$\frac{\partial L}{\partial \dot{\mathsf{X}}^j} = mc(-\dot{\mathsf{X}} \cdot \dot{\mathsf{X}})^{-1/2}\dot{\mathsf{X}}^j \tag{4.14}$$

(I've anticipated some notational conventions on components we'll employ later.) This yields the equation of motion for a space component

$$0 = \frac{d}{ds}\left(\frac{mc\dot{\mathsf{X}}^j}{(-\dot{\mathsf{X}} \cdot \dot{\mathsf{X}})^{1/2}}\right) \tag{4.15}$$

To change from ds back to $d\tau$, we need to evaluate $\dot{\mathsf{X}}$

$$\frac{d\mathsf{X}^j}{ds} = \frac{d\tau}{ds}\frac{d\mathsf{X}^j}{d\tau} \tag{4.16}$$

$$= \frac{(-\dot{\mathsf{X}} \cdot \dot{\mathsf{X}})^{1/2}}{(-\mathsf{U} \cdot \mathsf{U})^{1/2}}\frac{d\mathsf{X}^j}{d\tau} \tag{4.17}$$

$$= \frac{1}{c}(-\dot{\mathsf{X}} \cdot \dot{\mathsf{X}})^{1/2}\frac{d\mathsf{X}^j}{d\tau} \tag{4.18}$$

Substituting this into the equation of motion, we get

$$0 = \frac{d}{ds}\left(m\frac{d\mathsf{X}^j}{d\tau}\right) \tag{4.19}$$

$$= \frac{(-\dot{\mathsf{X}} \cdot \dot{\mathsf{X}})^{1/2}}{c}\frac{d}{d\tau}\left(m\frac{d\mathsf{X}^j}{d\tau}\right) \tag{4.20}$$

$$= \frac{(-\dot{\mathsf{X}} \cdot \dot{\mathsf{X}})^{1/2}}{c}m\frac{d^2\mathsf{X}^j}{d\tau^2} \tag{4.21}$$

In general, $(-\dot{\mathsf{X}} \cdot \dot{\mathsf{X}})$ is not zero, so this simplifies to the usual equation of motion.

$$0 = m\frac{d^2\mathsf{X}^j}{d\tau^2} \tag{4.22}$$

There are other possible Lagrangian forms, such as

$$L = \frac{1}{2}m\mathsf{U} \cdot \mathsf{U} \tag{4.23}$$

which looks surprisingly like the old non-relativistic form, though U is quite a different entity from $\mathbf{u}$. Indeed, one can use

$$L = mf(\mathsf{U} \cdot \mathsf{U}) \tag{4.24}$$

where $f(y)$ is any function such that

$$\left.\frac{\partial f}{\partial y}\right|_{y=-c^2} = \frac{1}{2} \tag{4.25}$$

4.1.1 Central force problem

Consider a conservative central force,

$$\mathbf{f}(\mathbf{r}) = f(r)\hat{\mathbf{r}} \tag{4.26}$$

which can therefore be written in the form of a potential $V(r)$, depending only on the radius $r^2 = x^2 + y^2 + z^2$. This is obviously not a purely relativistic problem, since we essentially have the instantaneous transmission of any changes from the source to the body. But we can still consider it as an approximation, and ask whether even at that level Special Relativity affects the system in a noticeable way. For instance, the solar system is clearly dominated by the Sun (which we can consider stationary), and the speeds of planets are not particularly relativistic, but there may still be effects.

One way to write the Lagrangian would be

$$\mathcal{L} = -mc^2\sqrt{1 - \frac{\dot{\mathbf{r}}^2}{c^2}} - V(r) \tag{4.27}$$

(I am switching font to distinguish from an L we'll define later.) In polar coordinates,

$$\dot{\mathbf{r}}^2 = \dot{x}^2 + \dot{y}^2 = \dot{r}^2 + r^2\dot{\phi}^2 \tag{4.28}$$

so we have for the Lagrangian

$$\begin{aligned}
\mathcal{L} &= -mc^2[1 - \frac{1}{c^2}(\dot{r}^2 + r^2\dot{\phi}^2)]^{1/2} - V(r) && (4.29)\\
p_r = \frac{\partial\mathcal{L}}{\partial\dot{r}} &= -mc^2\frac{1}{2}\gamma(-\frac{2}{c^2}\dot{r}) = \gamma m\dot{r} && (4.30)\\
\frac{\partial\mathcal{L}}{\partial r} &= -mc^2\frac{1}{2}\gamma(-\frac{2r}{c^2}\dot{\phi}^2) - \frac{\partial V}{\partial r} = \gamma mr\dot{\phi}^2 - \frac{\partial V}{\partial r} && (4.31)\\
L = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} &= -mc^2\frac{1}{2}\gamma(-\frac{2}{c^2}r^2\dot{\phi}) = \gamma mr^2\dot{\phi} && (4.32)\\
\frac{\partial\mathcal{L}}{\partial\phi} &= 0 && (4.33)
\end{aligned}$$

Since it's clear that even with the relativistic modification, ϕ is cyclic, and L is conserved. This is rather like angular momentum, but with a γ factor.

The Hamiltonian is

$$\begin{aligned}
\mathcal{H} &= p_r\dot{r} + L\dot{\phi} - \mathcal{L} && (4.34)\\
&= \gamma m\dot{r}^2 + \gamma mr^2\dot{\phi}^2 + mc^2/\gamma + V && (4.35)\\
&= \gamma m\dot{\mathbf{r}}^2 + mc^2/\gamma + V && (4.36)\\
&= \gamma m\left(\dot{\mathbf{r}}^2 + c^2(1 - \frac{\dot{\mathbf{r}}^2}{c^2})\right) + V && (4.37)\\
&= \gamma mc^2 + V && (4.38)
\end{aligned}$$

which is pretty much the energy we expect. And since we know that $\mathcal{L}$ doesn't depend explicitly on time, we have $\mathcal{H}$ conserved.

All this looks rather like the non-relativistic case, but with a few γ's thrown in. This introduces a new speed dependence which can complicate matters. But let's see how far we can push this by making it look as much as possible like a 1D non-relativistic problem in radius r (as you did before in CP1), with "radial" kinetic energy and an effective potential.

$$E_{\text{eff}} = \frac{1}{2}m\left(\frac{dr}{d\tau}\right)^2 + V_{\text{eff}} \tag{4.39}$$

We change to proper time: this is just to get rid of stray factors, rather than because we want to look at it in a particular frame. (In fact, with this kind of analysis, we're looking for things which will be true regardless of frame.)

$$\begin{aligned}
\frac{dr}{d\tau} &= \frac{dr}{dt}\frac{dt}{d\tau} = \dot{r}\gamma = \frac{p_r}{m} && (4.40)\\
E - V &= \frac{p_r^2}{\gamma m} + \frac{L^2}{\gamma m r^2} + \frac{mc^2}{\gamma} && (4.41)\\
&= \frac{1}{\gamma m}\left(m\frac{dr}{d\tau}\right)^2 + \frac{L^2}{\gamma m r^2} + \frac{mc^2}{\gamma} && (4.42)\\
\gamma(E - V) &= m\left(\frac{dr}{d\tau}\right)^2 + \frac{L^2}{mr^2} + mc^2 && (4.43)\\
\frac{1}{2}m\left(\frac{dr}{d\tau}\right)^2 &= \frac{\gamma}{2}(E - V) - \frac{L^2}{2mr^2} - \frac{1}{2}mc^2 && (4.44)\\
&= \frac{(E-V)^2}{2mc^2} - \frac{L^2}{2mr^2} - \frac{m^2c^4}{2mc^2} && (4.45)\\
&= \frac{E^2 - 2EV + V^2 - m^2c^4}{2mc^2} - \frac{L^2}{2mr^2} && (4.46)\\
&= \frac{E^2 - m^2c^4}{2mc^2} - \frac{V(2E - V)}{2mc^2} - \frac{L^2}{2mr^2} && (4.47)\\
&= E_{\text{eff}} - V_{\text{eff}} && (4.48)
\end{aligned}$$

where

$$\begin{aligned}
E_{\text{eff}} &= \frac{E^2 - m^2c^4}{2mc^2} && (4.49)\\
V_{\text{eff}} &= \frac{V(2E - V)}{2mc^2} + \frac{L^2}{2mr^2} && (4.50)
\end{aligned}$$

For a Coulomb-like potential,

$$\begin{aligned}
\mathbf{f}(r) &= -\frac{\alpha}{r^2}\hat{\mathbf{r}} && (4.51)\\
V(r) &= -\frac{\alpha}{r} && (4.52)
\end{aligned}$$

which plugs into the effective potential

$$V_{\text{eff}} = \frac{1}{2mc^2}\left(-\frac{\alpha}{r}(2E + \frac{\alpha}{r})\right) + \frac{L^2}{2mr^2} = \frac{1}{2mc^2}\left(\frac{L^2c^2 - \alpha^2}{r^2} - \frac{2\alpha E}{r}\right) \tag{4.53}$$

The r^{-1} term dominates at large r, and the r^{-2} term at small r.

There's clearly a critical value of the angular momentum at $L_c = \alpha/c$. For small $L < L_c$, even when nonzero, $V_{\text{eff}} < 0$ for all r, with no inner turning point, so the particle is sucked into the center (some of our approximations will break down as one approaches the center, of course).

For larger angular momentum, $L > L_c$, there are regions of r for which the first term in V_{eff} will be positive and compete against the second term to provide a centrifugal barrier which will prevent the particle from reaching the center.

If $E_{\text{eff}} > 0$, the motion is unbound.

If $E_{\text{eff}} < 0$, then the particle will be in an orbit of some kind. However, it won't be the tidy ellipse of a straight Coulomb-like force. Instead, there will be a small orbit precession, albeit smaller than the one predicted by General Relativity.

Chapter 5

Further kinematics

In this section, we'll look at a few relativistic effects which arise when boosts aren't conveniently lined up parallel to one another.

5.1 *Doppler effect

Another useful 4-vector is that of the frequency and wavenumber of a wave

$$\mathsf{K} = (\omega/c, \mathbf{k}) \tag{5.1}$$

The fact that frequency and wavenumber form a 4-vector shouldn't be a surprise. For a photon in free space, for instance, you've seen that

$$E = \hbar\omega \tag{5.2}$$
$$\mathbf{p} = \hbar\mathbf{k} \tag{5.3}$$
$$\omega = |\mathbf{k}|c \tag{5.4}$$

which fits naturally into this scheme. Also recall a typical plane wave solution

$$\cos(\mathbf{k}\cdot\mathbf{x} - \omega t) \tag{5.5}$$

for which the argument looks remarkably like a dot product of some 4-vector with X.

If we take this as a 4-vector, then one immediate consequence is the effect of the Lorentz transformation on the frequency, if the photon is coming towards you:

$$\omega' = \gamma(\omega - \beta c k_{\|}) = \gamma(\omega - \beta\omega) = \omega\gamma(1-\beta) \tag{5.6}$$

which gives the simplest form of the Doppler shift formula

$$\frac{\omega'}{\omega} = \sqrt{\frac{1-\beta}{1+\beta}} \tag{5.7}$$

Looking at the momentum (or wavenumber) transformation, you get another interesting but simple consequence. If you consider the Lorentz transformations along the photon direction, you find that

$$p' = \gamma(p - \beta E/c) = \gamma(p - \beta p) = \gamma p(1 - \beta) \tag{5.8}$$

which means that if a photon is going in some direction, there is no boost parallel to that direction which can make it appear to be going backwards. This is not the case for particles with mass, since in that case $E/c > p$.

If you aren't in the path of the wave, you have to do the Lorentz transformation itself. Let's put the source in the moving S_0 frame. In that frame, it emits some light at an angle θ_0 relative to the x_0 axis. The 4-wave vector is

$$\mathsf{K} = (\omega_0/c, k_0 \cos\theta_0, k_0 \sin\theta_0, 0) \tag{5.9}$$

The lab frame is S, with the corresponding axes aligned ("standard configuration"). The source moves with speed v along the x axis. The transformation from S_0 to S is then

$$\begin{aligned} \omega/c &= \gamma(\omega_0/c + \beta k_0 \cos\theta_0) && (5.10)\\ k_x &= \gamma(k_0 \cos\theta_0 + \beta\omega_0/c) && (5.11)\\ k_y &= k_0 \sin\theta_0 && (5.12) \end{aligned}$$

The observed frequency is thus

$$\omega = \gamma\omega_0(1 + \beta c(k_0/\omega_0)\cos\theta_0) \tag{5.13}$$

and the observed angle

$$\tan\theta = \frac{k_y}{k_x} = \frac{\sin\theta_0}{\gamma(\cos\theta_0 + \beta\omega_0/k_0 c)} \tag{5.14}$$

If we want to get the Doppler effect only in terms of lab-frame observables, consider $\mathsf{K} \cdot \mathsf{U}$, where U is the 4-velocity of the source. In the source (S_0) frame, this is simply

$$(\omega_0/c, \mathbf{k}_0) \cdot (c, 0) = -\omega_0 \tag{5.15}$$

This is a scalar, so we've lost the angular information within the source's frame. In the lab frame, the scalar is

$$(\omega/c, \mathbf{k}) \cdot (\gamma c, \gamma\mathbf{u}) = -\omega\gamma + \gamma\mathbf{k} \cdot \mathbf{u} = \gamma\omega(1 - \frac{ku}{\omega}\cos\theta) \tag{5.16}$$

and therefore

$$\frac{\omega}{\omega_0} = \frac{1}{\gamma(1 - (u/v_p)\cos\theta)} \tag{5.17}$$

where $v_p = \omega/k$ is the phase velocity.

5.2 *Aberration

One should note that the angles of emission and observation are different. We've already seen the relationship above; a simpler calculation can just use the E/c and p_x equations of the Lorentz transformation:

$$\begin{aligned} \omega &= \gamma(\omega_0 + \beta\omega_0 \cos\theta_0) && (5.18)\\ \omega\cos\theta &= \gamma(\omega_0 \cos\theta_0 + \beta\omega_0) && (5.19)\\ \cos\theta &= \frac{\cos\theta_0 + \beta}{1 + \beta\cos\theta_0} && (5.20) \end{aligned}$$

A typical situation is the observation of a star which is far away from the Sun, so we can consider it "at rest" in the Sun's frame. Does the difference between emission and observed angles affect our observations of the star from the Earth, which has a speed $v \ll c$? We want to see the change in angle as a function of the observing angle, so we need to switch the previous equation around.

$$\begin{aligned} \cos\theta_0 &= \frac{\cos\theta - \beta}{1 - \beta\cos\theta} && (5.21)\\ &= (\cos\theta - \beta)(1 + \beta\cos\theta + O(\beta^2)) && (5.22)\\ &= \cos\theta - \beta(1 - \cos^2\theta) + O(\beta^2) && (5.23)\\ &\simeq \cos\theta - \beta\sin^2\theta && (5.24) \end{aligned}$$

The largest difference in angle comes when $\sin\theta = \pm 1$, or when the Earth's velocity is perpendicular to the line from the Earth to the star. As a result, a star directly above appears to move in a circle, while at another angle, a star moves in an ellipse.

The size of the effect is small, $\approx 2\beta \approx 0.0002$ radians, or $0.01°$. But this kind of precision was achieved in 1727, with James Bradley's observation.

5.3 Headlight effect

The fact that the emission and observed angles differ should also get you asking about the observed angular distribution.

The starting point for figuring this out is to consider how an element of solid angle transforms. The element of solid angle is

$$d\Omega = \sin\theta d\theta d\phi = d\cos\theta d\phi \tag{5.25}$$

Let's orient the axes such that the boost is along the $\theta = 0$ direction. Then ϕ is a transverse angle, unaffected by the boost, and we just have to consider

$$\frac{d\Omega}{d\Omega_0} = \frac{d\cos\theta}{d\cos\theta_0} \tag{5.26}$$

The easiest way of dealing with this is to consider $\cos\theta$ as a function, so $f = \cos\theta$ and $f_0 = \cos\theta_0$.

$$f = \frac{f_0 + \beta}{1 + \beta f_0} \tag{5.27}$$

$$\frac{df}{df_0} = \frac{1}{1+\beta f_0} - \frac{f_0+\beta}{(1+\beta f_0)^2}\beta \tag{5.28}$$

$$= \frac{1 + \beta f_0 - \beta(f_0+\beta)}{(1+\beta f_0)^2} \tag{5.29}$$

$$= \frac{1}{\gamma^2(1+\beta f_0)^2} \tag{5.30}$$

$$\frac{d\Omega}{d\Omega_0} = \frac{d\cos\theta}{d\cos\theta_0} = \frac{1}{\gamma^2(1+\beta\cos\theta_0)^2} \tag{5.31}$$

But recall also that for light

$$\omega = \gamma\omega_0(1+\beta\cos\theta_0) \tag{5.32}$$

so that

$$\frac{d\Omega}{d\Omega_0} = \left(\frac{\omega_0}{\omega}\right)^2 \tag{5.33}$$

The number of photons observed entering an element of solid angle is therefore

$$\frac{dN}{d\Omega} = \frac{dN}{d\Omega_0}\frac{d\Omega_0}{d\Omega} = \frac{dN}{d\Omega_0}\left(\frac{\omega}{\omega_0}\right)^2 \tag{5.34}$$

where $dN/d\Omega_0$ is the angular distribution in the source's rest frame.

If we want to find the energy flux into a solid angle element, we need to keep in mind that this isn't simply proportional to the number of photons. In fact, we get

$$\frac{dP}{d\Omega} = \left(\frac{\omega}{\omega_0}\right)^4 \frac{dP}{d\Omega_0} \tag{5.35}$$

which is a very strong beaming effect.

5.4 Generators

We saw earlier that rotations and Lorentz transformations look rather similar.

Before we get onto our next topic, we should look at infinitesimal transformations. For a rotation through a small angle θ around an axis,

$$\mathbf{R}(\theta) = 1 - i\theta\mathbf{J} + O(\theta^2) \tag{5.36}$$

where $\mathbf{J}$ is called the "generator" of the rotation. (Some texts aren't very strict about this definition, so you have to be careful when reading about them, but we'll try to use the above

as the definition of a generator.) It's easy to verify that the generators for rotations around the axes are

$$\mathbf{J}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \tag{5.37}$$

$$\mathbf{J}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix} \tag{5.38}$$

$$\mathbf{J}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.39}$$

As you might remember from quantum mechanics last year, there's nothing special about the axes, so we can generalize this using a dot product. A finite rotation can then be written as

$$\mathbf{R} = e^{-i\boldsymbol{\theta}\cdot\mathbf{J}} \tag{5.40}$$

where the dot product is taken to mean

$$\boldsymbol{\theta}\cdot\mathbf{J} = \theta_1\mathbf{J}_1 + \theta_2\mathbf{J}_2 + \theta_3\mathbf{J}_3 \tag{5.41}$$

You can verify this easily along one axis (say the z axis) for a finite angle by plugging in:

$$\mathbf{R}_3(\theta) = e^{-i\theta\mathbf{J}_3} = I - i\theta\mathbf{J}_3 - \frac{\theta^2}{2!}\mathbf{J}_3^2 + \cdots \tag{5.42}$$

We also note that

$$\mathbf{J}_3^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.43}$$

so the series ends up as

$$\mathbf{R}_3(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.44}$$

as expected.

For Lorentz transformations, the generators mix time and space coordinates.

$$\mathbf{K}_1 = \begin{pmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.45}$$

$$\mathbf{K}_2 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.46}$$

$$\mathbf{K}_3 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \tag{5.47}$$

A finite boost is then

$$\mathbf{L} = e^{-i\boldsymbol{\eta}\cdot\mathbf{K}} \tag{5.48}$$

A full Lorentz transformation or rotation is a 4×4 matrix which is antisymmetric in the space-space part, and symmetric in the time-space part. So there are 16 elements, but 10 constraints. This leaves 6 free parameters, which is what we'd expect given that we need 3 rotations and 3 boosts.

If you're worried that these infinitesimal generators don't really result in finite rotations or boosts, consider applying the infinitesimal generator (to first order) n times:

$$\prod_{k=1}^{n}(1 - i\theta\mathbf{J}) = \sum_{k=0}^{n} \binom{n}{k} (-i\theta\mathbf{J})^k \tag{5.49}$$

$$= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}(-i\theta\mathbf{J})^k \tag{5.50}$$

$$= 1 - in\theta\mathbf{J} - \frac{n(n-1)}{2!}\theta^2\mathbf{J}^2 + i\frac{n(n-1)(n-2)}{3!}\theta^3\mathbf{J}^3\cdots \tag{5.51}$$

$$= 1 - in\theta\mathbf{J} - \frac{n}{n-1}\frac{(n\theta)^2}{2!}\mathbf{J}^2 + i\frac{n}{n-1}\frac{n}{n-2}\frac{(n\theta)^3}{3!}\mathbf{J}^3\cdots \tag{5.52}$$

As you take $n \to \infty$ while keeping $\Theta = n\theta$ a finite constant, the fractional coefficients tend to 1, and you're left with the exponential.

Besides being useful for calculations, the generators also express the fact that rotations and Lorentz transformations are continuous symmetries: with the generators, you connect one system to a physical equivalent system by a series of infinitesimal steps. This becomes important later on.

It's interesting to note that the boost generators only involve the top row and the leftmost

column. Try multiplying two infinitesimal boosts together:

$$(1 - i\theta\mathbf{K}_1)(1 - i\phi\mathbf{K}_2) = \begin{pmatrix} 1 & -\theta & 0 & 0 \\ -\theta & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -\phi & 0 \\ 0 & 1 & 0 & 0 \\ -\phi & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.53}$$

$$= \begin{pmatrix} 1 & -\theta & -\phi & 0 \\ -\theta & 1 & \theta\phi & 0 \\ -\phi & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.54}$$

and we see a non-zero element pop up in spatial rotation part of the matrix. So we can see that the boosts in 3D don't form a closed set under multiplication. We'll look at the physical implication of this next.

5.5 Thomas precession

This section contains a direct derivation which depends heavily on the classic derivation of Jackson (Sec 11.8). For a more physical discussion, see Steane (Sec 6.7).

Consider an electron moving with velocity $\mathbf{v}(t)$ with respect to a lab frame S. We also consider the electron's instantaneous rest frames S'. Since we know the electron motion, we simply define the velocity of the electron frame with respect to S as $\mathbf{v}(t)$. So the transformation from S to S' is

$$\mathsf{X}' = \mathbf{A}(\boldsymbol{\beta})\mathsf{X} \tag{5.55}$$

We then define frame S'' to be the instantaneous rest frame at the next instant of time, $t + \delta t$. We write the velocity at that time as

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t) + \delta\mathbf{v} \tag{5.56}$$

The boost from S to S'' is then

$$\mathsf{X}'' = \mathbf{A}(\boldsymbol{\beta} + \delta\boldsymbol{\beta})\mathsf{X} \tag{5.57}$$

We want to find the relationship between S' and S'', *i.e.*,

$$\mathsf{X}'' = \mathbf{A}_T\mathsf{X} \tag{5.58}$$

From the above relations, we see that

$$\mathbf{A}_T = \mathbf{A}(\boldsymbol{\beta} + \delta\boldsymbol{\beta})\mathbf{A}^{-1}(\boldsymbol{\beta}) = \mathbf{A}(\boldsymbol{\beta} + \delta\boldsymbol{\beta})\mathbf{A}(-\boldsymbol{\beta}) \tag{5.59}$$

Let's choose a convenient coordinate system, with the first boost along x, and the second in the xy plane. This gives

$$\mathbf{A}(-\boldsymbol{\beta}) = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.60}$$

To get the next boost in the xy plane, we start from the general Lorentz transformation, which can be written as

$$\mathbf{A}(\boldsymbol{\beta}) = \begin{pmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ -\gamma\beta_x & 1+\alpha\beta_x^2 & \alpha\beta_x\beta_y & \alpha\beta_x\beta_z \\ -\gamma\beta_y & \alpha\beta_x\beta_y & 1+\alpha\beta_y^2 & \alpha\beta_y\beta_z \\ -\gamma\beta_z & \alpha\beta_x\beta_z & \alpha\beta_y\beta_z & 1+\alpha\beta_z^2 \end{pmatrix} \tag{5.61}$$

where $\alpha = \gamma^2/(\gamma+1)$. The boost in $\boldsymbol{\beta}+\delta\boldsymbol{\beta}$ can be calculated to first order as follows:

$$\begin{aligned} \beta_x &\rightarrow \beta + \delta\beta_x & (5.62)\\ \beta_y &\rightarrow \delta\beta_y & (5.63)\\ \gamma &\rightarrow \gamma + \frac{\partial\gamma}{\partial\beta_x}\delta\beta_x = \gamma + \gamma^3\beta\delta\beta_x & (5.64) \end{aligned}$$

We also note that

$$\alpha\beta^2 = \frac{\gamma^2}{\gamma+1}\beta^2 = \frac{\gamma^2}{\gamma+1}\frac{\gamma^2-1}{\gamma^2} = \gamma - 1 \tag{5.65}$$

The result is

$$\mathbf{A}(\boldsymbol{\beta}+\delta\boldsymbol{\beta}) = \begin{pmatrix} \gamma+\gamma^3\beta\delta\beta_x & -(\gamma\beta+\gamma^3\delta\beta_x) & -\gamma\delta\beta_y & 0 \\ -(\gamma\beta+\gamma^3\delta\beta_x) & \gamma+\gamma^3\beta\delta\beta_x & \left(\frac{\gamma-1}{\beta}\right)\delta\beta_y & 0 \\ -\gamma\delta\beta_y & \left(\frac{\gamma-1}{\beta}\right)\delta\beta_y & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.66}$$

Multiplying the two together,

$$\mathbf{A}_T = \begin{pmatrix} 1 & -\gamma^2\delta\beta_x & -\gamma\delta\beta_y & 0 \\ -\gamma^2\delta\beta_x & 1 & \left(\frac{\gamma-1}{\beta}\right)\delta\beta_y & 0 \\ -\gamma\delta\beta_y & -\left(\frac{\gamma-1}{\beta}\right)\delta\beta_y & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.67}$$

In terms of infinitesimal generators

$$\mathbf{A}_T = I + i\left(\frac{\gamma-1}{\beta^2}\right)(\boldsymbol{\beta}\wedge\delta\boldsymbol{\beta})\cdot\mathbf{J} + i(\gamma^2\delta\boldsymbol{\beta}_\parallel + \gamma\delta\boldsymbol{\beta}_\perp)\cdot\mathbf{K} \tag{5.68}$$

We can separate this expression, to first order, into a product of a rotation and a boost.

$$\begin{aligned} \mathbf{A}_T &= \mathbf{R}(\Delta\boldsymbol{\Omega})\mathbf{A}(\Delta\boldsymbol{\beta}) & (5.69)\\ \mathbf{R}(\Delta\boldsymbol{\Omega}) &= I + i\Delta\boldsymbol{\Omega}\cdot\mathbf{J} & (5.70)\\ \mathbf{A}(\Delta\boldsymbol{\beta}) &= I + i\Delta\boldsymbol{\beta}\cdot\mathbf{K} & (5.71) \end{aligned}$$

where the new angles and rapidities are

$$\begin{aligned} \Delta\boldsymbol{\Omega} &= \left(\frac{\gamma-1}{\beta^2}\right)(\boldsymbol{\beta}\wedge\delta\boldsymbol{\beta}) = \frac{\gamma^2}{\gamma+1}(\boldsymbol{\beta}\wedge\delta\boldsymbol{\beta}) & (5.72)\\ \Delta\boldsymbol{\beta} &= \gamma^2\Delta\boldsymbol{\beta}_\parallel + \gamma\Delta\boldsymbol{\beta}_\perp & (5.73) \end{aligned}$$

We defined S, S', and S'' to have parallel axes. But since the two boosts have resulted in a rotation, it's natural to interpret the end result as one boost, but with a rotated coordinate system. So let's define S''', which only has that boost $\mathbf{A}(\Delta\boldsymbol{\beta})$:

$$\mathsf{X}''' = \mathbf{A}(\Delta\boldsymbol{\beta})\mathsf{X}' \tag{5.74}$$

Since we saw that

$$\mathbf{A}_T = \mathbf{A}(\boldsymbol{\beta}+\delta\boldsymbol{\beta})\mathbf{A}(-\boldsymbol{\beta}) = \mathbf{R}(\Delta\boldsymbol{\Omega})\mathbf{A}(\Delta\boldsymbol{\beta}) \tag{5.75}$$

we can isolate the boost part

$$\mathbf{A}(\Delta\boldsymbol{\beta}) = \mathbf{R}(-\Delta\boldsymbol{\Omega})\mathbf{A}(\boldsymbol{\beta}+\delta\boldsymbol{\beta})\mathbf{A}(-\boldsymbol{\beta}) \tag{5.76}$$

We then have the transformation

$$\begin{aligned}
\mathsf{X}''' &= \mathbf{R}(-\Delta\boldsymbol{\Omega})\mathbf{A}(\boldsymbol{\beta}+\delta\boldsymbol{\beta})\mathbf{A}(-\boldsymbol{\beta})\mathsf{X}' && (5.77)\\
&= \mathbf{R}(-\Delta\boldsymbol{\Omega})\mathbf{A}(\boldsymbol{\beta}+\delta\boldsymbol{\beta})\mathsf{X} && (5.78)\\
&= \mathbf{R}(-\Delta\boldsymbol{\Omega})\mathsf{X}'' && (5.79)
\end{aligned}$$

So we see that the S''' axes are rotated by $\Delta\boldsymbol{\Omega}$ relative to S''.

What happens to a physical vector, such as a spin, then? The angular frequency is defined as

$$\boldsymbol{\omega}_T = \lim_{\delta t\to 0}\frac{-\Delta\boldsymbol{\Omega}}{\delta t} = -\lim_{\delta t\to 0}\frac{\gamma^2}{\gamma+1}(\boldsymbol{\beta}\wedge\frac{\delta\boldsymbol{\beta}}{\delta t}) \tag{5.80}$$

where the velocities and accelerations are measured in the lab frame. Taking the limit,

$$\boldsymbol{\omega}_T = \frac{\gamma^2}{\gamma+1}\frac{\mathbf{a}\wedge\mathbf{v}}{c^2} \tag{5.81}$$

For a physical vector $\mathbf{G}$,

$$\left(\frac{d\mathbf{G}}{dt}\right)_{\text{non-rot}} = \left(\frac{d\mathbf{G}}{dt}\right)_{\text{rest}} + \boldsymbol{\omega}_T\wedge\mathbf{G} \tag{5.82}$$

which is the Thomas precession. This is a purely kinematic effect which occurs whenever an acceleration has a component perpendicular to the velocity. It's independent of other effects, such as the precession of a magnetic moment in a magnetic field.

The original problem was an early one in atomic physics, explaining the anomalous Zeeman effect without messing up how fine structure was understood. The original hypothesis (Uhlenbeck-Goudsmit) only considered the rest frame behavior of the electron spin, which resulted in an interaction potential

$$U = -\frac{ge}{2mc}\mathbf{s}\cdot\mathbf{B} + \frac{g}{2m^2c^2}(\mathbf{s}\cdot\mathbf{L})\frac{1}{r}\frac{dV}{dr} \tag{5.83}$$

The problem was that $g=2$ explained the anomalous Zeeman splitting, but applying $g=2$ to the spin-orbit coupling gave splittings twice as large as observed. Thomas realized that there was a missing step the effect of which was to reduce g in the spin-orbit term to $g-1$. The reduced factor thus restored the explanation of the spin-orbit coupling along with the anomalous Zeeman effect.

Chapter 6

Scalars, Vectors, and Tensors

Now that we've reviewed specifics of Special Relativity, let's look at the general structure.

We'll start by looking at things which have well-defined transformation properties under a change of frame.

6.1 Generalized Lorentz transformation

It was mentioned earlier that one can define a generalized transformation as one which preserves the norm of a 4-vector

$$s^2 = -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \tag{6.1}$$

We did this by defining the "dot product" in a matrix form

$$\mathsf{A} \cdot \mathsf{B} = \mathbf{A}^T \mathbf{g} \mathbf{B} \tag{6.2}$$

where

$$\mathbf{g} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.3}$$

This length is obviously related to, but is not identical to, the Euclidean norm. In fact, it's not positive-definite, so it can't really be one of those. Technically, it's a "bilinear" form, which is just a term used for a linear in both inputs, but doesn't have to be positive-definite.

Before, we just defined the dot product in this way, and you had to remember that you need to insert the matrix in that way. Let's generalize in such a way that this looks less *ad hoc*. A further advantage is that this structure can be taken directly into General Relativity.

6.1.1 Index notation

We'll restrict ourselves for now to a flat space with Cartesian coordinates. Of course one can also deal with different coordinate systems, such as cylindrical and spherical systems,

but curvilinear systems are perhaps most naturally discussed in a further course. Cartesian coordinates will best illustrate what we need at this point.

We designate 4-vectors with a superscripted symbol

$$\mathsf{X} \to x^\mu \tag{6.4}$$

where μ runs from 0 through 3. The components have the values

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) \tag{6.5}$$

When we use index notation, we'll define the components such that they all have the same units, in this case length.

If we need powers of such components, we usually put the superscripted symbol in parentheses. For instance, the norm of the vector above is

$$\mathsf{X} \cdot \mathsf{X} = -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 \tag{6.6}$$

With the explicit mention of components, this may actually seem like a step backwards: we've been trying to get you to think in terms of the abstract concepts, but suddenly the components have re-appeared. The main reason we use the indices, however, is not to write components explicitly, but to keep track of how the objects transform.

The "rank" of an object is the total number of indices the object has. So 4-vectors have a rank of 1, and scalars have a rank of 0. Objects with higher rank we usually call "tensors".

Notationally, if we're using indices, it will be obvious whether we're talking about a scalar, a vector, or a tensor, so we often drop any typographical conventions on the object itself, but we sometimes keep them if some additional clarity is needed.

Now for some things which the indices will tell us:

- the range:
  - Latin letters will indicate 3-space indices, ranging from 1 to 3, unless otherwise noted.
  - Greek letters will indicate 4-space indices, ranging from 0 to 3. The 0 component will be the time-like component (so $x^0 = ct$ to keep them all of them in the same units), and 1 to 3 will be the space-like components.
- the transformation style: this is indicated by whether the index is a superscript ("contravariant") or a subscript ("covariant").
  - Contravariant components transform in the way you'd normally expect: if the transformation is

    $$x^\mu \to x'^\mu = x'^\mu(\{x^\nu\}) \tag{6.7}$$

    (meaning that the transformed component x'^μ is a function of the set of untransformed components x^ν), then the components of a vector A transform as

    $$A'^\mu = \sum_{\nu=0}^{3} \frac{\partial x'^\mu}{\partial x^\nu} A^\nu \tag{6.8}$$

– Covariant components transform using the inverse transformation:

$$A'_\mu = \sum_{\nu=0}^{3} \frac{\partial x^\nu}{\partial x'^\mu} A_\nu \tag{6.9}$$

6.1.2 Summation convention

We will also introduce a "summation convention", by which we sum over indices which appear in both superscript and subscript. Since we don't usually perform operations on individual components, it's convenient to omit the redundant summation signs. So we can write

$$\mathsf{P} \cdot \mathsf{Q} = P^\mu Q_\mu = P_\mu Q^\mu \tag{6.10}$$

which we sometimes call "contracting" over the index. The last equality is useful in its own right; you can always swap upper and lower indices in this way.

We can recognize from this inner product that vectors with covariant components form a dual space with vectors with contravariant components. Note that this doesn't completely map into matrix form: there, the transpose vectors can be seen as the dual space, while the inner product explicitly includes the $\mathbf{g}$ matrix.

For space-only quantities, the upper and lower indices don't matter (for Special Relativity), so we can be relaxed and still sum over any repeated indices, whether super or subscript.

$$\mathbf{p} \cdot \mathbf{q} = p_i q_i \tag{6.11}$$

In general, if you're *not* supposed to sum over indices, we'll try to mention it.

6.1.3 Metric

We write the metric as $g_{\mu\nu}$, with the components as in the matrix. The form of the metric we're using is said to have a "signature" $(-1, 1, 1, 1)$, obviously given by the values of the diagonal elements.

It should be mentioned that other metrics are possible. For instance, the metric commonly used in particle physics has the opposite signature, $(1, -1, -1, -1)$. It's just a convention, but slightly more annoying than most, because it introduces different signs. If you are reading a book on Special or General Relativity, this is one of the first things you have to check in order to interpret any equations.

One of the important uses of the metric is to raise and lower indices:

$$x_\mu = g_{\mu\nu} x^\nu \tag{6.12}$$

(Remember the sum over ν.) By only summing over upper and lower indices, the summation convention provides a handy device to remember when you need to introduce a metric. If you have one vector with contravariant and another with covariant components, taking an inner

product is easy, and doesn't need any metric. If you have two vectors with contravariant components, however, it's obvious you need to lower one, and thus

$$\mathsf{P} \cdot \mathsf{Q} = g_{\mu\nu} P^\mu Q^\nu \tag{6.13}$$

When we write a vector in terms of its components, what we're writing are the forms of contravariant components. So, for instance, we have

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) \tag{6.14}$$

We could write covariant components instead, but whenever we do this we'll indicate it explicitly, either in text or by showing the lowered index:

$$x_\mu = (x_0, x_1, x_2, x_3) = (-ct, x, y, z) \tag{6.15}$$

Properties of the Minkowski metric:

- The matrix elements $g_{\mu\nu}$ are the same as $g^{\mu\nu}$.
- The metric is related to the Kronecker delta function

  $$g^{\mu\lambda} g_{\lambda\nu} = \delta^\mu_\nu \tag{6.16}$$

  which shouldn't be surprising given the matrix forms (and what we saw earlier).

6.2 Lorentz transformation matrices

As we saw in the beginning of the course, the Lorentz transformation (on contravariant components) takes the matrix form

$$\mathbf{X}' = \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \mathbf{L}\mathbf{X} \tag{6.17}$$

We rephrase this in index notation as follows:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{6.18}$$

where the matrix Λ is defined to be

$$\Lambda^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.19}$$

Even though matrices don't themselves have contravariant or covariant indices (they transform such indices), it's convenient to write them in the same style. The order of the indices

follows the usual convention, of the row being listed first, whether as an upper or lower index, followed by the column. Indices can be raised and lowered using the metric as with vectors (and tensors).

The covariant transformation must have the form

$$x'_\mu = x_\nu (\Lambda^{-1})^\nu{}_\mu \tag{6.20}$$

The order of the factors is simply to make the matrix-like multiplication explicit; the factors themselves commute, since they just involve numbers. The index order also makes explicit that

$$\Lambda^\mu{}_\nu (\Lambda^{-1})^\nu{}_\kappa = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\nu}{\partial x'^\kappa} = I \tag{6.21}$$

We can find more about the form of $(\Lambda^{-1})^\nu{}_\kappa$ by combining the two transforms:

$$\begin{aligned} x'_\mu &= g_{\mu\lambda} x'^\lambda && (6.22)\\ &= g_{\mu\lambda}\Lambda^\lambda{}_\kappa x^\kappa && (6.23)\\ &= g_{\mu\lambda}\Lambda^\lambda{}_\kappa g^{\kappa\nu} x_\nu && (6.24) \end{aligned}$$

which leads to the identification

$$(\Lambda^{-1})^\nu{}_\mu = g_{\mu\lambda}\Lambda^\lambda{}_\kappa g^{\kappa\nu} \tag{6.25}$$

Indeed, since the g's raise and lower indices,

$$(\Lambda^{-1})^\nu{}_\mu = \Lambda_\mu{}^\nu \tag{6.26}$$

This should not be mistaken for a transpose: it is better to consider it a mnemonic reminding you that you need to use the metric to raise and lower indices to get back to the familiar $\Lambda^\nu{}_\mu$, with the upper and lower indices in the right order.

In order to see how Λ^{-1} appears as a matrix, we go back to the more explicit formula

$$\begin{aligned} (\Lambda^{-1})^\nu{}_\kappa &= g^{\nu\mu} g_{\kappa\lambda}\Lambda^\lambda{}_\mu && (6.27)\\ &= g^{\nu\mu}(g_{\kappa\lambda}\Lambda^\lambda{}_\mu) && (6.28)\\ &= g^{\nu\mu}(g\Lambda)_{\kappa\mu} && (6.29)\\ &= g^{\nu\mu}(g\Lambda)^T_{\mu\kappa} && (6.30)\\ &= (g(g\Lambda)^T)^\nu{}_\kappa && (6.31) \end{aligned}$$

(We have slipped in a matrix transpose, which is safe in this case since the two indices are of one kind.) We then have as a matrix formula

$$\Lambda^{-1} = g\Lambda^T g \tag{6.32}$$

which also implies, since $gg = I$,

$$\Lambda^T = g\Lambda^{-1} g \tag{6.33}$$

Let's look at the invariance rule again, now in index notation.

$$g_{\mu\nu}A^\mu B^\nu = g_{\mu\nu}A'^\mu B'^\nu = g_{\mu\nu}(\Lambda^\mu{}_\kappa A^\kappa)(\Lambda^\nu{}_\lambda B^\lambda) = A^\kappa(\Lambda^\mu{}_\kappa g_{\mu\nu}\Lambda^\nu{}_\lambda)B^\lambda \tag{6.34}$$

If we match the components we sum over on the far left and far right sides (μ and ν on the left, and κ and λ on the right), then we have an equation which looks like a transformation of the metric:

$$g_{\mu\nu} = \Lambda^{\kappa}{}_{\mu}\Lambda^{\lambda}{}_{\nu}g_{\kappa\lambda} \tag{6.35}$$

However, in this case it's actually a more general way to define the Lorentz transformation Λ, as one which leaves the metric (and therefore any intervals) invariant.

In matrix form, one rewrites the relationship in the form

$$g_{\mu\nu} = \Lambda^{\kappa}{}_{\mu}g_{\kappa\lambda}\Lambda^{\lambda}{}_{\nu} \tag{6.36}$$

in order to make the index order match that of matrix multiplication. The first multiplication, however, combines two row indices (the κ index), which implies a multiplication by a transposed matrix:

$$g = \Lambda^{T} g \Lambda \tag{6.37}$$

which is the familiar result in matrix form.

6.3 Tensors

As we mentioned earlier, scalars are 0-rank tensors, and vectors 1-rank tensors. Scalars are left unchanged by Lorentz tranformations, while vectors are transformed (contravariantly or covariantly) by the application of one Λ. Higher rank objects are defined by similar transformation properties. A rank-2 tensor with contravariant components $M^{\mu\nu}$ transforms under the Lorentz transformation Λ as

$$M'^{\mu\nu} = \Lambda^{\mu}{}_{\kappa}\Lambda^{\nu}{}_{\lambda}M^{\kappa\lambda} \tag{6.38}$$

With covariant components,

$$M'_{\mu\nu} = (\Lambda^{-1})^{\kappa}{}_{\mu}(\Lambda^{-1})^{\lambda}{}_{\nu}M_{\kappa\lambda} \tag{6.39}$$

There are also tensors with mixed components:

$$M'^{\mu}{}_{\nu} = \Lambda^{\mu}{}_{\kappa}(\Lambda^{-1})^{\lambda}{}_{\nu}M^{\kappa}{}_{\lambda} \tag{6.40}$$

The order of indices can therefore be important. Individual indices can be raised and lowered using the metric as with vectors.

$$M^{\mu}{}_{\nu} = g_{\nu\kappa}M^{\mu\kappa} \tag{6.41}$$

Higher rank tensors can be made from lower-rank objects by taking an "outer" or "tensor" product. For instance, if you have tensors A^{μ} and B^{ν}, you can form

$$C^{\mu\nu} = A^{\mu}B^{\mu} \tag{6.42}$$

This obviously transforms as a rank-2 tensor.

Tensors with rank greater than 2 transform as you'd expect, by piling on more Λ's in the obvious way. The rank of the tensor is equal to the total number of indices.

The way to reduce the rank of tensors is to contract covariant and contravariant indices. So for instance, if you had a rank-3 tensor $M^{\mu\nu}{}_{\kappa}$, you can obtain rank-1 tensors $M^{\kappa\nu}{}_{\kappa}$ and $M^{\mu\kappa}{}_{\kappa}$.

Let's look again at the invariance of the metric, but this time distinguishing that original metric g with one "transformed" along with several vectors.

$$g_{\mu\nu}A^{\mu}B^{\nu} = g'_{\mu\nu}A'^{\mu}B'^{\nu} = g'_{\mu\nu}(\Lambda^{\mu}{}_{\kappa}A^{\kappa})(\Lambda^{\nu}{}_{\lambda}B^{\lambda}) = A^{\kappa}(\Lambda^{\mu}{}_{\kappa}g'_{\mu\nu}\Lambda^{\nu}{}_{\lambda})B^{\lambda} \tag{6.43}$$

To isolate $g'_{\mu\nu}$, we match the elements which multiply particular components of A^{μ} and B^{ν}, and then multiply by Λ^{-1}'s (note that the left-multiplication is of the matrix transpose):

$$\Lambda^{\kappa}{}_{\mu}g'_{\kappa\lambda}\Lambda^{\lambda}{}_{\nu} = g_{\mu\nu} \tag{6.44}$$

$$\Lambda^{\kappa}{}_{\mu}(\Lambda^{-1})^{\mu}{}_{\alpha}g'_{\kappa\lambda}\Lambda^{\lambda}{}_{\nu}(\Lambda^{-1})^{\nu}{}_{\beta} = (\Lambda^{-1})^{\mu}{}_{\alpha}g_{\mu\nu}(\Lambda^{-1})^{\nu}{}_{\beta} \tag{6.45}$$

The multiplication order on the left side looks odd, but remember that we are multiplying by numbers and the summing over the index μ. We do it in this way in order to reduce that part to an identity matrix.

$$\delta^{\kappa}_{\alpha}g'_{\kappa\lambda}\delta^{\lambda}_{\beta} = (\Lambda^{-1})^{\mu}{}_{\alpha}g_{\mu\nu}(\Lambda^{-1})^{\nu}{}_{\beta} \tag{6.46}$$

$$g'_{\alpha\beta} = (\Lambda^{-1})^{\mu}{}_{\alpha}(\Lambda^{-1})^{\nu}{}_{\beta}g_{\mu\nu} \tag{6.47}$$

So we notice that $g_{\mu\nu}$ has the transformation rule of two vectors with covariant components (though we know in the end that it's invariant). In this sense it is properly a rank-2 tensor with covariant components.

It should be mentioned that the tensors we'll be using in this course are exclusively Cartesian tensors, with Cartesian coordinates.

It is also worth noting that while we can represent rank-2 tensors with matrices, not all matrices are rank-2 tensors—even though they look very similar when we write them down, and we use the same index notation and summation convention on both. But they are rather different objects:

- Matrices such as $\Lambda^{\mu}{}_{\nu}$ represent linear operators, and the principal way to combine linear operators is to multiply them (though they do form linear spaces in their own right).
- Tensors represent elements of a linear space. As such, the way to combine tensors is to make linear combinations of them, *i.e.*, with addition and multiplication by scalars.

6.4 *4-gradient

We will see our tensors in differential equations and Euler-Lagrange equations, which means we'll need to take their derivatives. It turns out we can form a vector out of the differential operator. Consider the chain rule in light of a coordinate transformation:

$$\frac{\partial f}{\partial x'^{\mu}} = \frac{\partial x^{\nu}}{\partial x'^{\mu}}\frac{\partial f}{\partial x^{\nu}} \tag{6.48}$$

When we compare this to the tensor transformation rules, we find that the derivative with respect to contravariant components acts like a covariant vector. We denote the derivative as $\partial_\mu f$.

Similarly, derivatives with respect to covariant components act like a contravariant component:

$$\partial^\mu \equiv \frac{\partial}{\partial x_\mu} \tag{6.49}$$

For convenience, we will still assume non-indexed components are contravariant. So in terms of those usual components,

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{6.50}$$

$$\partial^\mu \equiv \frac{\partial}{\partial x_\mu} = g^{\mu\nu}\frac{\partial}{\partial x^\nu} = \left(-\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{6.51}$$

$$\tag{6.52}$$

The **d'Alembertian** operator is a 4-vector analogue of the ∇^2 operator:

$$\Box^2 \equiv \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} = \partial_\mu \partial^\mu \tag{6.53}$$

(In some textbooks, and in recent A2 notes, the d'Alembertian is simply $\Box$, but I prefer to preserve the indication of its nature as a second-derivative operator.)

Chapter 7

Groups

The reason we've gone into all this about tensors is that we want to write down actions which respect Lorentz symmetries. This implies that our actions (and Lagrangians) need to be scalars, *i.e.*, rank-0 tensors.

The main symmetry we've looked at is invariance with respect to Lorentz transformations, so the physics is the same in all inertial frames. Implied in Lorentz invariance is also invariance with respect to purely spatial rotations. We don't usually call rotations Lorentz transformations, but as we've seen with the Thomas precession, they play a role among Lorentz transformations.

Another symmetry is that of translational invariance, which we'll touch on later.

In order to understand these symmetries, we'll draw upon results from a branch of mathematics called "group theory".

A group is a set G and an operator $(\cdot)$ such that

1. Closure: for all $a, b \in G$, then $a \cdot b \in G$ as well.
2. Associative property: for all $a, b, c \in G$,
$$(a \cdot b) \cdot c = a \cdot (b \cdot c)$$
3. Identity element: there is an element $e \in G$ such that $a \cdot e = a$ for all $a \in G$.
4. Inverse element: for each element $a \in G$, there is an element $a^{-1} \in G$ such that $a \cdot a^{-1} = e$.

Since this will take a little effort, we'll start off by spoiling the punchline: a symmetry means there are equivalent configurations in a system, and groups allow us mathematically to "traverse" this space of equivalent configurations. One of the objectives is to find out what kinds of physical objects—such as scalars, vectors, and tensors—possess the symmetry, and therefore can be used to formulate physical laws. In the process, we'll find that there are more objects with such symmetry. Some of these are realized in Nature, while others are not—or, perhaps, just not yet. Symmetry thus becomes one of the guiding principles in the exploration of physics.

7.1 Example: permutation group

Let's look at a simple example to illustrate how we will use this construct. Consider a system of three identical bodies placed at the vertices of an equilaterial triangle. Label the bodies a, b, and c. It is obvious that there are 6 equivalent configurations: abc, bca, cab, acb, cba, and bac.

The configurations themselves don't form a group. After all, what would be the operation?

Instead, consider the transformations which get you from one configuration to another equivalent configuration. Now label the *positions* 1, 2, and 3. An explicit way to write a transformation is, for example,

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \tag{7.1}$$

which means that you take whatever body is in position 1 and move it to position 2, at the same time moving the body in position 2 to position 3, and the body in position 3 to position 1. This is, by the way, called a "cyclic permutation", and the group of operations called the "permutation group of degree 3".

The order of the first row is obviously arbitrary, so we'll have a conventional order 123. Then we can identify the group elements by the bottom row: 123 (which happens to be the identity e), 231, 312, 213, 132, and 321. Of course this looks like the configurations themselves, because we've (arbitrary) chosen one starting point and enumerated the ways to get to all the others.

You can form a multiplication table with these elements, with the first operation on the left, and the second operation listed on the top:

	e	231	312	213	132	321
e	e	231	312	213	132	321
231	231	312	e	321	213	132
312	312	e	231	132	321	213
213	213	132	321	e	231	312
132	132	321	213	312	e	231
321	321	213	132	231	312	e

We can see that the group is closed, and that every element has an inverse. It is obviously *not* commutative. (The term for a group with commutative multiplication is "Abelian"; the permutation group is "non-Abelian".)

It is also evident that the multiplication results are not randomly scattered everywhere. For instance, the elements e, 231, and 312 all multiply amongst themselves; they form a subset which we call a "subgroup". In fact, they are the cyclic permutations.

Also notice that the other quadrants are also self-contained. This reflects the fact that 213, 132, and 321 involve a single swap rather than a cyclic permutation; further cyclic permutations keep you within the quadrant.

Finally, it is evident you don't need all the elements in order to traverse the entire group from a single starting point. In fact, all you need is a cyclic permutation p and a swap s.

Then the elements can be written as e, p, pp, s, sp, and spp. Expressions such as ppp and ss get you back to e, so pp is clearly the inverse of p, and s is its own inverse.

These sorts of discrete groups are important in physics, and even more so in chemistry where you have crystals with these sorts of symmetries. It's really a couse in itself.

7.2 Rotations

In this course, however, we want to look at systems with continuous symmetries. For instance, think of the rotations in 3D of a rigid body like a block of wood with unequal sides. By analogy with the permutation group, choose a starting orientation, and assign to every orientation a group element which takes you from the starting to final orientation. There are clearly an infinite number of group elements, parameterized by continuous parameters. Such a group is commonly called a "Lie group". In this case there are 3 real parameters. And in fact we've seen the infinitesimal rotations already with the generators $\mathbf{J}_i$:

$$\mathbf{R}_i(\delta\theta) = 1 - i\delta\theta\mathbf{J}_i \tag{7.2}$$

From any starting orientation, you can use the three generators to get to any orientation of the body itself.

However, as in the case of the permutation group, there is a whole set of configurations which cannot be reached. These are the ones which are related to the original orientation by a parity transformation (or spatial inversion, or reflection)

$$(x, y, z) \to (-x, -y, -z) \tag{7.3}$$

With an odd number of spatial dimensions, there is no way to get there via infinitesimal rotations.

(In fact it may help to think of a set of vectors, or an extended body, rather than just one vector. The reason is that you can always make a single vector look as if it has gone through a reflection. For instance, you can rotate the vector $(1, 1, 1)$ to get $(-1, -1, -1)$ by rotating by an angle π around the axis pointing in the $(1, 1, -1)$ direction away from the origin. What won't be preserved is its relative orientation with respect to another vector.)

At the same time, the reflection clearly leaves all the internal distances within the body the same. This is an example of a discrete symmetry. So we could include all the orientations in the group, which we designate $O(3)$, the group of orthogonal transformations in 3D. At the same time, it's clear that if we omit the reflection operation, the remaining operations form their own group; we call this the "special" subgroup $SO(3)$.

There is a further caveat for 3D rotations: the generators encapsulate how to get from the identity element to its near neighbors. They don't necessarily capture global features of the group. For instance, there is a redundancy in the rotations, in that a rotation by π around an axis $\hat{\mathbf{n}}$ is equivalent to a rotation by π around the opposite axis $-\hat{\mathbf{n}}$. The group is considered *connected*, but not *simply connected*. In fact, it's *doubly connected*: one can visualize the issue by considering three parameters, the direction of $\hat{\mathbf{n}}$ being specified by angles $\theta \in [0, \pi]$

and $\phi \in [0, 2\pi]$, and the rotation angle by $\psi \in [0, 2\pi]$. The space of rotations can be thought of as a solid sphere with radius π, with $\hat{\mathbf{n}}$ the direction relative to the origin, and ψ the distance from the origin. The non-simple connection arises from the fact that points on the surface of the sphere are equivalent to opposite points on the surface.

Again, as in the case of the permutation group, we need to keep in mind that the group is that of rotations, not of configurations. This is implied in the use of generators: you rotate a little bit from orientation A to B, and then a little more from B to C. If you started from another orientation, you could go through the same set of rotations, but it would take you through a different series of orientations.

7.3 Representations

A group is an abstract concept, and don't have to refer to any specific physical entities. On the other hand, to handle group elements, and especially to do any calculations, we often need to play with concrete mathematical objects which must therefore have the same properties as the group.

For instance, the group of rotations can be represented by a group of rotation matrices which operate on 3-vectors. The two groups have exactly the same behavior, in that the matrix which is the product of two rotation matrices itself represents the resulting rotation. This is an example of an *isomorphism*, in which the elements of two groups G and H have a one-to-one correspondence, and that $g_1 g_2 = g_3$ in for elements in G is true if and only if $h_1 h_2 = h_3$ for the corresponding elements in H.

A *homomorphism* is a slightly looser construction: it preserves the multiplication rule, but loses the requirement that there be a one-to-one correspondence. Representations of groups are often homomorphic (rather than isomorphic) to the original group.

Physicists often don't make strong distinctions between abstract groups and their representations. Unfortunately, we also sometimes forget the distinction between the group representations (*i.e.*, the operators) and the objects on which they operate. Within the group, the rotation operators act not on vectors (which aren't members of the group, after all) but on other rotation operators. And indeed the rotations may not work on vectors, but other objects, which we'll explore in a little while. We'll try to refer to the space of objects as the "representation space" of the representation. The group analysis is independent of those concrete objects which lie outside the group.

7.3.1 Orthogonal matrices

To return to the rotation example, we can associate members of the 3D rotation group with the set of 3×3 orthogonal matrices, which have the property that

$$\mathbf{R}^T \mathbf{R} = I \tag{7.4}$$

A possible representation space is then the linear space of 3D vectors.

This also enables us to appreciate the relationship between rotations and spatial inversion. If we take the determinant of both sides of the above equation, we find that

$$(\det \mathbf{R})^2 = 1 \tag{7.5}$$
$$\det \mathbf{R} = \pm 1 \tag{7.6}$$

A normal rotation has $\det \mathbf{R} = +1$, but a reflection has $\det \mathbf{R} = -1$. Since all the infinitesimal rotations also have $\det \mathbf{R} = +1$, and the determinant of a product of two matrices is simply the product of the two determinants, it's obvious you can't get to a transformation with $\det \mathbf{R} = -1$.

Now let's look at the generators. You've already seen something very much like them in quantum mechanics when dealing with angular momentum. In particular, they do not commute, and the relations are familiar:

$$[\mathbf{J}_a, \mathbf{J}_b] \equiv \mathbf{J}_a\mathbf{J}_b - \mathbf{J}_b\mathbf{J}_a = i\epsilon_{abc}\mathbf{J}_c \tag{7.7}$$

This is sometimes called the "Lie algebra", and it determines the local behavior of the group—in other words, if you have one element, what is the relationship between the nearby elements? Since the continuous symmetries are based on infinitesimal transformations, it shouldn't be surprising that the Lie algebra determines most of the most important properties of the group.

We can derive representations of the $SO(3)$ group by considering its Lie algebra. We define the operators

$$\mathbf{J}^2 = \mathbf{J}_1^2 + \mathbf{J}_2^2 + \mathbf{J}_3^2 \tag{7.8}$$
$$\mathbf{J}_\pm = \mathbf{J}_1 \pm i\mathbf{J}_2 \tag{7.9}$$

The algebra is familiar from quantum mechanics: these are just linear operators, so it's all the same as before, but we recap here. The key observation, however, is that all this follows from the Lie algebra, the knowledge that the generators are Hermitian (which is true for rotations, but not for boosts), and that there is a finite basis set.

The first operator commutes with all the $\mathbf{J}_a$. For instance, consider its commutator with $\mathbf{J}_1$:

$$[\mathbf{J}^2, \mathbf{J}_1] = [\mathbf{J}_2^2, \mathbf{J}_1] + [\mathbf{J}_3^2, \mathbf{J}^1] \tag{7.10}$$
$$= \mathbf{J}_2[\mathbf{J}_2, \mathbf{J}_1] + [\mathbf{J}_2, \mathbf{J}_1]\mathbf{J}_2 + \mathbf{J}_3[\mathbf{J}_3, \mathbf{J}_1] + [\mathbf{J}_3, \mathbf{J}_1]\mathbf{J}_3 \tag{7.11}$$
$$= -i\mathbf{J}_2\mathbf{J}_3 - i\mathbf{J}_3\mathbf{J}_2 + i\mathbf{J}_2\mathbf{J}_3 + i\mathbf{J}_3\mathbf{J}_2 \tag{7.12}$$
$$= 0 \tag{7.13}$$

We call such operators "Casimir operators".

The other two operators are raising and lowering operators. By making them out of $\mathbf{J}_1$ and $\mathbf{J}_2$, we've chosen to use the eigenvectors of $\mathbf{J}_3$ as the basis. Therefore

$$\mathbf{J}_3|m\rangle = m|m\rangle \tag{7.14}$$

We calculate the following commutators:

$$\begin{aligned}
[\mathbf{J}_3, \mathbf{J}_\pm] &= [\mathbf{J}_3, \mathbf{J}_1] \pm i[\mathbf{J}_3, \mathbf{J}_2] && (7.15)\\
&= i\mathbf{J}_2 \mp i^2\mathbf{J}_1 && (7.16)\\
&= i\mathbf{J}_2 \pm \mathbf{J}_1 && (7.17)\\
&= \pm\mathbf{J}_\pm && (7.18)
\end{aligned}$$

This allows us to see the effect of the operators on the eigenvectors:

$$\begin{aligned}
\mathbf{J}_3\mathbf{J}_+|m\rangle - \mathbf{J}_+\mathbf{J}_3|m\rangle &= [\mathbf{J}_3, \mathbf{J}_+]|m\rangle = \mathbf{J}_+|m\rangle && (7.19)\\
\mathbf{J}_3\mathbf{J}_+|m\rangle &= (m+1)\mathbf{J}_+|m\rangle && (7.20)\\
\mathbf{J}_3\mathbf{J}_-|m\rangle - \mathbf{J}_-\mathbf{J}_3|m\rangle &= [\mathbf{J}_3, \mathbf{J}_-]|m\rangle = -\mathbf{J}_-|m\rangle && (7.21)\\
\mathbf{J}_3\mathbf{J}_-|m\rangle &= (m-1)\mathbf{J}_-|m\rangle && (7.22)
\end{aligned}$$

So we find that

$$\mathbf{J}_\pm|m\rangle \propto |m \pm 1\rangle \tag{7.23}$$

Since the basis is finite, there must be both a maximum and minimum m value. The maximum $m_{\max}$ is defined with

$$\begin{aligned}
\mathbf{J}_3|m_{\max}\rangle &= m_{\max}|m_{\max}\rangle && (7.24)\\
\mathbf{J}_+|m_{\max}\rangle &= 0 && (7.25)
\end{aligned}$$

To evaluate the $\mathbf{J}^2$ eigenvalue, we first note that

$$\begin{aligned}
\mathbf{J}_+\mathbf{J}_- &= (\mathbf{J}_1 + i\mathbf{J}_2)(\mathbf{J}_1 - i\mathbf{J}_2) && (7.26)\\
&= \mathbf{J}_1^2 + \mathbf{J}_2^2 + i[\mathbf{J}_2, \mathbf{J}_1] && (7.27)\\
&= \mathbf{J}^2 - \mathbf{J}_3^2 + \mathbf{J}_3 && (7.28)
\end{aligned}$$

and similarly

$$\mathbf{J}_-\mathbf{J}_+ = \mathbf{J}^2 - \mathbf{J}_3^2 - \mathbf{J}_3 \tag{7.29}$$

Then we have

$$\begin{aligned}
\mathbf{J}^2|m_{\max}\rangle &= (\mathbf{J}_3^2 + \mathbf{J}_3 + \mathbf{J}_-\mathbf{J}_+)|m_{\max}\rangle && (7.30)\\
&= (m_{\max}^2 + m_{\max} + 0)|m_{\max}\rangle && (7.31)\\
&= m_{\max}(m_{\max}+1)|m_{\max}\rangle && (7.32)
\end{aligned}$$

From this maximum m, we traverse backwards using successive applications of $\mathbf{J}_-$, stepping by -1 each time. We can show that the $\mathbf{J}^2$ eigenvalue applies to all the other eigenvectors, since $\mathbf{J}^2$ commutes with all the $\mathbf{J}_a$, and therefore all powers of $\mathbf{J}_-$:

$$\begin{aligned}
\mathbf{J}^2\mathbf{J}_-^n|m_{\max}\rangle &= \mathbf{J}_-^n\mathbf{J}^2|m_{\max}\rangle && (7.33)\\
&= m_{\max}(m_{\max}+1)\mathbf{J}_-^n|m_{\max}\rangle && (7.34)
\end{aligned}$$

Since we've now shown the role of $m_{\max}$, we can revert to the familiar terminology and call it j.

If we take all the $|m\rangle$ to be properly normalized, then we can find the coefficients left by the $\mathbf{J}_\pm$ operators:

$$\langle m|\mathbf{J}_+^\dagger \mathbf{J}_+|m\rangle = \langle m|\mathbf{J}_-\mathbf{J}_+|m\rangle = \langle m|(\mathbf{J}^2 - \mathbf{J}_3^2 - \mathbf{J}_3)|m\rangle = j(j+1) - m(m+1) \quad (7.35)$$

$$\langle m|\mathbf{J}_-^\dagger \mathbf{J}_-|m\rangle = \langle m|\mathbf{J}_+\mathbf{J}_-|m\rangle = \langle m|(\mathbf{J}^2 - \mathbf{J}_3^2 + \mathbf{J}_3)|m\rangle = j(j+1) - m(m-1) \quad (7.36)$$

and therefore, summarizing,

$$\mathbf{J}_\pm|m\rangle = [j(j+1) - m(m\pm 1)]^{1/2}|m\pm 1\rangle \quad (7.37)$$

At the minimum m value, we have

$$\begin{aligned}
\mathbf{J}_3|m_{\min}\rangle &= m_{\min}|m_{\min}\rangle && (7.38)\\
\mathbf{J}_-|m_{\min}\rangle &= 0 && (7.39)\\
0 &= \langle m_{\min}|\mathbf{J}_-^\dagger \mathbf{J}_-|m_{\min}\rangle && (7.40)\\
&= \langle m_{\min}|\mathbf{J}_+\mathbf{J}_-|m_{\min}\rangle && (7.41)\\
&= \langle m_{\min}|(\mathbf{J}^2 - \mathbf{J}_3^2 + \mathbf{J}_3)|m_{\min}\rangle && (7.42)\\
&= j(j+1) - m_{\min}(m_{\min} - 1) && (7.43)
\end{aligned}$$

If we take all the $|m\rangle$ to be properly normalized, then we find that $m_{\min} = -j$. For the series to be finite, j must be either an integer or half-integer.

So we get a set of eigenvectors (members of the representation space) with the eigenvalues j and m where

$$\begin{aligned}
\mathbf{J}^2|jm\rangle &= j(j+1)|jm\rangle && (7.44)\\
\mathbf{J}_3|jm\rangle &= m|jm\rangle && (7.45)\\
\mathbf{J}\pm|jm\rangle &= \sqrt{j(j+1) - m(m\pm 1)}|j, m\pm 1\rangle && (7.46)\\
j &= 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \cdots && (7.47)\\
m &= -j, -j+1, \cdots, j-1, j && (7.48)
\end{aligned}$$

This looks like quantum mechanics, but we're still dealing with classical physics. The important thing to remember here is that whatever happens to your physical intuition with quantum mechanics, it doesn't modify mathematics. (Another way of saying this is that the origin of this algebra is not quantum mechanical, but geometrical.)

The representations are characterized by the eigenvalue j, with the corresponding $2j+1$ eigenvectors spanning the representation space. Obviously most of these representations are homomorphic to the rotation group, rather than isomorphic. The simplest (trivial) example is $j = 0$, which maps all the rotations onto the identity element. The generators in this case (and in this representation basis, *not* the Cartesian basis we started with) are $\mathbf{J}_k = 0$ for all k, so they trivially satisfy the Lie algebra. Only the $j = 1$ representation is isomorphic to the rotation group itself.

We can check that the $j = 1$ representation behaves as we expect. To do so, we should be explicit about bases, because we've actually given the generators in two.

First, we used the Cartesian basis, which is convenient for its geometric roots. In this basis,

$$\mathbf{J}_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{7.49}$$

and the rotation matrix is

$$\mathbf{R}_3(\alpha) = e^{-i\alpha\mathbf{J}_3} = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7.50}$$

In the "canonical" basis of the $j = 1$ representation, the corresponding generator is

$$\mathbf{J}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{7.51}$$

so for a finite rotation this becomes

$$\mathbf{R}_3 = e^{-i\alpha\mathbf{J}_3} = 1 - i\sin\alpha\mathbf{J}_3 + (\cos\alpha - 1)\mathbf{J}_3^2 \tag{7.52}$$

Since we can see that $\mathbf{J}_3^2$ is "almost" an identity,

$$\mathbf{J}_3^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7.53}$$

we can write down the rotation matrix directly:

$$\mathbf{R}_3 = \begin{pmatrix} 1 - i\sin\alpha + \cos\alpha - 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 + i\sin\alpha + \cos\alpha - 1 \end{pmatrix} = \begin{pmatrix} e^{-i\alpha} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\alpha} \end{pmatrix} \tag{7.54}$$

Is this really a rotation matrix around the z axis? Let's examine how the matrix affects an appropriate representation space, such as the $l = 1$ spherical harmonics.

$$Y_1'^1(\theta,\phi) = Y_1{}^1(\theta,\phi)e^{i\alpha} = \frac{1}{2}\sqrt{\frac{3}{2\pi}}\sin\theta e^{-i(\phi-\alpha)} \tag{7.55}$$

$$Y_1'^0(\theta,\phi) = Y_1{}^0(\theta,\phi) = \frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\theta \tag{7.56}$$

$$Y_1'^{-1}(\theta,\phi) = Y_1{}^{-1}(\theta,\phi)e^{-i\alpha} = -\frac{1}{2}\sqrt{\frac{3}{2\pi}}\sin\theta e^{i(\phi-\alpha)} \tag{7.57}$$

$$\tag{7.58}$$

Since the xy dependence in these spherical harmonics is in the ϕ exponential, it appears the rotation matrix has rotated the basis functions by a consistent angle α.

7.3.2 Spinor representation

What of the $j = 1/2$ representation? This has 2 basis vectors with $m = \pm 1/2$, which we designate $|+\rangle$ and $|-\rangle$. Following the same procedure as before, we make $\mathbf{J}_3$ out of the eigenvalues:

$$\mathbf{J}_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.59}$$

The action of the raising and lower operators is

$$\mathbf{J}_+|-\rangle = |+\rangle \tag{7.60}$$
$$\mathbf{J}_-|+\rangle = |-\rangle \tag{7.61}$$

so the operators are

$$\mathbf{J}_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{7.62}$$
$$\mathbf{J}_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{7.63}$$
$$\tag{7.64}$$

which allow us to write down $\mathbf{J}_1$ and $\mathbf{J}_2$:

$$\mathbf{J}_1 = \frac{1}{2}(\mathbf{J}_+ + \mathbf{J}_-) = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{7.65}$$
$$\mathbf{J}_2 = -\frac{i}{2}(\mathbf{J}_+ - \mathbf{J}_-) = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{7.66}$$

In short,

$$\mathbf{J}_i = \frac{\sigma_i}{2} \tag{7.67}$$

where σ_i are the Pauli matrices.

A finite rotation around the z axis is

$$e^{i\theta\mathbf{J}_3} = 1 - i\theta\mathbf{J}_3 - \frac{\theta^2}{2!}\mathbf{J}_3^2 + i\frac{\theta^3}{3!}\mathbf{J}_3^3 + \cdots \tag{7.68}$$
$$= 1 - i\frac{\theta}{2}\sigma_3 - \frac{\theta^2}{2^2 2!}\sigma_3^2 + i\frac{\theta^3}{2^3 3!}\sigma_3^3 + \cdots \tag{7.69}$$
$$= \cos\frac{\theta}{2} - i\sigma_3 \sin\frac{\theta}{2} \tag{7.70}$$

taking advantage of the fact that $\sigma_j^2 = I$. In this case, we see that it takes a rotation through 4π to get back to the identity. A full rotation through 2π, on the other hand, gets you to -1.

It is worth pausing to consider what these results mean. The Pauli matrices are themselves a basis set for traceless, Hermitian 2×2 matrices, and the exponential of these matrices yield 2×2 unitary matrices. In fact, they yield a subset of unitary matrices we denote $SU(2)$:

the group of unitary matrices with determinant 1. This group has the same Lie algebra as $SO(3)$; after all, that's how we got it in the first place. This means that the local behavior of moving from one element of its representation space to another is identical to that of 3D rotations.

Moreover, $SU(2)$ is simply connected, unlike $SO(3)$. The easiest way to see this is to consider a general element of $SU(2)$ in the form

$$\mathbf{U} = \begin{pmatrix} a - ib & -c - id \\ -c + id & a + ib \end{pmatrix} \tag{7.71}$$

where a, b, c, and d are real parameters. One can verify that it's unitary

$$\begin{aligned} \mathbf{U}^\dagger\mathbf{U} &= \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}\begin{pmatrix} a - ib & -c - id \\ -c + id & a + ib \end{pmatrix} && (7.72)\\ &= \begin{pmatrix} a^2 + b^2 + c^2 + d^2 & 0 \\ 0 & a^2 + b^2 + c^2 + d^2 \end{pmatrix} && (7.73)\\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} && (7.74) \end{aligned}$$

with the constraint that

$$\det \mathbf{U} = a^2 + b^2 + c^2 + d^2 = 1 \tag{7.75}$$

The group manifold is therefore a unit (Euclidean) 4-sphere, with no equivalent points. Because $SU(2)$ is simply connected, any element can be written uniquely as an exponential of generators. It is actually a "covering group" of $SO(3)$, with identical local behavior but "unrolling" the latter's double connection into the double cover.

7.3.3 Spinor representation space

(Based on Steane's chapter on spinors)

Now, we might ask, what is the representation space of the $j = 1/2$ representation? In this case, we only have a representation in the canonical basis, whereas for $j = 1$ we had both canonical and and convenient Cartesian bases, the latter of which provided a more ready intuition as to what was going on.

On the other hand, we should have some sense of a rotation, so we expect that there can be some relationship with Cartesian components. Indeed, since we expect the 2-dimensional representation space to have complex coefficients, we certainly have enough degrees of freedom to represent a 3-vector—and more.

We can write the spinors in terms of 4 real parameters r, θ, ϕ, and α. The first 3 parameters define a usual 3-vector in polar coordinates. The last parameter then encodes an additional orientation, like a little flag flying from a 3-vector flagpole. The actual definition of the orientation doesn't mean much at this point, because the rotations of $SO(3)$ don't affect it.

The 2-component spinor is then

$$\mathsf{s} = \begin{pmatrix} a \\ b \end{pmatrix} = se^{-i\alpha/2}\begin{pmatrix} \cos\frac{\theta}{2}e^{-i\phi/2} \\ \sin\frac{\theta}{2}e^{i\phi/2} \end{pmatrix} \tag{7.76}$$

where

$$s^2 = |a|^2 + |b|^2 = r \tag{7.77}$$

The 3-vector components can be recovered from a and b as follows:

$$\begin{aligned} x &= ab^* + ba^* = \mathsf{s}^\dagger \sigma_x \mathsf{s} && (7.78)\\ y &= i(ab^* - ba^*) = \mathsf{s}^\dagger \sigma_y \mathsf{s} && (7.79)\\ z &= |a|^2 - |b|^2 = \mathsf{s}^\dagger \sigma_z \mathsf{s} && (7.80) \end{aligned}$$

We can confirm that these are rotated as expected by calculating the finite rotation matrices:

$$\begin{aligned} \mathbf{R}_1(\beta) &= e^{-i\beta \mathbf{J}_1} = e^{-i\beta\sigma_1/2} = \begin{pmatrix} \cos\frac{\beta}{2} & -i\sin\frac{\beta}{2} \\ -i\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix} && (7.81)\\ \mathbf{R}_2(\beta) &= e^{-i\beta \mathbf{J}_2} = e^{-i\beta\sigma_2/2} = \begin{pmatrix} \cos\frac{\beta}{2} & -\sin\frac{\beta}{2} \\ \sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix} && (7.82)\\ \mathbf{R}_3(\beta) &= e^{-i\beta \mathbf{J}_3} = e^{-i\beta\sigma_3/2} = \begin{pmatrix} e^{-i\beta/2} & 0 \\ 0 & e^{i\beta/2} \end{pmatrix} && (7.83) \end{aligned}$$

The rotation around z is easiest:

$$\begin{aligned} \mathbf{R}_3(\beta)\mathsf{s} &= s e^{-i\alpha/2} \begin{pmatrix} e^{-i\beta/2} & 0 \\ 0 & e^{i\beta/2} \end{pmatrix} \begin{pmatrix} \cos\frac{\theta}{2} e^{-i\phi/2} \\ \sin\frac{\theta}{2} e^{i\phi/2} \end{pmatrix} && (7.84)\\ &= s e^{-i\alpha/2} \begin{pmatrix} \cos\frac{\theta}{2} e^{-i(\phi+\beta)/2} \\ \sin\frac{\theta}{2} e^{i(\phi+\beta)/2} \end{pmatrix} && (7.85) \end{aligned}$$

which, as expected, adds to the azimuthal angle.

To simplify checking the rotation around y, we just consider spinors in the xz plane, *i.e.*, setting $\phi = 0$. The rotation then becomes

$$\begin{aligned} \mathbf{R}_2(\beta)\mathsf{s} &= s e^{-i\alpha/2} \begin{pmatrix} \cos\frac{\beta}{2} & -\sin\frac{\beta}{2} \\ \sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix} \begin{pmatrix} \cos\frac{\theta}{2} e^{-i\phi/2} \\ \sin\frac{\theta}{2} e^{i\phi/2} \end{pmatrix} && (7.86)\\ &\to s e^{-i\alpha/2} \begin{pmatrix} \cos\frac{\beta}{2} & -\sin\frac{\beta}{2} \\ \sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix} \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix} && (7.87)\\ &= s e^{-i\alpha/2} \begin{pmatrix} \cos\frac{\beta}{2}\cos\frac{\theta}{2} - \sin\frac{\beta}{2}\sin\frac{\theta}{2} \\ \sin\frac{\beta}{2}\cos\frac{\theta}{2} + \cos\frac{\beta}{2}\sin\frac{\theta}{2} \end{pmatrix} && (7.88)\\ &= s e^{-i\alpha/2} \begin{pmatrix} \cos\frac{\theta+\beta}{2} \\ \sin\frac{\theta+\beta}{2} \end{pmatrix} && (7.89) \end{aligned}$$

This increases the polar angle, which is what we expect.

For rotating around x, we check at $\phi = \pi/2$, *i.e.*, in the yz plane.

$$\begin{aligned}
\mathbf{R}_1(\beta)\mathsf{s} &= se^{-i\alpha/2}\begin{pmatrix} \cos\frac{\beta}{2} & -i\sin\frac{\beta}{2} \\ -i\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix}\begin{pmatrix} \cos\frac{\theta}{2}e^{-i\phi/2} \\ \sin\frac{\theta}{2}e^{i\phi/2} \end{pmatrix} && (7.90)\\
&\to \frac{se^{-i\alpha/2}}{\sqrt{2}}\begin{pmatrix} \cos\frac{\beta}{2} & -i\sin\frac{\beta}{2} \\ -i\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix}\begin{pmatrix} (1-i)\cos\frac{\theta}{2} \\ (1+i)\sin\frac{\theta}{2} \end{pmatrix} && (7.91)\\
&= \frac{se^{-i\alpha/2}}{\sqrt{2}}\begin{pmatrix} (1-i)\cos\frac{\beta}{2}\cos\frac{\theta}{2} - (1+i)i\sin\frac{\beta}{2}\sin\frac{\theta}{2} \\ -i(1-i)\sin\frac{\beta}{2}\cos\frac{\theta}{2} + (1+i)\cos\frac{\beta}{2}\sin\frac{\theta}{2} \end{pmatrix} && (7.92)\\
&= \frac{se^{-i\alpha/2}}{\sqrt{2}}\begin{pmatrix} (\cos\frac{\beta}{2}\cos\frac{\theta}{2} + \sin\frac{\beta}{2}\sin\frac{\theta}{2}) + i(-\cos\frac{\beta}{2}\cos\frac{\theta}{2} - \sin\frac{\beta}{2}\sin\frac{\theta}{2}) \\ (-\sin\frac{\beta}{2}\cos\frac{\theta}{2} + \cos\frac{\beta}{2}\sin\frac{\theta}{2}) + i(-\sin\frac{\beta}{2}\cos\frac{\theta}{2} + \cos\frac{\beta}{2}\sin\frac{\theta}{2}) \end{pmatrix} && (7.93)\\
&= \frac{se^{-i\alpha/2}}{\sqrt{2}}\begin{pmatrix} (1-i)\cos\frac{\theta-\beta}{2} \\ (1+i)\sin\frac{\theta-\beta}{2} \end{pmatrix} && (7.94)\\
&= se^{-i\alpha/2}\begin{pmatrix} e^{-i\pi/4}\cos\frac{\theta-\beta}{2} \\ e^{i\pi/4}\sin\frac{\theta-\beta}{2} \end{pmatrix} && (7.95)
\end{aligned}$$

In this case, the rotation is in the opposite sense of increasing the polar angle, so β is subtracted from θ.

7.3.4 Spinor representation with matrices

There is actually no particular reason the representation space has go be built out of column vectors. The requirement is that the space has to be linear, and that one can obtain a scalar (the norm) out of it.

In fact, it's (arguably) easier to represent spinors with 2×2 matrices. Associate each 3-vector $\mathbf{x}$ with a 2×2 unitary matrix as follows:

$$\mathbf{X} = x^i\sigma_i = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix} \tag{7.96}$$

The norm of the object is

$$|\mathbf{X}| = -\det\mathbf{X} = -z^2 - (x+iy)(x-iy) = -x^2 - y^2 - z^2 \tag{7.97}$$

Now, with each transformation $\mathbf{U}$, associate the similarity transformation:

$$\mathbf{X}' = \mathbf{U}\mathbf{X}\mathbf{U}^\dagger \tag{7.98}$$

Since $\mathbf{U} \in SU(2)$, the determinant of $\mathbf{U}$ is $+1$. We then have the relationship

$$\det\mathbf{X}' = \det\mathbf{X} \tag{7.99}$$

so the norm is clearly preserved by the transformation.

As a side note (which is actually rather important from other perspectives), we also see $\mathbf{U}$ and $-\mathbf{U}$ results in the same transformation. This is another manifestation of the double cover of $SU(2)$ over $SO(3)$.

Let's try out a rotation:

$$\begin{aligned}
\mathbf{R_3}(\beta)\mathbf{X}\mathbf{R_3}^\dagger(\beta) &= \begin{pmatrix} e^{-i\beta/2} & 0 \\ 0 & e^{i\beta/2} \end{pmatrix}\begin{pmatrix} z & x-iy \\ x+iy & -z \end{pmatrix}\begin{pmatrix} e^{i\beta/2} & 0 \\ 0 & e^{-i\beta/2} \end{pmatrix} && (7.100)\\
&= \begin{pmatrix} ze^{-i\beta/2} & (x-iy)e^{-i\beta/2} \\ (x+iy)e^{i\beta/2} & -ze^{i\beta/2} \end{pmatrix}\begin{pmatrix} e^{i\beta/2} & 0 \\ 0 & e^{-i\beta/2} \end{pmatrix} && (7.101)\\
&= \begin{pmatrix} z & (x-iy)e^{-i\beta} \\ (x+iy)e^{i\beta} & -z \end{pmatrix} && (7.102)
\end{aligned}$$

so we can see that z is unaffected, and the angle of $(x+iy)$ has been increased by β. One can test the other rotations as well.

Is there a relationship between these two spaces? We can form a 2×2 matrix from the column spinors by taking the outer product:

$$\begin{aligned}
2\mathsf{s}\mathsf{s}^\dagger &= 2s^2\begin{pmatrix} \cos\frac{\theta}{2}e^{-i\phi/2} \\ \sin\frac{\theta}{2}e^{i\phi/2} \end{pmatrix}\begin{pmatrix} \cos\frac{\theta}{2}e^{i\phi/2} & \sin\frac{\theta}{2}e^{-i\phi/2} \end{pmatrix} && (7.103)\\
&= 2r\begin{pmatrix} \cos^2\frac{\theta}{2} & \sin\frac{\theta}{2}\cos\frac{\theta}{2}e^{-i\phi} \\ \sin\frac{\theta}{2}\cos\frac{\theta}{2}e^{i\phi} & \sin^2\frac{\theta}{2} \end{pmatrix} && (7.104)\\
&= r\begin{pmatrix} 1+\cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & 1-\cos\theta \end{pmatrix} && (7.105)
\end{aligned}$$

from which we see that

$$\mathbf{X} = 2\mathsf{s}\mathsf{s}^\dagger - r\mathbf{1} \tag{7.106}$$

The transformation then follows:

$$\begin{aligned}
\mathbf{X}' &= \mathbf{U}\mathbf{X}\mathbf{U}^\dagger && (7.107)\\
&= 2\mathbf{U}\mathsf{s}\mathsf{s}^\dagger\mathbf{U}^\dagger - r\mathbf{U}\mathbf{U}^\dagger && (7.108)\\
&= 2\mathsf{s}'\mathsf{s}'^\dagger - r\mathbf{1} && (7.109)
\end{aligned}$$

We can then see that the transformation of $\mathbf{X}$ is closely related to the transformation of s.

7.3.5 Higher-spin representations

Representations with higher j values can be obtained through the same algebra, but they can also be obtained by combining lower-j representations. You're already familiar with the procedure from quantum mechanics, where it was called "addition of angular momenta."

As an example, you can take a direct product of two $j = 1/2$ representations. Elements of the direct product representation space have the form of the tensor product

$$\xi^a\zeta^b \tag{7.110}$$

where a and b are indices taking on values 1 or 2. The space is designated $\frac{1}{2}\otimes\frac{1}{2}$.

The direct product space can be broken up into subspaces with different j values (in group terminology, we are reducing the space into its irreducible representations, which are the

representations with a well-defined j). As expected, you get a $j = 0$ space and a $j = 1$ space. The $j = 0$ space transforms as a scalar, while the $j = 1$ space, with its 3 basis elements, transforms as a 3-vector. The matrix operators in the new basis are then block diagonal, with operators for the different j subspaces in each block. This combined representation is called a *direct sum* representation, and in this case is written $0 \oplus 1$. In fact, you have already seen something like this in action, in the relationship between the two spinor representation spaces; some authors make the analogy that a spinor is a sort of "square root" of a vector.

Chapter 8

Lorentz group

Having spent some time looking (albeit at some distance) at the properties of the rotation group $SO(3)$, we turn our attention to the full Lorentz group $SO(3,1)$. The "1" indicates the additional, time-like dimension.

The full Lorentz group consists of transformations which preserve the norm of a 4-vector. The elements $\Lambda^{\mu}{}_{\nu}$ satisfy the relation

$$g_{\mu\nu} = \Lambda^{\kappa}{}_{\mu}\Lambda^{\lambda}{}_{\nu}g_{\kappa\lambda} \tag{8.1}$$

where g is the metric tensor. In matrix form, this expression is

$$g = \Lambda^{T} g \Lambda \tag{8.2}$$

Rotations clearly form a subgroup of the full Lorentz group, since the Minkowski metric is invariant with respect to rotations in 3-space.

From the matrix equation, it's clear that

$$\det \Lambda = \pm 1 \tag{8.3}$$

For rotations, this meant that the group $O(3)$ divided into two parts, linked by the parity transformation. The generators traversed the two parts, but couldn't get from one to the other without the parity transformation. One sees this as well in the Lorentz group.

If we evaluate the relationship for $\mu = \nu = 0$, we get

$$\begin{aligned} g_{00} &= \Lambda^{\kappa}{}_{0}\Lambda^{\lambda}{}_{0}g_{\kappa\lambda} & (8.4)\\ -1 &= -(\Lambda^{0}{}_{0})^2 + \sum_{i=1}^{3}(\Lambda^{i}{}_{0})^2 & (8.5)\\ \Lambda^{0}{}_{0} &= \pm\sqrt{1 + \sum_{i=1}^{3}(\Lambda^{i}{}_{0})^2} & (8.6) \end{aligned}$$

There is therefore a gap between elements with positive and negative $\Lambda^{0}{}_{0}$, which, again, cannot be traversed using infinitesimal transformations.

There are therefore 4 divisions of the full Lorentz group. Of these, we'll be concerned with the division which forms a subgroup, *i.e.*, contains the identity element. This is called the "orthochronous" (preserving the normal time direction) Lorentz group:

$$\det \Lambda = +1 \tag{8.7}$$
$$\Lambda^0{}_0 \geq 1 \tag{8.8}$$

The subgroup is denoted $SO(3,1)^{\uparrow}_{+}$, where the superscript indicates it's the orthochronous division, and the subscript the parity. Since we'll pretty much just talk about this subgroup, we'll drop the decorations and call it $SO(3,1)$ by default; if we need another division, we'll mention it specifically.

8.1 Commutators

As with $SO(3)$, we start by looking at the Lie algebra, which tells us about the local behavior of the group's transformations. The commutators are

$$[\mathbf{J}_a, \mathbf{J}_b] = i\epsilon_{abc}\mathbf{J}_c \tag{8.9}$$
$$[\mathbf{K}_a, \mathbf{J}_b] = i\epsilon_{abc}\mathbf{K}_c \tag{8.10}$$
$$[\mathbf{K}_a, \mathbf{K}_b] = -i\epsilon_{abc}\mathbf{J}_c \tag{8.11}$$

The first commutator is simply the one for rotations. The third commutator shows that the difference in the order of two non-aligned Lorentz boosts is a rotation. We've already seen the physical effect of this non-zero commutator. (It's worth noting that since two non-aligned boosts aren't in general equivalent to a single boost, but a boost and a rotation, boosts don't by themselves form a group.)

The Lie algebra is already suggestive of rotations. It can be made even more suggestive by combining the operators:

$$\mathbf{M}_a = (\mathbf{J}_a + i\mathbf{K}_a)/2 \tag{8.12}$$
$$\mathbf{N}_a = (\mathbf{J}_a - i\mathbf{K}_a)/2 \tag{8.13}$$

The commutators then become

$$[\mathbf{M}_a, \mathbf{M}_b] = i\epsilon_{abc}\mathbf{M}_c \tag{8.14}$$
$$[\mathbf{N}_a, \mathbf{N}_b] = i\epsilon_{abc}\mathbf{N}_c \tag{8.15}$$
$$[\mathbf{M}_a, \mathbf{N}_b] = 0 \tag{8.16}$$

So we now have two, disjoint $SU(2)$ algebras. We can think of this as the direct product $SU(2) \times SU(2)$.

We saw before that $SU(2)$ was the cover group for $SO(3)$, in that it preserved the local relationships between the infinitesimal transformations, but had better global properties in that its group manifold was simply connected. Similarly, there is a (double) cover group for the Lorentz group, the "special linear group" $SL(2,\mathbb{C})$ of 2×2 matrices with determinant

+1. The group $SU(2)$ is obviously a subgroup of $SL(2,\mathbb{C})$, and indeed $SL(2,\mathbb{C})$ is the "complexified" direct product $SU(2) \times SU(2)$. And as with $SU(2)$ compared with $SO(3)$, it has the same local behavior but also better global behavior, *i.e.*, it is simply connected, unlike $SO(3,1)$.

The direct product structure allows us to enumerate the representations. Two Casimir operators immediately suggest themselves: $\mathbf{M}^2$, with eigenvalues $m(m+1)$, and $\mathbf{N}^2$, with eigenvalues $n(n+1)$. Both m and n are nonnegative integers or half-integers. So we write down the pair (m,n).

8.2 Fundamental representations

The structure $SU(2) \times SU(2)$ suggests looking at spin-1/2 representations. Let's start as we had with the rotation group, with one addition:

$$\mathbf{X} = x^\mu \sigma_\mu = \begin{pmatrix} ct+z & x-iy \\ x+iy & ct-z \end{pmatrix} \tag{8.17}$$

The components can be extracted using the trace

$$x^\mu = \frac{1}{2}\mathrm{tr}(\mathbf{X}\sigma_\mu) \tag{8.18}$$

The Lorentz transformation then takes the form

$$\mathbf{X}' = \mathbf{A}\mathbf{X}\mathbf{A}^\dagger \tag{8.19}$$

where $\mathbf{A} \in SL(2,\mathbb{C})$ is the matrix corresponding to the Lorentz transformation of x^μ. An elegant way of summarizing this is that

$$\mathbf{A}(x^\mu \sigma_\mu)\mathbf{A}^\dagger = (\Lambda x^\mu)\sigma_\mu \tag{8.20}$$

We'd like to write the vector as a direct product of spinors, $\mathbf{X} = \xi\xi^\dagger$. This tells us that the appropriate Lorentz transformation matrix is $\mathbf{A}$, since

$$\mathbf{X}' = \mathbf{A}\mathbf{X}\mathbf{A}^\dagger = \mathbf{A}\xi\xi^\dagger\mathbf{A}^\dagger = (\mathbf{A}\xi)(\mathbf{A}\xi)^\dagger = \xi'\xi'^\dagger \tag{8.21}$$

So a Lorentz transform Λ on a vector corresponds with the matrix $\mathbf{A}$ on a spinor.

This seems simple enough, but if you look in different texts on the subject, you'll see different conventions at work. The main difference is which transformation is taken as a starting point for deriving all the other transformation rules. We'll use a convention which has become fairly conventional in the larger physics community. Toward that end, we define

$$\psi_a \rightarrow \psi'_a = A_a{}^b \psi_b \tag{8.22}$$

$$\psi^a \rightarrow \psi'^a = \psi^b (A^{-1})_b{}^a \tag{8.23}$$

Notice that in this convention, it is the covariant spinor which transforms with the matrix $\mathbf{A}$, and the contravariant spinor with its inverse. In the end, all these differences reflect manipulations of some initial matrix; once the initial matrix is defined for a given transformation, the other forms follow.

First, let's find an invariant tensor that performs as a metric. It turns out that we can use the asymmetric tensor ε^{ab}:

$$\varepsilon^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{8.24}$$

The tensor ε plays a similar role to the metric g, though there are some subtleties about how it raises and lowers indices because it is, unlike g, antisymmetric:

$$\varepsilon_{ab} = -\varepsilon^{ab} \tag{8.25}$$

and as result,

$$\varepsilon^{ab}\psi_b \neq \psi_b\varepsilon^{ba} \tag{8.26}$$

A common convention is that ε raises/lowers from the left:

$$\psi^a = \varepsilon^{ab}\psi_b \tag{8.27}$$
$$\psi_a = \varepsilon_{ab}\psi^b \tag{8.28}$$

For matrix indices,

$$\varepsilon^{ab}A_b{}^c\varepsilon_{cd} = A^a{}_d \tag{8.29}$$

So if you write down matrix elements in one form, with indices in one configuration (such as $A_b{}^c$), but then need to use the matrix with indices in another configuration (such as $A^b{}_c$), you will need to use ε's to raise and lower to get the form you need.

In that spirit, the elements of the matrix form of ε_{cd} are defined by

$$\varepsilon^{ab}\varepsilon_{bc} = \varepsilon^a{}_c = \delta^a_c \tag{8.30}$$

as appropriate for a metric-like tensor, so

$$\varepsilon_{ab} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{8.31}$$

The two tensors clearly act as inverses of one another.

We now have two ways of writing the transformation of ψ^a. The first is the definition itself, in terms of $\mathbf{A}^{-1}$. The second is to use ε to relate it to the covariant transformation:

$$\psi'^a = \varepsilon^{ab}\psi'_b \tag{8.32}$$
$$= \varepsilon^{ab}A_b{}^c\psi_c \tag{8.33}$$
$$= \varepsilon^{ab}A_b{}^c\varepsilon_{cd}\psi^d \tag{8.34}$$

Comparison with the contravariant transformation rules yields

$$(A^{-1})_b{}^a = \varepsilon^{ac}A_c{}^d\varepsilon_{db} \tag{8.35}$$

Since this is in the form of a similarity transformation $\mathbf{SAS}^{-1}$, it is clear the covariant and contravariant spinor representations are equivalent.

Now let's see what happens when we take the complex conjugate of these transformations. We define

$$\bar{\chi}_{\dot{a}} = (\chi_a)^* \tag{8.36}$$
$$\bar{\chi}^{\dot{a}} = (\chi^a)^* \tag{8.37}$$

where we've dotted the indices to anticipate that we'll need to account for them separately from undotted ones. The transformation of the conjugate contravariant spinor follows:

$$\bar{\chi'}^{\dot{a}} = (\chi'^a)^* \tag{8.38}$$
$$= (\chi^b (A^{-1})_b{}^a)^* \tag{8.39}$$
$$= (\chi^b)^* ((A^{-1})_b{}^a)^* \tag{8.40}$$

In order to simplify the matrix part, we start from equivalence we found above:

$$(A^{-1})_b{}^a = \varepsilon^{ac} A_c{}^d \varepsilon_{db} = A^a{}_b \tag{8.41}$$
$$((A^{-1})_b{}^a)^* = (A^a{}_d)^* = (A^*)^{\dot{a}}{}_{\dot{b}} \tag{8.42}$$

So we can now write the conjugate contravariant spinor transformation as

$$\bar{\chi'}^{\dot{a}} = (A^*)^{\dot{a}}{}_{\dot{b}} \bar{\chi}^{\dot{b}} \tag{8.43}$$

We then transform the conjugate covariant spinor:

$$\bar{\chi'}_{\dot{a}} = (\chi'_a)^* \tag{8.44}$$
$$= (A_a{}^b \chi_b)^* \tag{8.45}$$
$$= (A_a{}^b)^* \bar{\chi}_{\dot{b}} \tag{8.46}$$

To find the matrix elements, we use the equivalence again:

$$(A^{-1})_b{}^a = \varepsilon^{ac} A_c{}^d \varepsilon_{db} \tag{8.47}$$
$$((A^{-1})_b{}^a)^* = \varepsilon^{ac} (A_c{}^d)^* \varepsilon_{db} \tag{8.48}$$
$$\varepsilon_{ma} (A^{*-1})_b{}^a \varepsilon^{bn} = \varepsilon_{ma} \varepsilon^{ac} (A_c{}^d)^* \varepsilon_{db} \varepsilon^{bn} \tag{8.49}$$
$$\varepsilon^{nb} (A^{*-1})_b{}^a \varepsilon_{ma} = (A^*)_m{}^n \tag{8.50}$$
$$(A^{*-1})^n{}_m = (A^*)_m{}^n \tag{8.51}$$

We have used the matrix form of ε liberally to evaluate its transpose and complex conjugate. The last relationship can be written in matrix notation as

$$\mathbf{A}^{*-1} = \mathbf{A}^\dagger \tag{8.52}$$

and the conjugate covariant spinor transformation is

$$\bar{\chi'}_{\dot{a}} = \bar{\chi}_{\dot{b}} (A^{*-1})^{\dot{b}}{}_{\dot{a}} \tag{8.53}$$

In summary,

$$\psi_a \rightarrow \psi'_a = A_a{}^b \psi_b \tag{8.54}$$
$$\psi^a \rightarrow \psi'^a = \psi^b (A^{-1})_b{}^a \tag{8.55}$$
$$\bar{\chi}^{\dot{a}} \rightarrow (A^*)^{\dot{a}}{}_{\dot{b}} \bar{\chi}^{\dot{b}} \tag{8.56}$$
$$\bar{\chi}_{\dot{a}} \rightarrow \bar{\chi}_{\dot{b}} (A^{*-1})^{\dot{b}}{}_{\dot{a}} \tag{8.57}$$

As we saw earlier, the first two of these transformation rules are equivalent, as are the last two. But it turns out that the first two are *not* equivalent with the latter two.

To see that, let's write down expressions for $\mathbf{A}$. We use the same rotation generators as before:

$$\mathbf{J}_a = \frac{1}{2}\sigma_a \tag{8.58}$$

For boosts, we can write down

$$\mathbf{K}_a = \frac{i}{2}\sigma_a \tag{8.59}$$

It is straightforward to verify that this works for boosts in the z direction.

$$\mathbf{A} = e^{\eta\sigma_3/2} = \cosh\frac{\eta}{2} + \sigma_3 \sinh\frac{\eta}{2} = \begin{pmatrix} e^{\eta/2} & 0 \\ 0 & e^{-\eta/2} \end{pmatrix} \tag{8.60}$$

Then we have $\mathbf{A}^\dagger = \mathbf{A}$, and

$$\mathbf{X}' = \mathbf{A}\mathbf{X}\mathbf{A}^\dagger \tag{8.61}$$
$$= \begin{pmatrix} e^{\eta/2} & 0 \\ 0 & e^{-\eta/2} \end{pmatrix} \begin{pmatrix} ct+z & x-iy \\ x+iy & ct-z \end{pmatrix} \begin{pmatrix} e^{\eta/2} & 0 \\ 0 & e^{-\eta/2} \end{pmatrix} \tag{8.62}$$
$$= \begin{pmatrix} (ct+z)e^{\eta/2} & (x-iy)e^{\eta/2} \\ (x+iy)e^{-\eta/2} & (ct-z)e^{-\eta/2} \end{pmatrix} \begin{pmatrix} e^{\eta/2} & 0 \\ 0 & e^{-\eta/2} \end{pmatrix} \tag{8.63}$$
$$= \begin{pmatrix} (ct+z)e^{\eta} & x-iy \\ x+iy & (ct-z)e^{-\eta} \end{pmatrix} \tag{8.64}$$
$$\tag{8.65}$$

The transformed coordinates can be evaluated using the trace:

$$ct' = \frac{1}{2}\mathrm{tr}\mathbf{X}' = \frac{1}{2}(ct(e^\eta + e^{-\eta}) + z(e^\eta - e^{-\eta})) = ct\cosh\eta + z\sinh\eta \tag{8.66}$$
$$z' = \frac{1}{2}\mathrm{tr}(\mathbf{X}'\sigma_3) = \frac{1}{2}(ct(e^\eta - e^{-\eta}) + z(e^\eta + e^{-\eta})) = ct\sinh\eta + z\cosh\eta \tag{8.67}$$

It is straightforward but tedious to verify this in boosts in the other directions. With these generators, a general transformation matrix can be written in the form

$$A_a{}^b = e^{-i(\theta_k \mathbf{J}_k + \eta_k \mathbf{K})} \tag{8.68}$$
$$= e^{-\frac{i}{2}(\theta_k\sigma_k + i\eta_k\sigma_k)} \tag{8.69}$$
$$= e^{-\frac{1}{2}(i\theta_k - \eta_k)\sigma_k} \tag{8.70}$$

Now let's take the complex conjugate.

$$(A^*)^{\dot{a}}{}_{\dot{b}} = \varepsilon^{\dot{a}\dot{c}}(A^*)_{\dot{c}}{}^{\dot{d}}\varepsilon_{\dot{d}\dot{b}} \tag{8.71}$$

where we've used the ε matrices to put the indices in the right place so we can use the expression we found for the $A_a{}^b$ matrix elements.

$$(A^*)_{\dot{c}}{}^{\dot{d}} = e^{-\frac{1}{2}(-i\theta_k - \eta_k)\sigma_k^*} \tag{8.72}$$

When we evaluate the exponential, we can insert $\varepsilon^{-1}\varepsilon$ in between all the σ_k^* matrices. We can evaluate these quickly:

$$\begin{aligned}
\varepsilon\sigma_1^*\varepsilon^{-1} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} && (8.73)\\
&= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} && (8.74)\\
&= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = -\sigma_1 && (8.75)\\
\varepsilon\sigma_2^*\varepsilon^{-1} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} && (8.76)\\
&= \begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} && (8.77)\\
&= \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = -\sigma_2 && (8.78)\\
\varepsilon\sigma_3^*\varepsilon^{-1} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} && (8.79)\\
&= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} && (8.80)\\
&= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -\sigma_3 && (8.81)
\end{aligned}$$

so in general

$$\varepsilon\sigma_k^*\varepsilon^{-1} = -\sigma_k \tag{8.82}$$

and we find

$$(A^*)^{\dot{a}}{}_{\dot{b}} = \varepsilon^{\dot{a}\dot{c}}(A^*)_{\dot{c}}{}^{\dot{d}}\varepsilon_{\dot{d}\dot{b}} = e^{-\frac{1}{2}(i\theta_k + \eta_k)\sigma_k^*} \tag{8.83}$$

We see then that using ε to transform $\mathbf{A}$ acts in different ways on rotations on boosts, as seen in the different signs in the exponent. This is an indication of the inequivalency between these two representations.

Since the sign of the boost is the difference, we can simply flip its sign for the conjugate representation:

$$\mathbf{K}_a = -\frac{i}{2}\sigma_a \tag{8.84}$$

The complexified generators are then

$$\mathbf{M}_a = \frac{1}{2}(\mathbf{J}_a + i\mathbf{K}_a) = \frac{1}{2}\sigma_a \tag{8.85}$$

$$\mathbf{N}_a = \frac{1}{2}(\mathbf{J}_a - i\mathbf{K}_a) = 0 \tag{8.86}$$

so we see that this is the $(\frac{1}{2}, 0)$ representation, whereas the original representation was $(0, \frac{1}{2})$.

Therefore in $SL(2,\mathbb{C})$ we have two conjugate but inequivalent representations. In order to keep them straight, physicists often decorate the spinor with a bar or dagger, and then also add dots to the indices. This is redundant, but is necessary for later shorthands (not that we'll get to them here, but it's worth mentioning in case you encounter this later). The main thing to remember about dotted and undotted spinor indices is that they are independent of one another, and are manipulated independently.

Conventionally, a spinor with an undotted index is called a "left-handed" Weyl spinor, whereas a dotted index indicates a "right-handed" one. In what sense are they handed (or "chiral")? If you specify a general transformation with all the rotation angles and boost directions, you will find that in the left-handed case a boost will be associated with a particular direction of rotation, while in the right-handed case, the same boost will come with an opposite sense of rotation.

It's also worth noting that we've put some time into discussing the algebra of spinors, but we haven't given a concrete form to the spinors themselves, as we did for $SU(2)$ by itself. The reason is that it's a whole topic in itself. One hint lies in the antisymmetric metric, since it implies that

$$|\psi|^2 = \varepsilon^{ab}\psi_a\psi_b = \psi_2\psi_1 - \psi_1\psi_2 \tag{8.87}$$

Therefore, for the norm not to vanish identically, the spinor components must *anti*-commute. These anti-commuting objects are called "Grassmann numbers", and are often used in physics (particularly quantum field theory) to describe fermion fields.

8.2.1 Direct product representations

Finally, we end up with the mixed rank-2 spinor

$$X'^{a\dot{b}} = (A^{-1})_c{}^a X^{c\dot{d}} (A^*)^{\dot{b}}{}_{\dot{d}} \tag{8.88}$$

but since we've seen before that

$$(A^{-1})_c{}^a = A^a{}_c \tag{8.89}$$

and that the transpose of the complex conjugate is

$$(A^*)^{\dot{b}}{}_{\dot{d}} = (A^\dagger)_{\dot{d}}{}^{\dot{b}} \tag{8.90}$$

then we have

$$X'^{a\dot{b}} = A^a{}_c X^{c\dot{d}} (A^\dagger)_{\dot{d}}{}^{\dot{b}} \tag{8.91}$$

$$\mathbf{X}' = \mathbf{A}\mathbf{X}\mathbf{A}^\dagger \tag{8.92}$$

which is what we started with. $X^{a\dot{b}}$ is a member of the $(\frac{1}{2},\frac{1}{2})$ representation. Since the pair of numbers expresses a direct product, we expect that, for instance, $(\frac{1}{2},\frac{1}{2})$ can be reduced to a direct sum in the usual way

$$\frac{1}{2}\otimes\frac{1}{2}=0\oplus 1 \tag{8.93}$$

In fact, we can get all the other (m,n) representations by taking direct products of $(0,\frac{1}{2})$ and $(\frac{1}{2},0)$.

8.3 Space inversion

Earlier, we talked of the parity transformation (or spatial inversion) as inaccessible to rotations in three dimensions. However, we know that most physics is invariant even under this transformation, so we'd like to enlarge our group to include this transformation as well. This would mean adding $SO(3,1)^{\uparrow}_{-}$ to the $SO(3,1)^{\uparrow}_{+}$ we've been considering so far. We'll just denote this in the utterly expected way, as $SO(3,1)^{\uparrow}$.

That said, we know that some physics actually *does* look different on spatial inversion. This phenomenon is called "parity violation", and is associated in paricular with the electroweak interaction in particle physics. For now, however, we'll take parity as a further symmetry.

The parity operator can be written in matrix form

$$\mathbf{P}=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix} \tag{8.94}$$

It is easy to see that this operator commutes with spatial rotations:

$$\mathbf{PRP}^{-1}=\mathbf{R} \tag{8.95}$$

which can be checked in matrix form, or by visualizing it.

For Lorentz transformations, we see, for example,

$$\mathbf{PL}_1\mathbf{P}^{-1} = \begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}\cosh\eta&\sinh\eta&0&0\\\sinh\eta&\cosh\eta&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix} \tag{8.96}$$

$$= \begin{pmatrix}\cosh\eta&-\sinh\eta&0&0\\-\sinh\eta&\cosh\eta&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix} \tag{8.97}$$

In general,

$$\mathbf{PL}(\eta)\mathbf{P}^{-1}=\mathbf{L}(-\eta)=\mathbf{L}^{-1}(\eta) \tag{8.98}$$

which makes sense, that 3-vectors should flip sign under spatial inversion.

In terms of generators,

$$\mathbf{P}\mathbf{J}_i\mathbf{P}^{-1} = \mathbf{J}_i \tag{8.99}$$
$$\mathbf{P}\mathbf{K}_i\mathbf{P}^{-1} = -\mathbf{K}_i \tag{8.100}$$

We can follow these into the complexified generators:

$$\mathbf{P}\mathbf{M}_j\mathbf{P}^{-1} = \mathbf{P}\left(\frac{\mathbf{J}_j + i\mathbf{K}_j}{2}\right)\mathbf{P}^{-1} = \frac{\mathbf{J}_j - i\mathbf{K}_j}{2} = \mathbf{N}_j \tag{8.101}$$
$$\mathbf{P}\mathbf{N}_j\mathbf{P}^{-1} = \mathbf{P}\left(\frac{\mathbf{J}_j - i\mathbf{K}_j}{2}\right)\mathbf{P}^{-1} = \frac{\mathbf{J}_j + i\mathbf{K}_j}{2} = \mathbf{M}_j \tag{8.102}$$

and thus, for the Casimir operators,

$$\mathbf{P}\mathbf{M}^2\mathbf{P}^{-1} = \mathbf{N}^2 \tag{8.103}$$
$$\mathbf{P}\mathbf{N}^2\mathbf{P}^{-1} = \mathbf{M}^2 \tag{8.104}$$

What this means is that the two parts of the direct product $SU(2) \times SU(2)$ are no longer completely factorized when considering spatial inversion as well. $\mathbf{P}$ connects them. As a result, a representation (u, v) becomes (v, u) under spatial inversion.

Since we expect physical objects to be in some representation space of $SO(3,1)^\uparrow$ (except for those cases where experiment has demonstrated otherwise!), we expect them to come in two classes: a direct sum of two representations $(u, v) \oplus (v, u)$, with $u \neq v$; and a self-conjugate representation, with $u = v$.

In the latter case, spatial inversion acts on the basis vectors (labelled by the eigenvalues m_u and m_v)

$$P|m_u\, m_v\rangle = \eta|m_v\, m_u\rangle \tag{8.105}$$

If we also demand that $P^2 = 1$, then we find that $\eta = \pm 1$. (In fact, with spinors it's also possible to use $P^2 = -1$, analogous to a rotation $\mathbf{R}(2\pi) = -1$.)

Scalars belong to the $(0, 0)$ representation space. If $\eta = 1$, then we have the normal scalars which are invariant under Lorentz transformations as well as spatial inversion. If $\eta = -1$, then it's still invariant under Lorentz transformations, but changes sign under spatial inversion; this is called a "pseudo-scalar".

4-vectors belong to the $(\frac{1}{2}, \frac{1}{2})$ space, as we saw earlier. Here again we have two cases: $\eta = 1$, which in this case means the sign changes under spatial inversion ("polar vectors"); and $\eta = -1$, in which case the sign doesn't change ("axial vectors"). Spacetime displacements, momentum, and vector potentials are polar vectors. We'll see examples of axial vectors later, though it's worth noting that a 3D version of an axial vector is the magnetic field.

Chapter 9

Poincaré group

So far, we've focused on rotations and boosts, but of course spacetime has translational symmetry, in that translations also keep the relativistic interval invariant. When we extend the symmetry group to include translations, we get the Poincaré group, which is sometimes called the "inhomogeneous Lorentz group".

We won't work out all the details, because what we're really aiming to do is to describe the representations.

A translation can be written

$$x^\mu \to x'^\mu = x^\mu + a^\mu \tag{9.1}$$

If we have a function of the spacetime event, we expand it to find the effect of an infinitesimal translation a^μ as follows:

$$f(x'^\mu) = f(x^\mu) + \frac{\partial f}{\partial x^\mu} a^\mu \tag{9.2}$$

We compare this with the generator definition

$$T(a^\mu) = 1 + i a^\mu P_\mu \tag{9.3}$$

which gives us the generators

$$P_\mu = -i\partial_\mu \tag{9.4}$$

These obviously commute amongst themselves. They're often designated the linear momentum operators.

Similarly, the angular momentum operators generate rotations; in Minkowski space, we also take in boosts. The spatial part of angular momentum can be written in differential form:

$$L_{\mu\nu} \equiv x_\mu P_\nu - x_\nu P_\mu = -i(x_\mu \partial_\nu - x_\nu \partial_\mu) \tag{9.5}$$

The commutator with the linear momentum operator is

$$\begin{aligned}
[L_{\mu\nu}, P_\rho] &= [-i(x_\mu\partial_\nu - x_\nu\partial_\mu), -i\partial_\rho] && (9.6)\\
&= -[x_\mu\partial_\nu, \partial_\rho] + [x_\nu\partial_\mu, \partial_\rho] && (9.7)\\
&= -x_\mu\partial_\nu\partial_\rho + \partial_\rho(x_\mu\partial_\nu) + x_\nu\partial_\mu\partial_\rho - \partial_\rho(x_\nu\partial_\mu) && (9.8)\\
&= (\partial_\rho g_{\mu\lambda}x^\lambda)\partial_\nu - (\partial_\rho g_{\nu\lambda}x^\lambda)\partial_\mu && (9.9)\\
&= g_{\mu\lambda}\delta^\lambda_\rho\partial_\nu - g_{\nu\lambda}\delta^\lambda_\rho\partial_\mu && (9.10)\\
&= g_{\mu\rho}\partial_\nu - g_{\nu\rho}\partial_\mu && (9.11)\\
&= i(g_{\mu\rho}(-i\partial_\nu) - g_{\nu\rho}(-i\partial_\mu)) && (9.12)\\
&= i(g_{\mu\rho}P_\nu - g_{\nu\rho}P_\mu) && (9.13)
\end{aligned}$$

Similarly, one can evaluate the commutators amongst the L's. The commutators are then

$$\begin{aligned}
[P_\mu, P_\nu] &= 0 && (9.14)\\
[L_{\mu\nu}, P_\rho] &= i(g_{\mu\rho}P_\nu - g_{\nu\rho}P_\mu) && (9.15)\\
[L_{\mu\nu}, L_{\kappa\lambda}] &= i(L_{\mu\lambda}g_{\nu\kappa} + L_{\nu\kappa}g_{\mu\lambda} - L_{\mu\kappa}g_{\nu\lambda} - L_{\nu\lambda}g_{\mu\kappa}) && (9.16)
\end{aligned}$$

Now, L is not the most general form of an angular momentum operator you can write. You can also add a term

$$J_{\mu\nu} = L_{\mu\nu} + S_{\mu\nu} \tag{9.17}$$

as long as the $S_{\mu\nu}$ commutes with the $L_{\mu\nu}$, and has all the same commutation relations—in other words, acts just like another angular momentum. We can then summarize the commutation relations with this generalized angular momentum instead:

$$\begin{aligned}
[P_\mu, P_\nu] &= 0 && (9.18)\\
[J_{\mu\nu}, P_\rho] &= i(g_{\mu\rho}P_\nu - g_{\nu\rho}P_\mu) && (9.19)\\
[J_{\mu\nu}, J_{\kappa\lambda}] &= i(J_{\mu\lambda}g_{\nu\kappa} + J_{\nu\kappa}g_{\mu\lambda} - J_{\mu\kappa}g_{\nu\lambda} - J_{\nu\lambda}g_{\mu\kappa}) && (9.20)
\end{aligned}$$

In matrix form, the $J_{\mu\nu}$ operators can be written in a fairly easy-to-remember antisymmetric form

$$(J_{\mu\nu})^{\rho\sigma} = i(\delta^\rho_\mu\delta^\sigma_\nu - \delta^\sigma_\mu\delta^\rho_\nu) \tag{9.21}$$

Note, however, that the matrix indices are both raised. To get this into the form with which we do matrix multiplication, we take

$$\begin{aligned}
(J_{\mu\nu})^\rho{}_\sigma &= g_{\sigma\lambda}(J_{\mu\nu})^{\rho\lambda} && (9.22)\\
&= ig_{\sigma\lambda}(\delta^\rho_\mu\delta^\lambda_\nu - \delta^\lambda_\mu\delta^\rho_\nu) && (9.23)\\
&= i(\delta^\rho_\mu g_{\sigma\nu} - \delta^\rho_\nu g_{\sigma\mu}) && (9.24)
\end{aligned}$$

The familiar generators are then

$$\begin{aligned}
J_i &= \frac{1}{2}\epsilon_{ijk}J_{jk} && (9.25)\\
K_i &= J_{0i} && (9.26)
\end{aligned}$$

One can confirm this results in the usual matrix generators.

9.1 Casimir operators

There are two Casimir operators. The most obvious is $P_\mu P^\mu$. Since the eigenvalues of P^μ are the momenta components p^μ, the eigenvalues of $P_\mu P^\mu$ are $p_\mu p^\mu$.

The second Casimir operator is a generalization of angular momentum, but now folded in with momentum. Its role is played by what is called the "Pauli-Lubanski" vector

$$W^\mu \equiv \frac{1}{2}\epsilon^{\mu\nu\kappa\lambda} P_\nu J_{\kappa\lambda} \tag{9.27}$$

where $\epsilon^{\mu\nu\kappa\lambda}$ is the 4-index Levi-Civita symbol. We've written this with the generalized angular momentum, but it is useful to break this up into orbital and internal (spin) angular momentum parts. The orbital angular momentum has the form

$$L_{\kappa\lambda} = x_\kappa P_\lambda - x_\lambda P_\kappa \tag{9.28}$$

so the orbital part of the Pauli-Lubanski vector is

$$\begin{aligned} W^\mu &= \frac{1}{2}\epsilon^{\mu\nu\kappa\lambda} P_\nu (x_\kappa P_\lambda - x_\lambda P_\kappa) && (9.29)\\ &= \frac{1}{2}\epsilon^{\mu\nu\kappa\lambda} (x_\kappa P_\nu P_\lambda - x_\lambda P_\nu P_\kappa) && (9.30)\\ &= -\frac{1}{2}\epsilon^{\mu\kappa\nu\lambda} (x_\kappa P_\nu P_\lambda - x_\lambda P_\nu P_\kappa) && (9.31) \end{aligned}$$

(the last step is simply to put the ν index next to κ and λ so it's only one index swap away). Since the P's commute, but ϵ is completely antisymmetric, each term sums to zero.

The same argument doesn't apply to the spin angular momentum. Instead, it is useful to evaluate the vector by component. First, we remind ourselves that we can get the space components of the spin vector from the definitions we used before:

$$S^i = \frac{1}{2}\epsilon^{ijk} S_{jk} \tag{9.32}$$

and $S^0 = 0$. The Pauli-Lubanski components are then

$$\begin{aligned} W^0 &= \frac{1}{2}\epsilon^{0ijk} P_i S_{jk} && (9.33)\\ &= P_i S^i && (9.34)\\ &= \mathbf{p}\cdot\mathbf{s} && (9.35)\\ W^i &= \frac{1}{2}\epsilon^{i\nu\kappa\lambda} P_\nu S_{\kappa\lambda} && (9.36)\\ &= \frac{1}{2}\epsilon^{i0\kappa\lambda} P_0 S_{\kappa\lambda} && (9.37)\\ &= -\frac{1}{2}\epsilon^{0ijk} P_0 S_{jk} && (9.38)\\ &= -P_0 S^i && (9.39)\\ &= (E/c)\mathbf{s} && (9.40) \end{aligned}$$

(keeping in mind that P_0 is the covariant component of the usual 4-momentum). In summary, the orbital angular momentum parts have dropped out and we are left with

$$\mathsf{W} = (\mathbf{p}\cdot\mathbf{s}, (E/c)\mathbf{s}) \tag{9.41}$$

The properties of W^μ are as follows:

$$\begin{aligned} W^\mu P_\mu &= 0 && (9.42)\\ [W^\mu, P^\mu] &= 0 && (9.43)\\ [W^\lambda, J^{\mu\nu}] &= i(W^\nu g^{\lambda\mu} - W^\mu g^{\nu\lambda}) && (9.44)\\ [W^\mu, W^\nu] &= i\epsilon^{\mu\nu\kappa\lambda} W_\kappa P_\lambda && (9.45) \end{aligned}$$

The Casimir operator is $W^\mu W_\mu$.

9.2 Representation space of the Poincaré group

The representation space of the Poincaré group now brings us back to physical particles, the inhabitants of the representation space of the symmetry groups we've been discussing.

The representations of the Poincaré group can be used to classify particles and fields (V Bargman, EP Wigner, *Proc Natl Ac Sci*, 34:5, 211 (1946)). The unitary representations are as follows:

1. $P^\mu P_\mu = -m^2$ where m is a real number. These are finite-mass particles, with spin values $s = 0, \frac{1}{2}, 1, \frac{3}{2}$, *etc.* The eigenvalue of $W^\mu W_\mu$ is $m^2 s(s+1)$. States are labelled by the spin z-component s_3 and the (continuous) 3-momentum $\mathbf{p}$.

2. $P^\mu P_\mu = 0$ and $W^\mu W_\mu = 0$. We can write these 4-vectors as

$$\begin{aligned} \mathsf{P} &= (|\mathbf{p}|, \mathbf{p}) && (9.46)\\ \mathsf{W} &= (w^0, \mathbf{w}) && (9.47) \end{aligned}$$

   with the 0th component of W

$$\begin{aligned} (w^0)^2 &= \mathbf{w}\cdot\mathbf{w} && (9.48)\\ w^0 &= \pm\sqrt{|\mathbf{w}|} && (9.49) \end{aligned}$$

   However, we also have

$$\begin{aligned} 0 &= \mathsf{P}\cdot\mathsf{W} && (9.50)\\ &= -w^0|\mathbf{p}| + \mathbf{p}\cdot\mathbf{w} && (9.51)\\ w^0|\mathbf{p}| &= |\mathbf{p}||\mathbf{w}|\cos\theta && (9.52) \end{aligned}$$

   where θ is the angle between $\mathbf{w}$ and $\mathbf{p}$. Clearly $\cos\theta = \pm 1$, so $\mathbf{w}$ and $\mathbf{p}$ are either parallel or anti-parallel.

The constant of proportionality between these vectors is the "helicity"

$$\lambda = \frac{w^0}{p^0} = \frac{\mathbf{s}\cdot\mathbf{p}}{|\mathbf{p}|} = \pm s \tag{9.53}$$

since $\mathbf{s}$ is proportional to $\mathbf{w}$, which is parallel or anti-parallel to $\mathbf{p}$. Therefore massless particles with non-zero spin have two helicities: photons with $s = \pm 1$, gravitons with $s = \pm 2$, and (in the massless approximation) neutrinos with $s = \pm\frac{1}{2}$.

3. $P^\mu P_\mu = 0$ but $W^\mu W_\mu > 0$. These are massless particles but with continuous spin. Particles of this type have not been found.

4. $P^\mu P_\mu > 0$. Particles in this category would be tachyons. Only found in two circumstances: science fiction, or virtual contributions to (non-fiction) scattering amplitudes.

9.2.1 Supersymmetry and spacetime

The Poincaré algebra is complete in itself: its generators all commute among one another and to no other generators.

The Coleman-Mandula theorem states that one cannot introduce any further generators into this algebra. In other words, if there are further symmetries, they can only be internal symmetries, and their Lie algebra will not involve non-trivial commutators with the generators of translations, rotations, and boosts. The linear space over which the generators operate is added to spacetime as a direct product.

The essence of proof is (apparently) fairly simple: in a hypothetical scattering experiment in which the input 4-momenta of 2 particles are known, the final state is known up to the scattering angle—in classical mechanics, this scattering angle is determined by the impact parameter, which of course is not specified among the 4-momenta. On the other hand, if there are further symmetries, then the final state would lose this degree of freedom, and would either be completely determined or overconstrained. (If you want to look it up, it may be worth reading later versions of the proof, for instance by Witten or Weinberg, rather than the original.)

However, there is a loophole in the theorem: it only applies to vectorial degrees of freedom. Spinors are not similarly constrained. As a result, spinorial operators can be added to the Poincaré algebra. This extension is called "supersymmetry". In quantum field theory, this results in certain famous results, among them the idea that for every fundamental fermion there must be a bosonic "superpartner". In a sense, supersymmetry is an extension of spacetime symmetries, and it's such a compelling idea that twenty years ago, some senior theoretical physicists expressed concern that there might be theory postgraduates out there who didn't realize it was an unproven idea.

Twenty years later, it remains unproven, so the role it *might* play in physics is so far undetermined. Its relevant energy scale is so far anywhere between whatever current experimental bounds exist, and the unification of gravity with all the other forces. Experimentally, the most obvious manifestation of supersymmetry is the flagrant overuse of the prefix "super".

9.3 Physics tensors

Let's continue to look at members of the representation space of our symmetry groups. There are not just particles, but also fields which transform as tensors.

9.3.1 3D tensors

We have already seen one rank-2 tensor, albeit as an operator: (spacetime) angular momentum.

$$L_{\mu\nu} = x_\mu p_\nu - x_\nu p_\mu \tag{9.54}$$

It's worth taking a moment to look at 3D tensors in connection with angular momentum. There is the angular momentum tensor itself

$$L_{ij} = x_i p_j - x_j p_i \tag{9.55}$$

as well as the momentum of inertia tensor.

$$I_{ij} = \int_V dm(r^2\delta_{ij} - r_i r_j) \tag{9.56}$$

Since we're dealing with the spatial dimensions, the symmetry group we're talking about is $SO(3)$ of rotations. Therefore the rotation takes the form

$$L'_{ij} = R_{im}R_{jn}L_{mn} \tag{9.57}$$

in which we recognize the matrix equation

$$\mathbf{L}' = \mathbf{R}\mathbf{L}\mathbf{R}^T \tag{9.58}$$

Moreover, since $\mathbf{R}$ is orthogonal, $\mathbf{R}^T = \mathbf{R}^{-1}$, so the transformation is the same as a similarity transform. In other words, you may have seen this rotation as a basis change for a matrix. Equivalently (for 3D), it's a legitimate tensor transformation.

9.3.2 *Transformation of electromagnetic fields

Now let's take a peek at electromagnetism. We have two fields, $\mathbf{E}$ and $\mathbf{B}$. How do they transform between frames?

One approach is to look at how the force transforms. After all, it's the force, rather the fields, which you "see" by their effect on other particles. We will see that the fields don't look like Lorentz 4-vectors.

The force per unit charge is

$$\mathbf{f} = q(\mathbf{E} + \mathbf{v} \wedge \mathbf{B}) \tag{9.59}$$

Notice that for $\mathbf{f}$ to be a normal polar vector which changes sign under spatial inversion, we need $\mathbf{E}$ to also be a polar vector. The cross product then means that $\mathbf{B}$ has to be an axial vector which doesn't change sign under spatial inversion.

The first scenario to consider: a single test charge q in a constant electric field $\mathbf{E}$ but with $\mathbf{B} = 0$. The test charge has a velocity $\mathbf{u}$ in frame S. The force in the test charge is thus

$$\mathbf{f} = q\mathbf{E} \tag{9.60}$$

Now let's transform into the frame S' which travels with velocity $\mathbf{v}$ in S. We know that the force should transform as follows:

$$\mathbf{f}'_{\parallel} = \frac{\mathbf{f}_{\parallel} - \left(\frac{\mathbf{v}}{c^2}\right)\frac{dE}{dt}}{1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \tag{9.61}$$

$$\mathbf{f}'_{\perp} = \frac{\mathbf{f}_{\perp}}{\gamma_v\left(1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}\right)} \tag{9.62}$$

For a pure force (when the rest mass stays constant, so the force goes into the kinetic energy of the particle), we have

$$0 = \mathsf{U}\cdot\mathsf{F} = (\gamma_u c, \gamma_u\mathbf{u})\cdot\left(\frac{\gamma_u}{c}\frac{dE}{dt}, \gamma_u\mathbf{f}\right) = \gamma_u^2\left(-\frac{dE}{dt} + \mathbf{u}\cdot\mathbf{f}\right) \tag{9.63}$$

so we see that

$$\frac{dE}{dt} = \mathbf{f}\cdot\mathbf{u} \tag{9.64}$$

Plugging this into the formula for $\mathbf{f}_{\parallel}$, we get

$$\mathbf{f}'_{\parallel} = \frac{\mathbf{f}_{\parallel} - \mathbf{v}(\mathbf{f}\cdot\mathbf{u})/c^2}{1 - \mathbf{u}\cdot\mathbf{v}/c^2} \tag{9.65}$$

If we choose $\mathbf{v} = \mathbf{u}$ such that the test particle is at rest in S', we have

$$\mathbf{f}_{\parallel} = q\mathbf{E}_{\parallel} \tag{9.66}$$

$$\mathbf{f}\cdot\mathbf{u} = q\mathbf{E}\cdot\mathbf{u} \tag{9.67}$$

$$\mathbf{f}'_{\parallel} = \frac{q(\mathbf{E}_{\parallel} - \mathbf{v}(\mathbf{E}\cdot\mathbf{u})/c^2)}{1 - \mathbf{u}\cdot\mathbf{v}/c^2} \tag{9.68}$$

$$= \frac{q(\mathbf{E}_{\parallel} - \mathbf{E}_{\parallel}\beta^2)}{1 - \beta^2} \tag{9.69}$$

$$= q\mathbf{E}_{\parallel} \tag{9.70}$$

$$\mathbf{f}'_{\perp} = \frac{q\mathbf{E}_{\perp}}{\gamma(1 - \beta^2)} \tag{9.71}$$

$$= \gamma q\mathbf{E}_{\perp} \tag{9.72}$$

from which we infer that

$$\mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel} \tag{9.73}$$

$$\mathbf{E}'_{\perp} = \gamma\mathbf{E}_{\perp} \tag{9.74}$$

It is clear that $\mathbf{E}_{\parallel}$ doesn't transform at all like a 4-vector, which would boost the longitudinal component, and leave the transverse part alone.

And indeed there is also a magnetic field in S', though for the $\mathbf{v} = \mathbf{u}$ case, the resulting force is zero, since the test charge is at rest. To see some effect, keep $\mathbf{v}$ parallel to $\mathbf{u}$, but now $\mathbf{v} \neq \mathbf{u}$. The velocity of the test charge in S' is now $\mathbf{u}'$, which must still be parallel to $\mathbf{v}$. But since the force in S' must be of the form

$$\mathbf{f}' = q(\mathbf{E}' + \mathbf{u}' \wedge \mathbf{B}') \tag{9.75}$$

we will only pick out the component of $\mathbf{B}'$ which is perpendicular to $\mathbf{v}$.

Using the transformation formula for the transverse force,

$$\begin{aligned} \mathbf{f}'_\perp &= \frac{\mathbf{f}_\perp}{\gamma_v(1 - \mathbf{u}\cdot\mathbf{v}/c^2)} && (9.76)\\ &= \frac{q\mathbf{E}_\perp}{\gamma_v(1 - \mathbf{u}\cdot\mathbf{v}/c^2)} && (9.77) \end{aligned}$$

We'd like to write this in terms of quantities in S'. The denominator can be changed by considering the velocity addition formula:

$$\begin{aligned} \mathbf{u}'_\parallel &= \frac{\mathbf{u}_\parallel - \mathbf{v}}{1 - \mathbf{u}\cdot\mathbf{v}/c^2} && (9.78)\\ \mathbf{u}'_\parallel \cdot \mathbf{v} &= \frac{\mathbf{u}_\parallel \cdot \mathbf{v} - v^2}{1 - \mathbf{u}\cdot\mathbf{v}/c^2} && (9.79)\\ 1 + \mathbf{u}'_\parallel \cdot \mathbf{v}/c^2 &= \frac{1 - \mathbf{u}\cdot\mathbf{v}/c^2 + \mathbf{u}\cdot\mathbf{v}/c^2 - v^2/c^2}{1 - \mathbf{u}\cdot\mathbf{v}/c^2} && (9.80)\\ &= \frac{1 - v^2/c^2}{1 - \mathbf{u}\cdot\mathbf{v}/c^2} && (9.81)\\ \gamma_v^2(1 + \mathbf{u}'_\parallel \cdot \mathbf{v}/c^2) &= \frac{1}{1 - \mathbf{u}\cdot\mathbf{v}/c^2} && (9.82) \end{aligned}$$

so plugging in,

$$\mathbf{f}'_\perp = q\mathbf{E}_\perp\gamma_v(1 + \mathbf{u}'_\parallel \cdot \mathbf{v}/c^2) \tag{9.83}$$

Comparing this with the force equation, we see that

$$\mathbf{u}' \wedge \mathbf{B}' = \mathbf{E}_\perp\gamma_v u' v/c^2 \tag{9.84}$$

This is consistent with a perpendicular magnetic field component

$$\mathbf{B}'_\perp = -\gamma(\mathbf{v} \wedge \mathbf{E})/c^2 \tag{9.85}$$

In any case, it's clear that the Lorentz transformation is mixing up electric and magnetic fields.

9.3.3 *The Maxwell field tensor

What is happening? It turns out that $\mathbf{E}$ and $\mathbf{B}$ aren't parts of independent 4-vectors, but rather of a single tensor, which we can write

$$F^{\mu\nu} \equiv \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{pmatrix} \tag{9.86}$$

This is the "Maxwell field tensor". In terms of Lorentz group representation, it belongs to $(0,1) \oplus (1,0)$, which brings out the involvement of two types of 3-vectors.

A Lorentz transformation takes the form

$$F'^{\mu\nu} = \Lambda^{\mu}{}_{\kappa}\Lambda^{\nu}{}_{\lambda}F^{\kappa\lambda} \tag{9.87}$$

or in matrix form (suppressing c for now)

$$\mathbf{F}' = \Lambda\mathbf{F}\Lambda^T \tag{9.88}$$

$$= \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix} \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{9.89}$$

$$= \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\beta\gamma E_x & \gamma E_x & E_y & E_z \\ -\gamma E_x & \beta\gamma E_x & B_z & -B_y \\ -\gamma E_y + \beta\gamma B_z & \beta\gamma E_y - \gamma B_z & 0 & B_x \\ -\gamma E_z - \beta\gamma B_y & \beta\gamma E_z + \gamma B_y & -B_x & 0 \end{pmatrix} \tag{9.90}$$

$$= \begin{pmatrix} 0 & \gamma^2 E_x - \beta^2\gamma^2 E_x & \gamma E_y - \beta\gamma B_z & \gamma E_z + \beta\gamma B_y \\ \beta^2\gamma^2 E_x - \gamma^2 E_x & 0 & -\beta\gamma E_y + \gamma B_z & -\beta\gamma E_z - \gamma B_y \\ -\gamma E_y + \beta\gamma B_z & \beta\gamma E_y - \gamma B_z & 0 & B_x \\ -\gamma E_z - \beta\gamma B_y & \beta\gamma E_z + \gamma B_y & -B_x & 0 \end{pmatrix} \tag{9.91}$$

which results in the following transformed fields:

$$E'_x = E_x \tag{9.92}$$
$$E'_y = \gamma(E_y - \beta B_z) \tag{9.93}$$
$$E'_z = \gamma(E_z + \beta B_y) \tag{9.94}$$
$$B'_x = B_x \tag{9.95}$$
$$B'_y = \gamma(B_y + \beta E_z) \tag{9.96}$$
$$B'_z = \gamma(B_z - \beta E_y) \tag{9.97}$$

If we rotate our transformation to another axis, we get the following more general forms:

$$\mathbf{E}_{\parallel} \rightarrow \mathbf{E}_{\parallel} \tag{9.98}$$
$$\mathbf{E}_{\perp} \rightarrow \gamma(\mathbf{E}_{\perp} + \mathbf{v} \wedge \mathbf{B}) \tag{9.99}$$
$$\mathbf{B}_{\parallel} \rightarrow \mathbf{B}_{\parallel} \tag{9.100}$$
$$\mathbf{B}_{\perp} \rightarrow \gamma(\mathbf{B}_{\perp} - \mathbf{v} \wedge \mathbf{E}) \tag{9.101}$$
$$\tag{9.102}$$

Chapter 10

Classical fields

We've now looked at what kinds of objects transform, and how they transform. In the process we found that there are some additional objects which obey the transformation, *i.e.*, are consistent with Special Relativity. Now we'll look at how we join these things together in equations of motion.

The main lesson: Lagrangians are functionals which can be made to result in a scalar. They have to be scalars in order for the search for an extremal path to have a meaning; if they weren't, there would be no ordering and therefore no extreme value. And because the scalar is invariant with respect to Lorentz transformations, we'll find that the extremal path is also invariant.

This isn't quite as possible with Hamiltonians, since in many cases Hamiltonians are a component of a 4-vector.

We have taken the symmetries of Special Relativity as the focus of our study, but it's not the only possible set of symmetries. In fact, an important avenue for research is to understand other symmetries we observe, sometimes only indirectly, and sometimes only in approximation. In order to incorporate them, however, we need to find an action which has both old and new symmetries. So finding "new physics" is a matter of elucidating the form of the action. Symmetries guide and constrain the form of such actions.

10.1 The field viewpoint

In previous lectures, we've seen how a Lagrangian can be written to yield results consistent with Special Relativity. Those Lagrangians were limited to the behavior of individual particles in some external potential which was not itself form-invariant. This is of somewhat limited applicability. What we'd like to do is write Lagrangians which are entirely form-invariant, including all the interactions.

The interactions present a particular problem, since in Special Relativity the interactions themselves have a propagation speed. But there is a way to do this, by folding those interactions into a continous field. The interactions then propagate as field changes: an interaction can start in one location, find its way to another, and then cause its effect.

This then raises the other main motivation for a field viewpoint: locality. With the field viewpoint, we have been able to phrase all known laws of Nature as local interactions, *i.e.*, interactions happen where things overlap in spacetime. We have thus been able to eliminate notions of action at a distance, which was a rather unsatisfactory model even as Newton proposed it for gravity.

10.2 Continuous systems

(Based on Goldstein chapter 12)

Let's take an example of a discrete system from which we can extract a continuous limit: a line of coupled uniform harmonic oscillators.

The non-relativistic Lagrangian:

$$T = \frac{1}{2}\sum_i m\dot{\eta}_i^2 \tag{10.1}$$

$$V = \frac{1}{2}\sum_i k(\eta_{i+1}-\eta_i)^2 \tag{10.2}$$

$$L = T - V \tag{10.3}$$

$$= \frac{1}{2}\sum_i (m\dot{\eta}_i^2 - k(\eta_{i+1}-\eta_i)^2) \tag{10.4}$$

where η_i are the displacements of the ith particles from their equilibrium positions.

Introduce the parameter a, the equilibrium separation between masses:

$$L = \frac{1}{2}\sum_i a\left(\frac{m}{a}\dot{\eta}_i - ka\left(\frac{\eta_{i+1}-\eta_i}{a}\right)^2\right) = \sum_i aL_i \tag{10.5}$$

So in the limit of infinitesimal a, we have $m/a \to \mu$, the mass density, and $ka \to Y$, the Young's modulus (extension per unit length of an elastic rod, which is proportional to the force exerted on the rod). Then we take the sum into an integral

$$L = \frac{1}{2}\int\left(\mu\left(\frac{\partial\eta}{\partial t}\right)^2 - Y\left(\frac{\partial\eta}{\partial x}\right)^2\right)dx \tag{10.6}$$

where now $\eta \equiv \eta(x,t)$. Note that now that η is a function of both x and t, we need to use explicit partial derivatives.

The most important point here is that the position coordinate x is no longer one of the generalized coordinates of a path. Instead, it serves as a label: it replaces the label i. So x becomes like t in its status in the Lagrangian. Likewise, the generalized coordinates are now η and its derivatives with respect to x and t.

10.3 Lagrangian density

Now let's look at this more generically.

The Lagrangian we want would integrate the field over all the space "labels":

$$L = \int d^3x \mathcal{L}(\eta, \partial_\mu \eta, x^\mu) \tag{10.7}$$

where

$$\partial_\mu \eta \equiv \frac{\partial \eta}{\partial x^\mu} \tag{10.8}$$

is a shorthand for a particular kind of partial derivative where you differentiate with respect to one "label" coordinate (such as x or t) while holding the other "label" coordinates constants. This isn't much of an issue for a field η, which here is only a function of those label-type coordinates. However, $\mathcal{L}$ is a function of the fields, the field derivatives, and the coordinates. So

$$\partial_\mu \mathcal{L} = \frac{\partial}{\partial x^\mu} \mathcal{L}(\eta(x^\nu), \partial_\mu \eta(x^\nu), x^\nu) \tag{10.9}$$

will be interpreted as the partial derivative of $\mathcal{L}$ considering it as a function of the "label" coordinates. In other words, it differentiates through the fields via the chain rule. This has rather clumsily been called a "total partial derivative".

The action to minimize is then

$$I[\eta(x^\mu)] = \int L dt = \iint_V d^4x \mathcal{L}(\eta, \partial_\mu \eta, x^\mu) \tag{10.10}$$

$\mathcal{L}$ is technically the "Lagrangian density", but it's often called the "Lagrangian" anyway, the distinction usually being clear in context. The spacetime volume V is such that the field values of η are fixed on its boundaries.

As we noted before, the action isn't a property of a particle or a field: its only job is to give the equations of motion via the Euler-Lagrange equations. There are therefore any number of possible actions to describe some given physics. On the other hand, experience has given us some guidelines as to some reasonable constraints (Ramond, p 24):

1. the fields in the Lagrangian depend only on one spacetime point, *i.e.*, we consider only "local" field theories. Indeed, we use local field theories to describe non-local phenomena as well.

2. the action I is real-valued. Complex-valued actions—specifically, potentials—tend to result in matter disappearing, which isn't very satisfying classical physics (at least).

3. the Lagrangian depends on no higher than second derivatives. Higher-order differential equations tend to lead to non-causal solutions. The implication is that Lagrangians tend to have products of first derivatives, which result in second-order derivatives in the equations of motion.

4. the action reflects other symmetries. For instance, most of the actions and Lagrangians we'll see in this course are relativistically invariant, *i.e.*, they are scalars. But there may be further symmetries with respect to other degrees of freedom, such as electric charge. In quantum field theories, there will be further internal degrees of freedom, such as the phase of the quantum field at a given spacetime point.

Let's illustrate evaluating the stationary field configuration for one field $\eta(x,t)$ in x and t.

$$\mathcal{L} \equiv \mathcal{L}(\eta, \partial_x\eta, \partial_t\eta, x, t) \tag{10.11}$$

The variation is then

$$0 = \delta I = \delta \int dxdt\mathcal{L} \tag{10.12}$$

It should be noted that the variation changes the field and its derivative, not x and t, which are just labels for the field values. In a sense, the extension of classical field theory to special relativity is actually rather simple, since the Lorentz transformation only affects the labels. As long as the quantities which make up $\mathcal{L}$ transform appropriately, you can get away without worrying too much about Special Relativity itself—except for a subtle point about integration limits.

Since we know the fields at the x and t boundary, we take the variation to be zero at the boundary. Let

$$\eta(x,t;\alpha) = \eta(x,t;0) + \alpha\zeta(x,t) \tag{10.13}$$

where $\eta(x,t;0)$ is the correct function which satisfies Hamilton's Principle, and $\zeta(x,t)$ is any well-behaved (continuous) function which vanishes at the boundary.

Take the derivative of the integral

$$\begin{aligned} 0 &= \frac{dI}{d\alpha} = \int dxdt \frac{d\mathcal{L}}{d\alpha} && (10.14)\\ &= \int dxdt \left(\frac{\partial\mathcal{L}}{\partial\eta}\frac{\partial\eta}{\partial\alpha} + \frac{\partial\mathcal{L}}{\partial(\partial_t\eta)}\frac{\partial(\partial_t\eta)}{\partial\alpha} + \frac{\partial\mathcal{L}}{\partial(\partial_x\eta)}\frac{\partial(\partial_x\eta)}{\partial\alpha} \right) && (10.15)\\ &= \int dxdt \left(\frac{\partial\mathcal{L}}{\partial\eta}\zeta + \frac{\partial\mathcal{L}}{\partial(\partial_t\eta)}\frac{\partial\zeta}{\partial t} + \frac{\partial\mathcal{L}}{\partial(\partial_x\eta)}\frac{\partial\zeta}{\partial x} \right) && (10.16) \end{aligned}$$

At this point most texts refer blithely to "integrating by parts", but let's be a bit more explicit here because it'll help us understand the nature of derivatives we take.

In multivariable calculus, integrating by parts is essentially an application of Stokes' Theorem (or its specialization or generalization, depending on the number of dimensions you're working in). In this case we're specializing to two dimensions, so we'll use Stokes' Theorem in x and y (to keep our minds on the geometric nature of this integral):

$$\int_S \nabla \wedge \mathbf{F} \cdot d\mathbf{a} = \oint_C \mathbf{F} \cdot d\mathbf{s} \tag{10.17}$$

we get for a field $\mathbf{F} = (P, Q, 0)$, where P and Q are both functions of x and y,

$$\int_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy = \oint_C (Pdx + Qdy) \tag{10.18}$$

Now let's go back into $x - -t$ space by changing y to t, and then substituting

$$P = -\frac{\partial \mathcal{L}}{\partial(\partial_t \eta)}\zeta \tag{10.19}$$

$$Q = \frac{\partial \mathcal{L}}{\partial(\partial_x \eta)}\zeta \tag{10.20}$$

Now we need to careful with derivatives here. As far as Stokes' Theorem is concerned, P and Q are functions only of x and t (or y), but $\mathcal{L}$ is a function of η and its derivatives as well. So when we take the derivative of Q with respect to x, we actually need to carry the derivative through η and its derivatives (like a full derivative), while keeping t constant (like a partial derivative). In other words, the derivative we need is the "total partial derivative" we described earlier.

$$\frac{\partial Q}{\partial x} \to \partial_x Q = \frac{\partial Q}{\partial x} + \frac{\partial Q}{\partial \eta}\frac{\partial \eta}{\partial x} + \frac{\partial Q}{\partial(\partial_x \eta)}\frac{\partial(\partial_x \eta)}{\partial x} + \frac{\partial Q}{\partial(\partial_t \eta)}\frac{\partial(\partial_t \eta)}{\partial x} \tag{10.21}$$

So we end up with

$$\int_S (\partial_x Q - \partial_t P)dxdt = \int_S \left(\partial_x \left(\frac{\partial \mathcal{L}}{\partial(\partial_x \eta)}\right)\zeta + \frac{\partial \mathcal{L}}{\partial(\partial_x \eta)}\frac{\partial \zeta}{\partial x} + \partial_t \left(\frac{\partial \mathcal{L}}{\partial(\partial_t \eta)}\right)\zeta + \frac{\partial \mathcal{L}}{\partial(\partial_t \eta)}\frac{\partial \zeta}{\partial t}\right)dxdt \tag{10.22}$$

and

$$\oint_C (Pdx + Qdt) = \oint_C \zeta\left(-\frac{\partial \mathcal{L}}{\partial(\partial_t \eta)}dx + \frac{\partial \mathcal{L}}{\partial(\partial_x \eta)}dt\right) \tag{10.23}$$

which is zero because $\zeta = 0$ on the boundary.

Thus we end up with

$$0 = \int dxdt \left[\frac{\partial \mathcal{L}}{\partial \eta} - \partial_t \left(\frac{\partial \mathcal{L}}{\partial(\partial_t \eta)}\right) - \partial_x \left(\frac{\partial \mathcal{L}}{\partial(\partial_x \eta)}\right)\right]\zeta \tag{10.24}$$

In this case it doesn't help to "choose" ζ to be zero anywhere, because the integral is throughout the spacetime volume of paths.

But since the path variation ζ is arbitrary, the square bracketed part must vanish.

So in general, for each field ϕ_k,

$$0 = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_k)}\right) - \frac{\partial \mathcal{L}}{\partial \phi_k} \tag{10.25}$$

These are the Euler-Lagrange equations for fields.

It is also worth reminding oneself that the derivative with respect to a contravariant coordinate/label x^μ is covariant, and the derivative of $\mathcal{L}$ with respect to the (covariant) field derivative is contravariant. Thus the sum in the first term is a simple sum—no metric needed.

In summary, the Lagrangian approach for fields is very similar to those for particles, but with coordinates q_i and $\dot{q}_i$ replaced with an infinite number of values indexed by space-time

position.

$$
\begin{aligned}
i &\rightarrow x^{\mu}, k && (10.26)\\
q_i &\rightarrow \phi_k(x) && (10.27)\\
\dot{q}_i &\rightarrow \partial_{\mu}\phi_k(x) && (10.28)\\
L = \sum_i L_i(q_i, \dot{q}_i) &\rightarrow \int \mathcal{L}(\phi_k, \partial_{\mu}\phi_k) d^3x && (10.29)
\end{aligned}
$$

Finally, in order to guarantee Lorentz form-invariance, $\mathcal{L}$ has to be a scalar: it has to be constructed from elements of the representation space for the Poincaré symmetry group.

Chapter 11

Relativistic field equations

Now we'll take a look at some relativistic field equations.

11.1 *Classical Klein-Gordon equation

The simplest field should be that of the $(0,0)$ representation, which is just a scalar field $\phi(x^\mu)$. We can also form vectors out of the derivatives $\partial_\mu\phi$, though these would have to be contracted to give a scalar in the Lagrangian.

If we also want to have a linear equation of motion, we can only involve up to quadratic terms in ϕ and $\partial_\mu\phi$. So this suggests a Lagrangian

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m^2\phi^2 \tag{11.1}$$

where the signs and factors anticipate later interpretation. Note that the mass-like factor m here has units of L^{-1}, but we keep it this way for simplicity; in normal units (reintroducing c and $\hbar$, the factor would be $mc/\hbar$.

Since we know how to differentiate $\partial_\mu\phi$, rather than the contravariant version, we rewrite the Lagrangian with the metric

$$\mathcal{L} = -\frac{1}{2}g^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\phi) - \frac{1}{2}m^2\phi^2 \tag{11.2}$$

The derivatives are then

$$\frac{\partial\mathcal{L}}{\partial\phi} = -m^2\phi \tag{11.3}$$

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} = -g^{\mu\nu}\partial_\nu\phi \tag{11.4}$$

The "total partial derivatives" will be simple, since the second formula above only involves $\partial\phi/\partial x^\mu$. So the total partial derivative is the same as the partial derivative with respect to spacetime labels. The Euler-Lagrange equation is then

$$0 = -\partial_\mu\partial^\mu\phi + m^2\phi \tag{11.5}$$

This is the Klein-Gordon equation for a single scalar field.

There are two ways to look at interpreting this equation. On the one hand, we can compare it with the Lagrangian density of the coupled harmonic oscillators,

$$\mathcal{L} = \frac{1}{2}\mu(\partial_t\phi)^2 - \frac{1}{2}Y(\partial_x\phi)^2 \tag{11.6}$$

in which we see something like the derivative term. The speed of a propagating wave must be something like Y/μ.

For the second term of the Klein-Gordon equation, an interpretation is suggested when we replace the partial derivatives with momentum operators

$$0 = (-i\partial_\mu)(-i\partial^\mu)\phi + m^2\phi = (P_\mu P^\mu + m^2)\phi \tag{11.7}$$

which looks like the "on-shell" condition of a particle with mass m. This is even more explicit when we plug in "free particle" solutions

$$\phi = Ae^{-ik_\mu x^\mu} \tag{11.8}$$

which results in the equation

$$0 = k_\mu k^\mu + m^2 \tag{11.9}$$

Classically, this is a travelling wave with an effective mass, *i.e.*, a certain resistance to changes in inertia.

Another type of solution can be found by postulating a spherical solution, such as one might have for emissions from a point source. It's instructive to look at the time-independent case:

$$\nabla^2\phi(\mathbf{x}) = m^2\phi(\mathbf{x}) \tag{11.10}$$

which we turn into a radial equation

$$\frac{1}{r}\frac{\partial^2}{\partial r^2}(r\phi(r)) = m^2\phi(r) \tag{11.11}$$

$$\frac{\partial^2}{\partial r^2}(r\phi(r)) = m^2(r\phi(r)) \tag{11.12}$$

$$\tag{11.13}$$

We postulate the form of the solution to be

$$r\phi(r) = Ae^{-mr} + Be^{mr} \tag{11.14}$$

Since we are interested in the source solution, we take $B = 0$, and we end up with

$$\phi(r) = A\frac{e^{-mr}}{r} \tag{11.15}$$

If this was a photon with $m = 0$, this looks like the Coulomb field. And indeed, if you want to add a Coulomb interaction to a Lagrangian, you add to it the product of the two fields: the field (or particle), and the Coulomb field.

If $m \neq 0$, then we have a field which is exponentially suppressed, with length scale m^{-1}: it's as if the non-zero mass limits the range of the interaction. This idea lies behind the Yukawa explanation of the short-range forces which keep the nucleus together. Since the nuclear length scale is on the order of 10^{-15} m, we get a mass on the order of a hundred or so MeV.

The pion, which was discovered soon afterwards, has a mass a little under 140 MeV. However, a cautionary tale: it wasn't the first. The muon (mass 105 MeV) was discovered first, and was hailed as the Yukawa meson. The problem was that it had *none* of the properties of a force-carrying particle: it was a fermion, and, worse yet, it didn't interact much with nucleons. In some older textbooks (not just those published before the π was discovered), the muon is still referred to as a μ-meson. The lesson is that one really has to check the properties of new particles, not just their masses. There is still a lot of work to do, for instance, in measuring the properties of the Higgs boson.

11.1.1 Complex-valued fields

If the field is complex-valued, we can describe two fields in a single Lagrangian.

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu \phi^*)(\partial^\mu \phi) - \frac{1}{2}m^2 \phi^* \phi \tag{11.16}$$

It turns out that we can treat a field and its complex conjugate as two independent fields. But how should we differentiate with respect to ϕ and ϕ^*? One way to see how to do this is to separate the two into independent real components:

$$\phi = u + iv \tag{11.17}$$

Then if we have a function $f(\phi, \phi^*)$ to extremize,

$$0 = \delta f = \frac{\partial f}{\partial \phi}(\delta u + i\delta v) + \frac{\partial f}{\partial \phi^*}(\delta u - i\delta v) \tag{11.18}$$

from which we conclude (since δu and δv are arbitrary),

$$\begin{aligned} 0 &= \frac{\partial f}{\partial \phi} + \frac{\partial f}{\partial \phi^*} \\ 0 &= \frac{\partial f}{\partial \phi} - \frac{\partial f}{\partial \phi^*} \end{aligned}$$

or, equivalently,

$$\begin{aligned} 0 &= \frac{\partial f}{\partial \phi} \\ 0 &= \frac{\partial f}{\partial \phi^*} \end{aligned}$$

Therefore we can proceed to differentiate with respect to ϕ and ϕ^* as if they were independent. The derivatives are then

$$\frac{\partial \mathcal{L}}{\partial \phi} = -\frac{1}{2}m^2\phi^* \tag{11.19}$$

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = -\frac{1}{2}g^{\mu\nu}\partial_\nu \phi^* \tag{11.20}$$

$$\frac{\partial \mathcal{L}}{\partial \phi^*} = -\frac{1}{2}m^2\phi \tag{11.21}$$

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^*)} = -\frac{1}{2}g^{\mu\nu}\partial_\nu \phi \tag{11.22}$$

and the Euler-Lagrange equations

$$0 = \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right) - \frac{\partial \mathcal{L}}{\partial \phi} \tag{11.23}$$

$$0 = -\frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\nu\phi^* + \frac{1}{2}m^2\phi^* \tag{11.24}$$

$$0 = -\partial_\mu\partial^\mu\phi^* + m^2\phi^* \tag{11.25}$$

Similarly,

$$0 = -\partial_\mu\partial^\mu\phi + m^2\phi \tag{11.26}$$

Thus we have two independent fields satisfying the same equation of motion and with the same mass.

11.2 Dirac equation

The next "simplest" (or at least next most fundamental) field should involve the members of the $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ representation. Their equation of motion is known as the Dirac equaiton.

What will be done here is in all likelihood profoundly unhistorical: Dirac was trying to figure out how to get rid of the second derivative in time implied by the Klein-Gordon equation, and when he found that he couldn't solve his new equation with single-valued functions, he postulated multi-valued functions—often written as 4-component column matrices—which were found to correspond to spin-1/2 particles.

Approaching from group theoretical considerations, one immediately "sees" a 4-component representation, because we have the direct sum of two 2-component spinors from the conjugate (and inequivalent) representations. In the Weyl (chiral) basis, we write a Dirac spinor as

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} \phi \\ \bar{\chi} \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \bar{\chi}^{\dot{1}} \\ \bar{\chi}^{\dot{2}} \end{pmatrix} \tag{11.27}$$

where we need to keep in mind that the ϕ, $\bar{\chi}$, ψ_L, and ψ_R are 2-component spinors. In the first version, we broke up ψ into left-handed and right-handed components. These correspond to the undotted and dotted spinors we see on the right side.

Then it's a matter of coming up with a suitable equation of motion which is first-order in time. We have the vector $\partial_\mu\phi$, but contracting it with itself will result in a second-order equation. Instead, we need to look at how to construct invariants with Dirac spinors. If we only had rotations to deal with, then a spinor $\eta^\dagger \cdot \eta$ would be fine, but it would not be invariant with respect to the parity transformation. Instead, we form the linear combination

$$(\bar{\chi})^\dagger \cdot \phi + \bar{\chi} \cdot \phi^\dagger \tag{11.28}$$

or, to make the left and right-handed explicit,

$$\psi_L^\dagger \cdot \psi_R + \psi_L \psi_R^\dagger \tag{11.29}$$

To write this in matrix form with the Dirac spinor itself, we define the matrix

$$\gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \tag{11.30}$$

and thus the invariant scalar is

$$\psi^\dagger \gamma^0 \psi = \psi_L^\dagger \cdot \psi_R + \psi_R^\dagger \cdot \psi_L \tag{11.31}$$

This will work for the mass term in the Lagrangian.

Notational caution: the adjoint of the Dirac spinor is also often designated with a bar

$$\bar{\psi} \equiv \psi^\dagger \gamma^0 \tag{11.32}$$

The reason it's useful here is that when we take derivatives of the Lagrangian, we'll take derivatives with respect to ψ and its (Dirac) adjoint $\bar{\psi}$. For this course, we'll just have to put up with this a short while.

For the kinetic term (which we want to keep linear in ∂_t), we form a scalar using the other Dirac matrices in the Weyl basis: $\bar{\psi}\gamma^\mu\partial_\mu\psi$, where

$$\gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix} \tag{11.33}$$

so the Dirac Lagrangian is

$$\begin{aligned} \mathcal{L} &= i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi & (11.34)\\ &= \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi & (11.35) \end{aligned}$$

The derivatives are then

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\bar{\psi}} &= i\gamma^\mu\partial_\mu\psi - m\psi & (11.36)\\ \frac{\partial\mathcal{L}}{\partial\psi} &= -m\bar{\psi} & (11.37)\\ \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)} &= i\bar{\psi}\gamma^\mu & (11.38) \end{aligned}$$

The first derivative gives

$$0 = i\gamma^\mu \partial_\mu \psi - m\psi \tag{11.39}$$

directly. This is the Dirac equation, the equation of motion for spin-1/2 particles. The other derivatives give

$$\begin{aligned} 0 &= \partial_\mu(i\bar{\psi}\gamma^\mu) + m\bar{\psi} && (11.40)\\ &= i\partial_\mu(\psi^\dagger\gamma^0\gamma^\mu) + m\psi^\dagger\gamma^0 && (11.41)\end{aligned}$$

which results in the Dirac equation for the adjoint field

$$0 = i\partial_\mu\bar{\psi}\gamma^\mu + m\bar{\psi} \tag{11.42}$$

We can confirm that solutions to the Dirac equation also satisfy the "on-shell" condition implicit in the Klein-Gordon equation by operating on the left of the Dirac equation by $(-i\gamma^\nu\partial_\nu - m)$:

$$\begin{aligned} 0 &= (-i\gamma^\nu\partial_\nu - m)(i\gamma^\mu\partial_\mu - m)\psi && (11.43)\\ &= (\gamma^\nu\gamma^\mu\partial_\nu\partial_\mu + m^2)\psi && (11.44)\end{aligned}$$

Since the sum takes in both $\gamma^\nu\gamma^\mu\partial_\nu\partial_\mu$ and $\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu$, we calculate the "anti-commutator"

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu \tag{11.45}$$

We can find by direct calculation that the anticommutator is zero when $\mu \neq \nu$, and for the others

$$\begin{aligned} \{\gamma^0, \gamma^0\} &= 2\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} = 2I && (11.46)\\ \{\gamma^i, \gamma^i\} &= -2\begin{pmatrix} -(\sigma^i)^2 & 0 \\ 0 & -(\sigma^i)^2 \end{pmatrix} = -2I && (11.47)\end{aligned}$$

which we summarize as

$$\{\gamma^\mu, \gamma^\nu\} = -2g^{\mu\nu} \tag{11.48}$$

(In fact, Dirac actually used this anti-commutation relation to *define* his matrices, rather than calculating them after the fact.) The left-multiplied Dirac equation is then

$$\begin{aligned} 0 &= (\frac{1}{2}\{\gamma^\mu, \gamma^\nu\}\partial_\mu\partial_\nu + m^2)\psi && (11.49)\\ &= (-g^{\mu\nu}\partial_\mu\partial_\nu + m^2)\psi && (11.50)\\ &= (-\partial_\mu\partial^\mu + m^2)\psi && (11.51)\end{aligned}$$

thus satisfying the Klein-Gordon equation.

11.3 Weyl equation

We can write the Dirac equation in terms of its 2-component parts:

$$0 = (i\gamma^\mu \partial_\mu - m)\psi = \begin{pmatrix} -m & i(\partial_0 + \boldsymbol{\sigma}\cdot\nabla) \\ i(\partial_0 - \boldsymbol{\sigma}\cdot\nabla) & -m \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{11.52}$$

We can see that with $m \neq 0$, the left and right-handed parts of the Dirac spinor get mixed up. However, in the massless limit, the equations decouple, and we get

$$0 = i(\partial_0 - \boldsymbol{\sigma}\cdot\nabla)\psi_L \tag{11.53}$$
$$0 = i(\partial_0 + \boldsymbol{\sigma}\cdot\nabla)\psi_R \tag{11.54}$$

which are the Weyl equations for massless chiral fermions. If we change these into operators,

$$0 = (E/c + \mathbf{s}\cdot\mathbf{p})\psi_L \tag{11.55}$$
$$0 = (E/c - \mathbf{s}\cdot\mathbf{p})\psi_R \tag{11.56}$$

which be rephrased as eigenvalue equations:

$$\mathbf{s}\cdot\mathbf{p}\psi_L = -(E/c)\psi_L \tag{11.57}$$
$$\mathbf{s}\cdot\mathbf{p}\psi_R = +(E/c)\psi_R \tag{11.58}$$

Since we're dealing with $m = 0$, then E is simply the magnitude of the momentum. We have here again the helicity

$$\lambda = \frac{\mathbf{s}\cdot\mathbf{p}}{|\mathbf{p}|} \tag{11.59}$$

The sign should (finally) make clear why we called one spinor "left-handed" and the other "right-handed".

Chapter 12

Electromagnetism

The equations of electromagnetism were the original relativistic field equations. As mentioned earlier, the electric and magnetic fields are part of the Maxwell field tensor $F^{\mu\nu}$, which comes from the $(0,1) \oplus (1,0)$ representation. To see how the field equations arise, however, it is useful to reintroduce the vector potential.

12.1 Revision: Maxwell's equations and potentials

Maxwell's equations without matter are as follows:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \tag{12.1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{12.2}$$

$$\nabla \wedge \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{12.3}$$

$$\nabla \wedge \mathbf{B} = \mu_0 \mathbf{j} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{12.4}$$

The Lorentz force is

$$\mathbf{f} = q(\mathbf{E} + \mathbf{v} \wedge \mathbf{B}) \tag{12.5}$$

A note on units: it's true that MKS units are probably the most useful for engineers. On the other hand, we're not engineers. We're trying to understand the structure of the theory in order to see if there are further insights. We'll try to keep with MKS units, but also to keep this all in perspective. As one senior Oxford physicist has said, "All systems of units are absurd".

12.2 *Electromagnetic potential as a 4-vector

The equation $\nabla \cdot \mathbf{B} = 0$ indicates that $\mathbf{B}$ can be written in terms of a 3-vector potential

$$\mathbf{B} = \nabla \wedge \mathbf{A} \tag{12.6}$$

We can plug this into the equation for the electric field

$$\begin{aligned}
\nabla \wedge \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} && (12.7)\\
&= -\frac{\partial}{\partial t}(\nabla \wedge \mathbf{A}) && (12.8)\\
&= -\nabla \wedge \left(\frac{\partial \mathbf{A}}{\partial t}\right) && (12.9)\\
0 &= \nabla \wedge \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right) && (12.10)
\end{aligned}$$

The last equation then indicates that we can introduce another potential function

$$\begin{aligned}
\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} &= -\nabla\phi && (12.11)\\
\mathbf{E} &= -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} && (12.12)
\end{aligned}$$

With two equations automatically satisfied, we can write the $\nabla \cdot \mathbf{E}$ equation in the form

$$\begin{aligned}
\frac{\rho}{\epsilon_0} &= \nabla \cdot \mathbf{E} && (12.13)\\
&= \nabla \cdot \left(-\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}\right) && (12.14)\\
-\frac{\rho}{\epsilon_0} &= \nabla^2\phi + \frac{\partial}{\partial t}(\nabla \cdot \mathbf{A}) && (12.15)
\end{aligned}$$

The last Maxwell equation is then

$$\begin{aligned}
\nabla \wedge \mathbf{B} &= \mu_0 \mathbf{j} + \mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} && (12.16)\\
\nabla \wedge (\nabla \wedge \mathbf{A}) &= \mu_0 \mathbf{j} + \frac{1}{c^2}\frac{\partial}{\partial t}\left(-\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}\right) && (12.17)\\
\nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} &= \mu_0 \mathbf{j} - \frac{1}{c^2}\frac{\partial}{\partial t}\nabla\phi - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} && (12.18)\\
-\mu_0 \mathbf{j} &= \nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla\left(\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t}\right) && (12.19)
\end{aligned}$$

These equations still seem rather complicated, though we can make them look more similar by adding/subtracting a derivative of ϕ to the earlier equation:

$$-\frac{\rho}{\epsilon_0} = \nabla^2\phi - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} + \frac{\partial}{\partial t}\left(\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t}\right) \quad (12.20)$$

We can decouple the equations by taking advantage of an ambiguity in the definitions of the potentials: we can modify them by a "gauge transformation"

$$\begin{aligned}
\mathbf{A} &\rightarrow \mathbf{A} + \nabla\chi && (12.21)\\
\phi &\rightarrow \phi - \frac{\partial \chi}{\partial t} && (12.22)
\end{aligned}$$

where χ is a single-valued function of position and time. We start by writing

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = f(x^\mu) \tag{12.23}$$

We then apply the gauge transformation above, which gives us

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} + \nabla^2 \chi - \frac{1}{c^2}\frac{\partial^2 \chi}{\partial t^2} = f(x^\mu) \tag{12.24}$$

so if we choose χ such that

$$\nabla^2 \chi - \frac{1}{c^2}\frac{\partial^2 \chi}{\partial t^2} = f(x^\mu) \tag{12.25}$$

(which can always be done), then we have transformed the potentials such that they satisfy the "Lorenz gauge condition":

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0 \tag{12.26}$$

This decouples the remaining two Maxwell equations to give

$$\nabla^2 \phi - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\varepsilon_0} \tag{12.27}$$

$$\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{j} \tag{12.28}$$

These are the wave equations for electromagnetic radiation due to a source.

The structure of these equations suggest something else: we have two more 4-vectors representing the potential and the current

$$A^\mu = (\phi/c, \mathbf{A}) \tag{12.29}$$

$$J^\mu = (\rho c, \mathbf{j}) \tag{12.30}$$

Current conservation can be written

$$0 = \nabla \cdot \mathbf{j} - \frac{\partial \rho}{\partial t} = \partial_\mu J^\mu \tag{12.31}$$

The Lorenz gauge condition can be written

$$\partial_\mu A^\mu = 0 \tag{12.32}$$

which is actually more natural since one doesn't transform either $\mathbf{A}$ or ϕ, but both simultaneously. The gauge transformation itself can be written in 4-vector form

$$A^\mu \to A^\mu + \partial^\mu \chi \tag{12.33}$$

(recall that $\partial^0 = -\partial_t$). The equations of motion can be written

$$\partial_\mu \partial^\mu A^\nu = -\mu_0 J^\nu \tag{12.34}$$

To get the Maxwell field tensor in terms of the potential, we write

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{12.35}$$

It is not difficult to confirm that this gives the **E** and **B** we saw earlier. The equation of motion in terms of $F^{\mu\nu}$ is then

$$\partial_\mu F^{\mu\nu} = \partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = \partial_\mu \partial^\mu A^\nu - \partial^\nu(\partial_\mu A^\mu) = -\mu_0 J^\nu \tag{12.36}$$

The parenthesis near the last equality disappears in the Lorenz gauge, and one gets the wave equations back.

12.3 *Gauge invariance

"Gauge invariance" refers to the fact that the choice of gauge shouldn't affect any physics. For the purpose of problem solving, this gives us a good excuse to choose whatever gauge is simplest for the problem. Choosing the Lorenz gauge condition to decouple the wave equations is one example.

Besides the Lorenz gauge, another gauge condition has historically been very useful:

$$\nabla \cdot \mathbf{A} = 0 \tag{12.37}$$

This is sometimes called the "Coulomb gauge", and normally it would be frame-dependent. However, it also doesn't fix the gauge entirely, and one can then add the condition that ϕ is independent of time. This would then satisfy both the Coulomb and Lorenz gauge conditions simultaneously.

Another aspect of invariance, which shouldn't surprise you by now, is that there are restrictions on what can be physical, *i.e.*, not everything we can write down in mathematical form is valid for physics. For instance, one might add a "mass" term (analogous to $m^2\phi^2$ in the Klein-Gordon equation) to a Lagrangian along the lines of

$$m^2 A_\mu A^\mu \tag{12.38}$$

but a gauge transformation would leave additional terms

$$\begin{aligned} A_\mu A^\mu \to A'_\mu A'^\mu &= (A_\mu + \partial_\mu \chi)(A^\mu + \partial^\mu \chi) && (12.39)\\ &= A_\mu A^\mu + 2A_\mu \partial^\mu \chi + (\partial_\mu \chi)(\partial^\mu \chi) && (12.40) \end{aligned}$$

which would enter into the equations of motion. On the other hand,

$$\begin{aligned} F^{\mu\nu} \to F'^{\mu\nu} &= \partial^\mu A'^\nu - \partial^\nu A'^\mu && (12.41)\\ &= \partial^\mu A^\nu + \partial^\mu \partial^\nu \chi - \partial^\nu A^\mu - \partial^\nu \partial^\mu \chi && (12.42)\\ &= F^{\mu\nu} && (12.43) \end{aligned}$$

so a Lagrangian could be made from $F^{\mu\nu}$, as long as one could make a scalar out of it.

12.4 Lagrangian for em fields, equations of motion

The only proper scalar which can be obtained from the Maxwell field tensor is $F_{\mu\nu}F^{\mu\nu}$.

$$F^{\mu\nu} \equiv \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{pmatrix} \tag{12.44}$$

$$F_{\mu\nu} = g_{\mu\kappa}g_{\nu\lambda}F^{\kappa\lambda} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{pmatrix} \tag{12.45}$$

$$\tag{12.46}$$

The scalar contraction amounts to the sum of the product of every element with its corresponding element in the other matrix.

$$\frac{1}{2}F_{\mu\nu}F^{\mu\nu} = \mathbf{B}^2 - \frac{\mathbf{E}^2}{c^2} \tag{12.47}$$

There is also a pseudo-scalar which we can obtain using the "dual" Maxwell tensor:

$$\bar{F}^{\mu\nu} \equiv \frac{1}{2}\epsilon^{\mu\nu\kappa\lambda}F_{\kappa\lambda} = \begin{pmatrix} 0 & B_x & B_y & B_z \\ -B_x & 0 & -E_z/c & E_y/c \\ -B_y & E_z/c & 0 & -E_x/c \\ -B_z & -E_y/c & E_x/c & 0 \end{pmatrix} \tag{12.48}$$

The dual tensor swaps $\mathbf{E} \to \mathbf{B}$ and $\mathbf{B} \to -\mathbf{E}$. The pseudo-scalar is then

$$F_{\mu\nu}\bar{F}^{\mu\nu} = -\frac{4\mathbf{E}\cdot\mathbf{B}}{c} \tag{12.49}$$

We expect that the Lagrangian can be quadratic in $\partial_\mu A^\nu$. Gauge invariance implies that it can only include such derivatives in the combinations contained in $F^{\mu\nu}$. So we write down the only proper scalar

$$\mathcal{L} = -\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu} \tag{12.50}$$

$$= -\frac{1}{4\mu_0}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) \tag{12.51}$$

$$= -\frac{1}{4\mu_0}(\partial_\mu A_\nu - \partial_\nu A_\mu)g^{\mu\kappa}g^{\nu\lambda}(\partial_\kappa A_\lambda - \partial_\lambda A_\kappa) \tag{12.52}$$

$$\tag{12.53}$$

The derivatives are then

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial(\partial_\beta A_\alpha)} &= -\frac{1}{4\mu_0} g^{\mu\kappa} g^{\nu\lambda} (\delta^\beta_\mu \delta^\alpha_\nu - \delta^\beta_\nu \delta^\alpha_\mu)(\partial_\kappa A_\lambda - \partial_\lambda A_\kappa) && (12.54)\\
&\quad -\frac{1}{4\mu_0} (\partial_\mu A_\nu - \partial_\nu A_\mu) g^{\mu\kappa} g^{\nu\lambda} (\delta^\beta_\kappa \delta^\alpha_\lambda - \delta^\beta_\lambda \delta^\alpha_\kappa) && (12.55)\\
&= -\frac{1}{4\mu_0} (g^{\beta\kappa} g^{\alpha\lambda} - g^{\alpha\kappa} g^{\beta\lambda}) F_{\kappa\lambda} - \frac{1}{4\mu_0} F_{\mu\nu} (g^{\mu\beta} g^{\nu\alpha} - g^{\mu\alpha} g^{\nu\beta}) && (12.56)\\
&= -\frac{1}{4\mu_0} (F^{\beta\alpha} - F^{\alpha\beta}) - \frac{1}{4\mu_0} (F^{\beta\alpha} - F^{\alpha\beta}) && (12.57)\\
&= -\frac{1}{2\mu_0} (F^{\beta\alpha} - F^{\alpha\beta}) && (12.58)\\
&= \frac{1}{\mu_0} F^{\alpha\beta} && (12.59)
\end{aligned}$$

and the Euler-Lagrange equations

$$0 = \partial_\beta \left(\frac{\partial \mathcal{L}}{\partial(\partial_\beta A_\alpha)} \right) = \frac{1}{\mu_0} \partial_\beta F^{\alpha\beta} \tag{12.60}$$

for a source-free field.

To introduces sources, we need to add a term to the Lagrangian which contracts J^μ with the vector field A_μ:

$$\mathcal{L} = -\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} + J^\mu A_\mu \tag{12.61}$$

This Lagrangian is not itself gauge invariant, but we can plow ahead assuming we can fix it up later. (Goldstein, following Dirac, calls this a "weak" condition, and in a sense is like varying $\mathsf{U} \cdot \mathsf{U}$ even while knowing that in the end it should end up as $-c^2$.) The additional derivative is then

$$\frac{\partial \mathcal{L}}{\partial A_\alpha} = J^\alpha \tag{12.62}$$

resulting in

$$\partial_\beta F^{\alpha\beta} = \mu_0 J^\alpha \tag{12.63}$$

which is "fixed up" as far as gauge invariance is concerned.

It's also easy to verify for μ, ν, and λ all different,

$$\partial_\lambda F_{\mu\nu} + \partial_\nu F_{\lambda\mu} + \partial_\mu F_{\nu\lambda} = 0 \tag{12.64}$$

since the derivative operators commute. This gives the Maxwell equations such as $\nabla \cdot \mathbf{B} = 0$ which have been automatically satisfied by introducing the potential.

12.4.1 Use of invariants

As before, invariants $\mathbf{E}^2/c^2 - \mathbf{B}^2$ and $\mathbf{E} \cdot \mathbf{B}$ allow us to make some frame-independent observations.

- If the $\mathbf{E}$ and $\mathbf{B}$ are perpendicular in some frame ($\mathbf{E}\cdot\mathbf{B}=0$), they are perpendicular in all frames. This is particularly relevant for electromagnetic radiation.
- If the magnitudes of the fields are the same in one frame, *i.e.*, $\mathbf{E}^2/c^2-\mathbf{B}^2=0$, they are the same in all frames.
- If $\mathbf{E}^2/c^2-\mathbf{B}^2\neq 0$, it is possible to choose a frame in which one of them is zero.

Let's examine the case where the fields are perpendicular. How do we choose a frame in which one field disappears?

If $\mathbf{E}^2/c^2-\mathbf{B}^2>0$, we'd like to choose a frame with $\mathbf{B}=0$. Since

$$\mathbf{B}'_{\parallel}=\mathbf{B}_{\parallel} \tag{12.65}$$

we need to choose a frame in which $\mathbf{B}_{\parallel}=0$, *i.e.*, the component parallel to $\mathbf{v}$ is zero. This is easy: just choose $\mathbf{v}$ to be perpendicular to $\mathbf{B}$. Then we have for the perpendicular component

$$\mathbf{B}'_{\perp}=\gamma\left(\mathbf{B}_{\perp}-\frac{\mathbf{v}\wedge\mathbf{E}}{c^2}\right) \tag{12.66}$$

To make this zero, we want

$$\mathbf{B}=\frac{\mathbf{v}\wedge\mathbf{E}}{c^2} \tag{12.67}$$

We already have $\mathbf{v}$ perpendicular to $\mathbf{B}$, and $\mathbf{E}$ also perpendicular to $\mathbf{B}$. We can choose $\mathbf{v}$ to be perpendicular to $\mathbf{E}$ as well. Therefore $\mathbf{v}$ must be proportional to $\mathbf{E}\wedge\mathbf{B}$:

$$\begin{aligned}
\mathbf{v} &= k(\mathbf{E}\wedge\mathbf{B}) && (12.68)\\
\mathbf{B} &= \frac{\mathbf{v}\wedge\mathbf{E}}{c^2}=\frac{k}{c^2}(\mathbf{E}\wedge\mathbf{B})\wedge\mathbf{E} && (12.69)\\
&= \frac{k}{c^2}(\epsilon_{ijk}E_jB_k)\epsilon_{aib}E_b && (12.70)\\
&= -\frac{k}{c^2}(\epsilon_{ijk}\epsilon_{iab})E_jB_kE_b && (12.71)\\
&= -\frac{k}{c^2}(\delta_{ja}\delta_{kb}-\delta_{jb}\delta_{ka})E_jB_kE_b && (12.72)\\
&= -\frac{k}{c^2}(\mathbf{E}(\mathbf{B}\cdot\mathbf{E})-(\mathbf{E}\cdot\mathbf{E})\mathbf{B}) && (12.73)\\
&= \frac{k}{c^2}E^2\mathbf{B} && (12.74)\\
\mathbf{v} &= \frac{c^2}{E^2}(\mathbf{E}\wedge\mathbf{B}) && (12.75)
\end{aligned}$$

Then we want to calculate the new electric fields. We have

$$\mathbf{E}'_{\parallel}=\mathbf{E}_{\parallel}=0 \tag{12.76}$$

because we chose $\mathbf{v}\perp\mathbf{E}$. The perpendicular component is then the whole electric field $\mathbf{E}$:

$$\mathbf{E}'=\gamma(\mathbf{E}+\mathbf{v}\wedge\mathbf{B}) \tag{12.77}$$

To calculate the cross product, we note that because the three vectors are mutually perpendicular, $\mathbf{E}$ must be proportional to $\mathbf{v} \wedge \mathbf{B}$.

$$\begin{aligned} \lambda \mathbf{E} &= \mathbf{v} \wedge \mathbf{B} && (12.78)\\ \lambda \mathbf{E} \cdot \mathbf{E} &= \mathbf{E} \cdot (\mathbf{v} \wedge \mathbf{B}) && (12.79)\\ \lambda E^2 &= \mathbf{v} \cdot (\mathbf{B} \wedge \mathbf{E}) && (12.80)\\ &= -\mathbf{v} \cdot (\mathbf{E} \wedge \mathbf{B}) && (12.81)\\ \lambda &= -\mathbf{v} \cdot \frac{\mathbf{E} \wedge \mathbf{B}}{E^2} = -\frac{\mathbf{v} \cdot \mathbf{v}}{c^2} = -\beta^2 && (12.82) \end{aligned}$$

Therefore

$$\mathbf{E}' = \gamma(\mathbf{E} + \lambda \mathbf{E}) = \gamma \mathbf{E}(1 - \beta^2) = \mathbf{E}/\gamma \tag{12.83}$$

Similarly, if $\mathbf{E}^2/c^2 - \mathbf{B}^2 < 0$, then we choose a frame moving with velocity

$$\mathbf{v} = \frac{\mathbf{E} \wedge \mathbf{B}}{B^2} \tag{12.84}$$

in which case the transformed fields are

$$\begin{aligned} \mathbf{E}' &= 0 && (12.85)\\ \mathbf{B}' &= \mathbf{B}/\gamma && (12.86) \end{aligned}$$

12.4.2 Motion in an electromagnetic field

The Lagrangian for a free single particle can be written as

$$L = -mc(-U_\mu U^\mu)^{1/2} \tag{12.87}$$

How do we add interactions with the field?

We can start from the Lagrangian density we wrote down earlier

$$\mathcal{L} = -\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} + J^\mu A_\mu \tag{12.88}$$

and insert for the current density a moving point charge

$$\begin{aligned} \rho(\mathbf{r}) &= q\delta(\mathbf{r} - \mathbf{s}) && (12.89)\\ \mathbf{j}(\mathbf{r}) &= q\delta(\mathbf{r} - \mathbf{s})\mathbf{v}(\mathbf{r}) && (12.90) \end{aligned}$$

The Lagrangian (not the density this time) is

$$\begin{aligned} L &= \int d^3x \left(-\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} - \rho\phi + \mathbf{j} \cdot \mathbf{A} \right) && (12.91)\\ &= \int d^3x \left(-\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} \right) - q\phi + q\mathbf{v} \cdot \mathbf{A} && (12.92)\\ &= \int d^3x \left(-\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} \right) + qU_\mu A^\mu && (12.93) \end{aligned}$$

This suggests that for a single particle in an electromagnetic field,

$$L = -mc(-U_\mu U^\mu)^{1/2} + qU_\mu A^\mu + \int d^3x \left(-\frac{1}{4\mu_0} F_{\mu\nu}F^{\mu\nu}\right) \tag{12.94}$$

which combines both discrete and continuous systems in one common Lagrangian.

If we concentrate on the single-particle coordinates, we have for the derivatives

$$\begin{aligned} L &= -mc(-g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu)^{1/2} + q\dot{x}^\mu A_\mu & (12.95)\\ \frac{\partial L}{\partial \dot{x}^\alpha} &= mc(-g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu)^{-1/2} g_{\alpha\beta}\dot{x}^\beta + qA_\alpha & (12.96)\\ \frac{\partial L}{\partial x^\alpha} &= q\dot{x}^\mu \frac{\partial A_\mu}{\partial x^\alpha} & (12.97) \end{aligned}$$

where we've used the dot to indicate a full derivative with respect to proper time τ. Combined (and taking the full derivative of $A_\mu(x^\nu)$), we get

$$\begin{aligned} 0 &= \frac{d}{d\tau}\left(m\frac{dx_\mu}{d\tau} + qA_\mu\right) - q\frac{dx^\nu}{d\tau}\frac{\partial A_\nu}{\partial x^\mu} & (12.98)\\ &= m\frac{d^2x_\mu}{d\tau^2} + q\frac{dx^\nu}{d\tau}\frac{\partial A_\mu}{\partial x^\nu} - q\frac{dx^\nu}{d\tau}\frac{\partial A_\nu}{\partial x^\mu} & (12.99)\\ &= m\frac{d^2x_\mu}{d\tau^2} + qU^\nu(\partial_\nu A_\mu - \partial_\mu A_\nu) & (12.100)\\ m\frac{dU_\mu}{d\tau} &= qU^\nu F_{\mu\nu} & (12.101) \end{aligned}$$

We extract the spatial part:

$$\begin{aligned} \gamma m\frac{du_j}{dt} &= q(\gamma c F_{j0} + \gamma u^k F_{jk}) & (12.102)\\ m\frac{du_j}{dt} &= q(cF_{j0} + u^k F_{jk}) & (12.103) \end{aligned}$$

Checking $j = 1$ for the x part,

$$f_x = q(E_x + v_y F_{12} + v_z F_{13}) = q(E_x + v_y B_z - v_z B_y) \tag{12.104}$$

which is one part of the familiar vector equation

$$\mathbf{f} = q(\mathbf{E} + \mathbf{v} \wedge \mathbf{B}) \tag{12.105}$$

And thus we see that all the electromagnetic relations are aspects of a properly Lorentz-invariant theory.

Chapter 13

Radiation

In this section, we'll look at some results concerning the fields themselves.

13.1 Conservation of energy

One can derive a continuity equation on the fields directly from the Maxwell equations. We start with the curl equations:

$$0 = \nabla \wedge \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} \tag{13.1}$$

$$0 = \nabla \wedge \mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} - \mu_0 \mathbf{j} \tag{13.2}$$

We take dot products of these two equations and subtract them to obtain

$$0 = \mathbf{B} \cdot (\nabla \wedge \mathbf{E}) + \mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} - \mathbf{E} \cdot (\nabla \wedge \mathbf{B}) + \frac{1}{c^2}\mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{E} \cdot \mathbf{j} \tag{13.3}$$

To simplify this, we note that

$$\begin{aligned}
\nabla \cdot (\mathbf{E} \wedge \mathbf{B}) &= \partial_i \epsilon_{ijk} E_j B_k && (13.4)\\
&= \epsilon_{ijk} \partial_i (E_j B_k) && (13.5)\\
&= \epsilon_{ijk} (E_j \partial_i B_k + B_k \partial_i E_j) && (13.6)\\
&= -\epsilon_{ijk} E_j \partial_i B_k + \epsilon_{kij} B_k \partial_i E_j && (13.7)\\
&= -\mathbf{E} \cdot (\nabla \wedge \mathbf{B}) + \mathbf{B} \cdot (\nabla \wedge \mathbf{E}) && (13.8)
\end{aligned}$$

We can also simplify the time derivatives

$$\frac{\partial B^2}{\partial t} = \frac{\partial}{\partial t}\mathbf{B} \cdot \mathbf{B} = 2\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} \tag{13.9}$$

and similarly for $\mathbf{E}$. Using these, we obtain the equation

$$0 = \nabla \cdot (\mathbf{E} \wedge \mathbf{B}) + \frac{1}{2}\frac{\partial B^2}{\partial t^2} + \frac{1}{2c^2}\frac{\partial E^2}{\partial t^2} + \mu_0 \mathbf{j} \cdot \mathbf{E} \tag{13.10}$$

$$-\mathbf{j} \cdot \mathbf{E} = \nabla \cdot \left(\frac{1}{\mu_0}\mathbf{E} \wedge \mathbf{B}\right) + \frac{\partial}{\partial t}\left(\frac{1}{2\mu_0}(B^2 + \frac{E^2}{c^2})\right) \tag{13.11}$$

The quantity

$$u \equiv \frac{1}{2\mu_0}\left(B^2 + \frac{E^2}{c^2}\right) \tag{13.12}$$

is the energy density of the electromagnetic field. The vector

$$\mathbf{N} \equiv \frac{1}{\mu_0}\mathbf{E} \wedge \mathbf{B} \tag{13.13}$$

is known as Poynting's vector, and represents energy flow carried by the field. The continuity equation is then

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{N} = -\mathbf{j} \cdot \mathbf{E} \tag{13.14}$$

which says that the energy lost in the fields is due to the work done by the electric field on sources. In the absence of sources, energy is conserved, flowing in and out of a volume like a fluid.

13.2 Plane waves in vacuum

This suggests that the absence of sources doesn't imply a trivial solution to the Maxwell equations. There could be propagating waves which carry energy and momentum.

For convenience, we'll use complex forms for the planes waves in vacuum:

$$\mathbf{E} = \mathbf{E}_0 e^{i\mathsf{K}\cdot\mathsf{X}} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \tag{13.15}$$

$$\mathbf{B} = \mathbf{B}_0 e^{i\mathsf{K}\cdot\mathsf{X}} = \mathbf{B}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \tag{13.16}$$

To get the physical fields, we take the real parts of the field components.

If $\mathbf{E}_0$ and $\mathbf{B}_0$ are both real, we have linearly polarized light. If $\mathbf{E}_0$ and $\mathbf{B}_0$ are themselves complex, we could form any polarization, subject to the Maxwell equations.

To find out how the fields are related, we plug into the vacuum equations:

$$i\mathbf{k} \cdot \mathbf{E}_0 = 0 \tag{13.17}$$

$$i\mathbf{k} \cdot \mathbf{B}_0 = 0 \tag{13.18}$$

$$i\mathbf{k} \wedge \mathbf{E}_0 = i\omega\mathbf{B}_0 \tag{13.19}$$

$$i\mathbf{k} \wedge \mathbf{B}_0 = -i\omega\mathbf{E}_0/c^2 \tag{13.20}$$

The first two equations tell us that $\mathbf{E}$ and $\mathbf{B}$ are orthogonal to the wave direction $\mathbf{k}$, *i.e.*, if the plane waves are to satisfy the Maxwell equations, they have to be orthogonal. Moreover, the third equation tells us that $\mathbf{E}$ and $\mathbf{B}$ are also mutually orthogonal.

We also have

$$E_0 = cB_0 \tag{13.21}$$

as long as $\omega = kc$, which is fine for vacuum.

Let's now look at this in terms of the 4-potential. We saw earlier that each component satisfies a wave equation

$$\partial_\mu \partial^\mu A^\nu = 0 \tag{13.22}$$

We can write the solution

$$A^\mu = a\epsilon^\mu e^{ik_\nu x^\nu} \tag{13.23}$$

where a is a constant, and ϵ^μ is the polarization vector. This appears to have 4 degrees of freedom, but we expect only two. What are the other constraints?

One comes from the Lorenz gauge condition itself.

$$0 = \partial_\mu A^\mu = a\epsilon^\mu i k_\mu e^{ik_\nu x^\nu} \tag{13.24}$$

which implies that

$$k_\mu \epsilon^\mu = 0 \tag{13.25}$$

Since k_μ is a given vector, this fixes one of the ϵ^μ components in terms of the others.

We can make one more choice under the Lorenz condition. In particular, we can always choose

$$\phi/c = A^0 = 0 \tag{13.26}$$

The way we do this is to set

$$\chi = \int \phi dt \tag{13.27}$$

and use this to modify a potential which already satisfies the Lorenz condition. The gauge transformation is

$$A^\mu \to A^\mu + \partial^\mu \chi \tag{13.28}$$

which when separated (remember this is ∂^μ, not ∂_μ, so there's a sign change) is

$$\phi \to \phi' = \phi - \frac{\partial \chi}{\partial t} \tag{13.29}$$

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla\chi \tag{13.30}$$

Let's plug in to see if still satisfies the condition:

$$\partial_\mu A'^\mu = \frac{1}{c^2}\frac{\partial \phi'}{\partial t} + \nabla \cdot \mathbf{A}' \tag{13.31}$$

$$= \frac{1}{c^2}\frac{\partial}{\partial t}\left(\phi - \frac{\partial}{\partial t}\int \phi dt\right) + \nabla \cdot \left(\mathbf{A} + \nabla \int \phi dt\right) \tag{13.32}$$

$$= \left(\frac{1}{c^2}\frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A}\right) + \int dt \left(-\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} + \nabla^2 \phi\right) \tag{13.33}$$

The first parenthesis is just the Lorenz condition on the original potential, and the second just the wave equation, which ϕ already satisfied. So both are zero, and thus the transformed potential still satisfies the gauge condition.

This means we can always fix $\epsilon^0 = 0$ (since we have set $A^0 = 0$ via the second constraint). Combined with the first constraint, we obtain

$$0 = k_\mu \epsilon^\mu = \mathbf{k} \cdot \boldsymbol{\epsilon} \tag{13.34}$$

which implies that the 3-polarization is always transverse to $\mathbf{k}$, *i.e.*, the field is always polarized transverse to the direction of travel.

In terms of the 4-potential, the magnetic field is

$$\mathbf{B} = \nabla \wedge \mathbf{A} = i\mathbf{k} \wedge \mathbf{A} \tag{13.35}$$

which implies that $\mathbf{B}$ is perpendicular to $\mathbf{A}$ as well as $\mathbf{k}$. (Note that the elements of $\mathbf{A}$ can be complex, so the i indicates the phase relationship.) The electric field is

$$\begin{aligned} \mathbf{E} &= -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} && (13.36) \\ &= -a\epsilon^0 i\mathbf{k} e^{ik_\nu x^\nu} - a\boldsymbol{\epsilon} i k_0 e^{ik_\nu x^\nu} && (13.37) \end{aligned}$$

Since we can set $\epsilon^0 = 0$, we are left with

$$\mathbf{E} = -i\frac{\omega}{c}\mathbf{A} \tag{13.38}$$

which means $\mathbf{E}$ is parallel to $\mathbf{A}$.

Chapter 14

Fields with sources

In this section, we'll start looking at solutions to the Maxwell equations for a given source.

14.1 *Fields of a uniformly moving charge

We want to obtain the force due to a charge moving with a constant velocity $\mathbf{v}$ in frame S. Let's say it is moving in the positive $\hat{\mathbf{x}}$ direction.

An easy way to start is to consider it in a frame S' in which the charge is at rest at the origin. This field is well known:

$$\mathbf{E}' = \frac{q}{4\pi\varepsilon_0 r'^3}\mathbf{r}' \tag{14.1}$$

$$\mathbf{B}' = 0 \tag{14.2}$$

Then we transform the fields (inverting the Lorentz transform, so the signs are flipped):

$$E_x = E'_x = \frac{q}{4\pi\varepsilon_0}\frac{x'}{r'^3} \tag{14.3}$$

$$E_y = \gamma(E'_y - (\mathbf{v}\wedge\mathbf{B}')_y) = \frac{\gamma q}{4\pi\varepsilon_0}\frac{y'}{r'^3} \tag{14.4}$$

But we also want the fields in terms of quantities in S:

$$r'^2 = x'^2 + y'^2 + z'^2 = \gamma^2(x-vt)^2 + y^2 + z^2 \tag{14.5}$$

We can choose to define $t = 0$ as the time of observation, with the origins in the two frames coinciding, leaving us with

$$\mathbf{E} = \frac{\gamma q\mathbf{r}}{4\pi\varepsilon_0(\gamma^2x^2 + y^2 + z^2)^{3/2}} \tag{14.6}$$

The magnetic field in S' was zero, so we get in S

$$\mathbf{B}_\parallel = \mathbf{B}'_\parallel = 0 \tag{14.7}$$

$$\mathbf{B}_\perp = \gamma(\mathbf{B}'_\perp + \frac{\mathbf{v}\wedge\mathbf{E}'}{c^2}) = \gamma\frac{\mathbf{v}\wedge\mathbf{E}'_\perp}{c^2} = \frac{\mathbf{v}\wedge\mathbf{E}_\perp}{c^2} = \frac{\mathbf{v}\wedge\mathbf{E}}{c^2} \tag{14.8}$$

where the last equality uses the fact that

$$\mathbf{v} \wedge \mathbf{E} = \mathbf{v} \wedge (\mathbf{E}_\perp + \mathbf{E}_\parallel) \tag{14.9}$$

but $\mathbf{v} \wedge \mathbf{E}_\parallel = 0$.

Some observations, just to get a picture of what the fields look like:

- The magnetic fields in S circulates around the charge, as one might have expected from the Biot-Savart law. There is no magnetic field along the direction of motion.
- The electric fields in S are "flattened" transverse to the direction of motion. The transverse fields are increased by γ, while they are decreased by γ in the direction of motion. The reason for the latter is length contraction along the direction of motion, rather than the fields changing (which doesn't happen for $\mathbf{E}_\parallel$).
- The electric field lines remain radial, rather than reflecting the position of the charge when the field was "emitted". This reflects uniform motion.

In terms of 4-potentials, we have

$$\begin{aligned} \phi' &= \frac{q}{4\pi\varepsilon_0 r'} && (14.10)\\ \mathbf{A}' &= 0 && (14.11) \end{aligned}$$

In frame S, the position of the charge as a function of time is

$$\mathbf{r}_p = (vt, 0, 0) \tag{14.12}$$

So the potentials in S are

$$\begin{aligned} r' &= [\gamma^2(x-vt)^2 + y^2 + z^2]^{1/2} && (14.13)\\ \phi &= \gamma(\phi' + \beta A'_x) = \frac{\gamma q}{4\pi\varepsilon_0[\gamma^2(x-vt)^2 + y^2 + z^2]^{1/2}} && (14.14)\\ A_x &= \gamma(A'_x + \beta\phi') = \beta\phi && (14.15)\\ A_y &= A_z = 0 && (14.16) \end{aligned}$$

which gives us fields as we saw before.

14.2 *Retarded potentials

Consider a system of charges and currents which are varying in time. Since these can always be analyzed with Fourier transformations, we can restrict our study to sinusoidal variations. The total behavior is then the inverse Fourier transform.

$$\begin{aligned} \rho(x^\mu) &= \rho(\mathbf{x})e^{-i\omega t} && (14.17)\\ \mathbf{J}(x^\mu) &= \mathbf{J}(\mathbf{x})e^{-i\omega t} && (14.18) \end{aligned}$$

The resulting 4-potential satisfies the equation

$$\partial_\mu \partial^\mu A^\nu = -\mu_0 J^\nu \tag{14.19}$$

Of course, as we saw before, one can also add source-free solutions, but here we're interested in the effects of the sources.

To simplify the notation somewhat, we consider just one component of the wave equation in the form

$$\nabla^2 \psi - \frac{1}{c^2}\frac{\partial^2 \psi}{\partial t^2} = -f(t, \mathbf{x}) \tag{14.20}$$

We can write the field and the source in terms of their Fourier components:

$$\psi(t, \mathbf{x}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \psi(\omega, \mathbf{x}) e^{-i\omega t} d\omega \tag{14.21}$$

$$f(t, \mathbf{x}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\omega, \mathbf{x}) e^{-i\omega t} d\omega \tag{14.22}$$

Substituting this into the wave equation,

$$\nabla^2 \psi = \frac{1}{2\pi}\int_{-\infty}^{\infty} \nabla^2 \psi(\omega, \mathbf{x}) e^{-i\omega t} d\omega \tag{14.23}$$

$$-\frac{1}{c^2}\frac{\partial^2 \psi}{\partial t^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \psi(\omega, \mathbf{x}) \left(\frac{\omega^2}{c^2}\right) e^{-i\omega t} d\omega \tag{14.24}$$

which results in

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} (\nabla^2 + k^2)\psi(\omega, \mathbf{x}) e^{-i\omega t} d\omega = -\frac{1}{2\pi}\int_{-\infty}^{\infty} f(\omega, \mathbf{x}) e^{-i\omega t} d\omega \tag{14.25}$$

where $k = \omega/c$. We thus obtain the inhomogeneous Helmholtz equation for each frequency component

$$(\nabla^2 + k^2)\psi(\omega, \mathbf{x}) = -f(\omega, \mathbf{x}) \tag{14.26}$$

By focusing on each of the frequency components, we have removed the time dependence of the problem. We can then work on a static solution for each frequency, which we can label either by ω or k. We do this using a Green's function $G_k(\mathbf{x}, \mathbf{x}')$, which is the solution to the equation

$$(\nabla^2 + k^2) G_k(\mathbf{x}, \mathbf{x}') = -\delta(\mathbf{x} - \mathbf{x}') \tag{14.27}$$

In this equation, $\mathbf{x}$ is the observation point (and the subject of the ∇^2), while $\mathbf{x}'$ is a source point. The solution for a given frequency component is then

$$\psi(\omega, \mathbf{x}) = \frac{1}{4\pi}\int d^3x' G_k(\mathbf{x}, \mathbf{x}') f(\omega, \mathbf{x}') \tag{14.28}$$

The integral is over the source volume.

The problem is spherically symmetric, so G_k must have the form

$$G_k(\mathbf{x}, \mathbf{x}') = G_k(|\mathbf{x} - \mathbf{x}'|) = G(r) \tag{14.29}$$

where $\mathbf{r} = \mathbf{x} - \mathbf{x}'$ and $r = |\mathbf{r}|$.

When we use the spherical polar form of ∇^2, we get the differential equation

$$\frac{1}{r}\frac{d^2}{dr^2}(rG_k) + k^2 G_k = -\delta(\mathbf{r}) \tag{14.30}$$

or, away from the point source ($r = 0$),

$$\frac{d^2}{dr^2}(rG_k) + k^2(rG_k) = 0 \tag{14.31}$$

so the solutions are

$$rG_k(r) = Ae^{ikr} + Be^{-ikr} \tag{14.32}$$

The first term is a diverging spherical wave, while the second term is one which is converging on the source. To fix the overall normalization of the solution, we note that as $r \to 0$, we know from electrostatics that

$$\lim_{kr\to 0} G_k(r) = \frac{1}{r} \tag{14.33}$$

so we must have $A + B = 1$. We designate the two solutions for a given frequency component

$$G_k^{(\pm)}(r) = \frac{e^{\pm ikr}}{r} \tag{14.34}$$

Returning to the time-domain problem, we look for a (new but related) Green's function $G(t, \mathbf{x}; t', \mathbf{x}')$ for a point in spacetime. This function is a solution to the equation

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) G(t, \mathbf{x}; t', \mathbf{x}') = -\delta(\mathbf{x} - \mathbf{x}')\delta(t - t') \tag{14.35}$$

The Fourier transform of the left side is

$$(\nabla^2 + k^2) G_k(\mathbf{x}; t', \mathbf{x}') \tag{14.36}$$

while for the right side it is

$$-\int_{-\infty}^{\infty} \delta(\mathbf{x} - \mathbf{x}')\delta(t - t')e^{i\omega t} dt = -\delta(\mathbf{x} - \mathbf{x}')e^{i\omega t'} \tag{14.37}$$

So the equation in the frequency domain is

$$(\nabla^2 + k^2) G_k(\mathbf{x}; t', \mathbf{x}') = -\delta(\mathbf{x} - \mathbf{x}')e^{i\omega t'} \tag{14.38}$$

Again, t has been integrated out, and the only dependence on t' on the left side is in G_k. Therefore we can equate the t' dependence on both sides, and write G_k in terms of the static G_k:

$$G_k(\mathbf{x}; t', \mathbf{x}') = G_k(\mathbf{x}, \mathbf{x}')e^{i\omega t'} \tag{14.39}$$

Then we go back to the time domain:

$$G^{(\pm)}(t,\mathbf{x};t',\mathbf{x}') = \frac{1}{2\pi}\int_{-\infty}^{\infty} G_k^{(\pm)}(\mathbf{x};t',\mathbf{x}')e^{-i\omega t}d\omega \tag{14.40}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} G_k^{(\pm)}(\mathbf{x},\mathbf{x}')e^{i\omega t'}e^{-i\omega t}d\omega \tag{14.41}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{\pm ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|}e^{-i\omega(t-t')}d\omega \tag{14.42}$$

We can simplify this with the substitutions

$$r \equiv |\mathbf{x}-\mathbf{x}'| \tag{14.43}$$

$$\tau \equiv t-t' \tag{14.44}$$

which results in

$$G^{(\pm)}(\tau,r) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{\pm ikr}}{r}e^{-i\omega\tau}d\omega \tag{14.45}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{r}e^{\pm\frac{i\omega r}{c}-i\omega\tau}d\omega \tag{14.46}$$

$$= \frac{1}{2\pi r}\int_{-\infty}^{\infty} e^{-i\omega(\tau\mp\frac{r}{c})}d\omega \tag{14.47}$$

$$= \frac{1}{r}\delta(\tau\mp\frac{r}{c}) \tag{14.48}$$

In full form,

$$G^{(\pm)}(t,\mathbf{x};t',\mathbf{x}') = \frac{1}{|\mathbf{x}-\mathbf{x}'|}\delta\left(t'-[t\mp\frac{|\mathbf{x}-\mathbf{x}'|}{c}]\right) \tag{14.49}$$

$G^{(+)}$ is called the "retarded Green function", giving behavior after the source event. The "advanced Green function" $G^{(-)}$ describes behavior before the source event, but we'll restrict ourselves to the more common circumstance that there is no incoming wave, and the source is active in a finite interval around $t'=0$. Then we have

$$\psi(t,\mathbf{x}) = \frac{1}{4\pi}\int d^3x'dt'G^{(+)}(t,\mathbf{x};t',\mathbf{x}')f(t',\mathbf{x}') \tag{14.50}$$

$$= \frac{1}{4\pi}\int d^3x'\frac{[f(t',\mathbf{x}')]_{\text{ret}}}{|\mathbf{x}-\mathbf{x}'|} \tag{14.51}$$

where $[f(t',\mathbf{x}')]_{\text{ret}}$ indicates that t' is to be evaluated at the retarded time

$$t' = t-\frac{|\mathbf{x}-\mathbf{x}'|}{c} \tag{14.52}$$

In other words, the fields ψ at some point $\mathbf{x}$ and time t are the result of past source events which can be causally connected to $(t,\mathbf{x})$. More specifically, these source events lie on the "past light cone" of $(t,\mathbf{x})$, since their effects travel at the speed of light.

We can translate the above expressions back into something closer to Steane's notation

$$\mathsf{A}(t,\mathbf{r}) = \frac{1}{4\pi\varepsilon_0 c^2}\int \frac{\mathsf{J}(t-|\mathbf{r}-\mathbf{r}_s|/c,\mathbf{r}_s)}{|\mathbf{r}-\mathbf{r}_s|}d\mathbf{r}_s \tag{14.53}$$

where the integral is over source points $\mathbf{r}_s$. A further definition is common

$$r_{\rm sf} \equiv |\mathbf{r}-\mathbf{r}_s| \tag{14.54}$$

in which case the fields are written

$$\mathsf{A}(t,\mathbf{r}) = \frac{1}{4\pi\varepsilon_0 c^2}\int \frac{\mathsf{J}(t-r_{\rm sf}/c,\mathbf{r}_s)}{r_{\rm sf}}d\mathbf{r}_s \tag{14.55}$$

14.3 Arbitrarily moving charge

We can obtain an expression for the potential of a particle undergoing arbitrary motion. The expression obtained above is a good starting point: it is a relationship between the 4-vector A and another 4-vector J; this implies that the differential element $d\mathbf{r}_s/r_{\rm sf}$ is invariant, to preserve the transformation properties. (Formally, this follows from the Quotient Rule.)

For a point charge, J should incorporate the charge's 4-velocity U. The denominator also suggests a vector

$$\mathsf{R}_{\rm sf} = \mathsf{R} - \mathsf{R}_s \tag{14.56}$$

which is the difference between the source 4-position R_s and the field observation 4-position R. Since the source 4-position must be on the past light cone of the observation point, $\mathsf{R}_{\rm sf}$ must be a null vector with

$$0 = (\mathbf{r}-\mathbf{r}_s)^2 - c^2(t-t_s)^2 \tag{14.57}$$

$$t = t_s + \frac{|\mathbf{r}-\mathbf{r}_s|}{c} \tag{14.58}$$

Let's look at some limiting cases. If the charge isn't moving, we know the potential is

$$\phi(t,\mathbf{r}) = \frac{q}{4\pi\varepsilon_0[x^2+y^2+z^2]^{1/2}} \tag{14.59}$$

In the case of uniform motion with velocity $\mathbf{v} = v\hat{\mathbf{x}}$, we have

$$\phi = \gamma(\phi' + \beta A'_x) = \frac{\gamma q}{4\pi\varepsilon_0[\gamma^2(x-vt)^2+y^2+z^2]^{1/2}} \tag{14.60}$$

$$A_x = \gamma(A'_x + \beta\phi') = \beta\phi \tag{14.61}$$

$$A_y = A_z = 0 \tag{14.62}$$

Again, we see what looks like a 4-velocity in the numerator. The denominator is also modified by a γ, which suggests more than an $\mathsf{R}\cdot\mathsf{R}$. So instead we propose

$$\mathsf{A} = \frac{q}{4\pi\varepsilon_0}\frac{\mathsf{U}/c}{(-\mathsf{R}_{\rm sf}\cdot\mathsf{U})} \tag{14.63}$$

and check that it matches the limiting cases. (In fact, you can also get this by careful consideration of how one integrates the delta function describing the source's position on the past light cone.)

In the stationary case,

$$\begin{aligned} \mathsf{U} &= (c, 0) && (14.64)\\ \mathsf{R}_s &= (ct_s, 0, 0, 0) && (14.65)\\ \mathsf{R} &= (ct, x, y, z) && (14.66)\\ \mathsf{R}_{\rm sf} &= (c(t-t_s), x, y, z) && (14.67)\\ \mathsf{R}_{\rm sf}\cdot\mathsf{U} &= -c^2(t-t_s) = -c|\mathbf{r}-\mathbf{r}_s| && (14.68) \end{aligned}$$

which results in

$$\phi(t,\mathbf{r}) = \frac{qc}{4\pi\varepsilon_0}\frac{c/c}{(-\mathsf{R}_{\rm sf}\cdot\mathsf{U})} = \frac{q}{4\pi\varepsilon_0}\frac{1}{|\mathbf{r}-\mathbf{r}_s|} \qquad (14.69)$$

the same as before.

In the uniform motion case,

$$\begin{aligned} \mathsf{U} &= \gamma(c, \mathbf{v}) && (14.70)\\ \mathsf{R}_s &= (ct_s, \mathbf{r}_s) = (ct_s, vt_s, 0, 0) && (14.71)\\ \mathsf{R} &= (ct, \mathbf{r}) = (ct, x, y, z) && (14.72)\\ \mathsf{R}_{\rm sf} &= (r_{\rm sf}, \mathbf{r}_{\rm sf}) = (c(t-t_s), x-vt_s, y, z) && (14.73)\\ & && (14.74) \end{aligned}$$

This takes a little more care: we want to express the result in terms of t, without reference to the time t_s when the radiation was emitted. We can write

$$\begin{aligned} \mathsf{R}_{\rm sf}\cdot\mathsf{U} &= \gamma\mathbf{r}_{\rm sf}\cdot\mathbf{v} - \gamma r_{\rm sf}c && (14.75)\\ &= \gamma(\mathbf{r}_{\rm sf}\cdot\mathbf{v} - r_{\rm sf}c) && (14.76) \end{aligned}$$

A naive calculation is tedious, but we can simplify things by introducing another position and its displacement from the observation point

$$\begin{aligned} \mathbf{r}_p &= \mathbf{r}_s + \mathbf{v}(t-t_s) && (14.77)\\ \mathbf{r}_{\rm pf} &= \mathbf{r} - \mathbf{r}_p && (14.78)\\ &= \mathbf{r} - \mathbf{r}_s - \mathbf{v}(t-t_s) && (14.79)\\ &= \mathbf{r}_{\rm sf} - \mathbf{v}r_{\rm sf}/c && (14.80) \end{aligned}$$

If we draw a triangle illustrating the last relation between $\mathbf{r}_{\rm pf}$, $\mathbf{r}_{\rm sf}$, and $\mathbf{v}(t-t_s)$, we see that

$$\begin{aligned} (y^2+z^2)^{1/2} &= r_{\rm sf}\sin\alpha = r_{\rm pf}\sin\theta && (14.81)\\ (x-vt) &= r_{\rm sf}\cos\alpha - v(t-t_s) = r_{\rm pf}\cos\theta && (14.82) \end{aligned}$$

where α is the angle between $\mathbf{v}$ and $\mathbf{r}_s$, and θ is that between $\mathbf{v}$ and $\mathbf{r}_p$. Then the dot product is

$$\begin{aligned}
(\mathbf{r}_{\rm sf}\cdot\mathbf{v})^2 &= r_{\rm sf}^2 v^2\cos^2\alpha && (14.83)\\
&= r_{\rm sf}^2 v^2(1-\sin^2\alpha) && (14.84)\\
&= r_{\rm sf}^2 v^2 - r_{\rm pf}^2 v^2\sin^2\theta && (14.85)\\
-\frac{r_{\rm pf}^2 v^2}{c^2}\sin^2\theta &= \left(\frac{\mathbf{r}_{\rm sf}\cdot\mathbf{v}}{c}\right)^2 - \frac{r_{\rm sf}^2 v^2}{c^2} && (14.86)
\end{aligned}$$

We then square the expression for $\mathbf{r}_{\rm pf}$

$$\begin{aligned}
\mathbf{r}_{\rm pf}\cdot\mathbf{r}_{\rm pf} &= \mathbf{r}_{\rm sf}\cdot\mathbf{r}_{\rm sf} + \mathbf{v}\cdot\mathbf{v}\frac{r_{\rm sf}^2}{c^2} - 2\frac{r_{\rm sf}}{c}\mathbf{r}_{\rm sf}\cdot\mathbf{v} && (14.87)\\
r_{\rm pf}^2 &= r_{\rm sf}^2 + \frac{r_{\rm sf}^2 v^2}{c^2} - 2\frac{r_{\rm sf}}{c}\mathbf{r}_{\rm sf}\cdot\mathbf{v} && (14.88)\\
& && (14.89)
\end{aligned}$$

and add this to the previous expression

$$\begin{aligned}
r_{\rm pf}^2 - \frac{r_{\rm pf}^2 v^2}{c^2}\sin^2\theta &= \left(\frac{\mathbf{r}_{\rm sf}\cdot\mathbf{v}}{c}\right)^2 + r_{\rm sf}^2 + -2\frac{r_{\rm sf}}{c}\mathbf{r}_{\rm sf}\cdot\mathbf{v} && (14.90)\\
r_{\rm pf}^2(1-\beta^2\sin^2\theta) &= \left(r_{\rm sf} - \frac{\mathbf{r}_{\rm sf}\cdot\mathbf{v}}{c}\right)^2 && (14.91)
\end{aligned}$$

We can then write

$$\begin{aligned}
\mathsf{R}_{\rm sf}\cdot\mathsf{U} &= \gamma(\mathbf{r}_{\rm sf}\cdot\mathbf{v} - r_{\rm sf}c) && (14.92)\\
&= -\gamma c r_{\rm pf}(1-\beta^2\sin^2\theta)^{1/2} && (14.93)\\
&= -\gamma c r_{\rm pf}(\cos^2\theta + \sin^2\theta - \beta^2\sin^2\theta)^{1/2} && (14.94)\\
&= -\gamma c r_{\rm pf}(\cos^2\theta + (1-\beta^2)\sin^2\theta)^{1/2} && (14.95)\\
&= -c r_{\rm pf}(\gamma^2\cos^2\theta + \sin^2\theta)^{1/2} && (14.96)\\
&= -c(\gamma^2(x-vt)^2 + y^2 + z^2)^{1/2} && (14.97)\\
& && (14.98)
\end{aligned}$$

Plugging into the expression for A

$$\begin{aligned}
\mathsf{A} &= \frac{q}{4\pi\varepsilon_0}\frac{\mathsf{U}/c}{(-\mathsf{R}_{\rm sf}\cdot\mathsf{U})} && (14.99)\\
&= \frac{q}{4\pi\varepsilon_0 c}\frac{\gamma(1,\beta,0,0)}{(\gamma^2(x-vt)^2 + y^2 + z^2)^{1/2}} && (14.100)
\end{aligned}$$

The expression for A is the Liénard-Weichert potential for an arbitrarily moving point charge.

Chapter 15

Accelerated charge

15.1 Slowly oscillating dipole

We now look at the radiation of a charge moving at a speed much less than c. This actually accounts for most of the light we see. Also note that in atomic physics, the selection rules you study are dipole rules; higher-order terms are highly suppressed—enough that they've come to be called "forbidden".

Consider a dipole of two charges, $-q$ and q, with $-q$ at the origin and q placed at $\mathbf{x}_q(t)$ along the $\hat{\mathbf{z}}$ axis. $\mathbf{x}_0$ is the maximum displacement of q away from $-q$. The velocity is $\mathbf{v} = \dot{\mathbf{x}}_q$.

Start from the Liénard-Weichert potential:

$$\mathsf{A} = \frac{q}{4\pi\varepsilon_0} \left.\frac{\mathsf{U}/c}{(-\mathsf{R}\cdot\mathsf{U})}\right|_{\text{ret}} \tag{15.1}$$

which is to be evaluated at the retarded time, so

$$\mathsf{U} = \gamma(c, \mathbf{v}) \tag{15.2}$$

$$\mathsf{R} = (c(t - t_s), \mathbf{r}_{\text{sf}} = \mathbf{x} - \mathbf{x}_q) \tag{15.3}$$

The R vector is a null vector, *i.e.*, $\mathsf{R}\cdot\mathsf{R} = 0$. The vector potential is then (still exact)

$$\mathsf{A} = \frac{q}{4\pi\varepsilon_0} \frac{(c, \mathbf{v})}{c(r_{\text{sf}}c - \mathbf{r}_{\text{sf}}\cdot\mathbf{v})} \tag{15.4}$$

We now make the approximation $v \ll c$, which implies

$$\frac{d}{\lambda} \approx \frac{v\Delta t}{c\Delta t} = \frac{v}{c} \ll 1 \tag{15.5}$$

where the numerator is the characteristic length of the dipole over one oscillation, and λ the wavelength (characteristic length of light over the time of one oscillation). Let

$$\mathbf{x}_q(t_s) = \mathbf{x}_0 \sin\omega t_s = \mathbf{x}_0 \sin\omega(t - r_{\text{sf}}/c) = \mathbf{x}_0 \sin(\omega t - k r_{\text{sf}}) \tag{15.6}$$

$$\mathbf{r}_{\text{sf}} = \mathbf{x} - \mathbf{x}_q(t - r_{\text{sf}}/c) \tag{15.7}$$

$$\mathbf{v} = \dot{\mathbf{x}}_q = \mathbf{x}_0\omega\cos(\omega t - k r_{\text{sf}}) \tag{15.8}$$

Note that $\mathbf{x}$ is the location where the field is being observed, while $\mathbf{x}_q(t - r_{\text{sf}}/c)$ is the position of the source when the field was emitted. We also assume that we're observing from a long distance away (far field approximation), *i.e.*, $r \gg x_q$, so $r \approx r_{\text{sf}}$ and $\mathbf{r}_{\text{sf}} \cdot \mathbf{v} \ll r_{\text{sf}}c$. This leaves us with

$$\begin{aligned} \mathsf{A} &\approx \frac{q}{4\pi\varepsilon_0 c^2}\frac{(c, \mathbf{x}_0\omega\cos(\omega t - kr_{\text{sf}}))}{r_{\text{sf}}} && (15.9)\\ &\approx \frac{q}{4\pi\varepsilon_0 c^2}\frac{(c, \mathbf{x}_0\omega\cos(\omega t - kr))}{r} && (15.10)\end{aligned}$$

We can then evaluate the magnetic field, keeping in mind that

$$\nabla \wedge f\hat{\mathbf{z}} = \left(\frac{\partial f}{\partial y}, -\frac{\partial f}{\partial x}, 0\right) \tag{15.11}$$

so that

$$\begin{aligned} \mathbf{B} &= \nabla \wedge \mathbf{A} && (15.12)\\ &= \frac{q}{4\pi\varepsilon_0 c^2}\nabla \wedge \left(\frac{\mathbf{x}_0\omega\cos(\omega t - kr)}{r}\right) && (15.13)\\ B_x &= \frac{q}{4\pi\varepsilon_0 c^2}\frac{\partial}{\partial y}\left(\frac{\omega\cos(\omega t - kr)}{r}\right) && (15.14)\\ &= \frac{q\omega}{4\pi\varepsilon_0 c^2}\left(-\frac{y\cos(\omega t - kr)}{r^3} - \frac{\sin(\omega t - kr)}{r}(-k)\frac{y}{r}\right) && (15.15)\\ &= \frac{1}{4\pi\varepsilon_0 c^2}\left(-\frac{y}{r^3}\dot{d}_{\text{ret}} - \frac{y}{cr^2}\ddot{d}_{\text{ret}}\right) && (15.16)\end{aligned}$$

where we've made the following source-related definitions:

$$\begin{aligned} \mathbf{d} &= q\mathbf{x}_0\sin\omega t_s && (15.17)\\ \dot{\mathbf{d}}_{\text{ret}} &= q\omega\mathbf{x}_0\cos(\omega t - kr) && (15.18)\\ \ddot{\mathbf{d}}_{\text{ret}} &= -q\omega^2\mathbf{x}_0\sin(\omega t - kr) && (15.19)\\ \dot{\mathbf{d}}_{\text{ret}} \wedge \mathbf{r} &= \dot{d}_{\text{ret}}(\hat{\mathbf{z}} \wedge \mathbf{r}) = (-y, x, 0) && (15.20)\end{aligned}$$

Then by replacing $-y \to x$,

$$B_y = \frac{1}{4\pi\varepsilon_0 c^2}\left(\frac{x}{r^3}\dot{d}_{\text{ret}} + \frac{x}{cr^2}\ddot{d}_{\text{ret}}\right) \tag{15.21}$$

and combining

$$\mathbf{B} = \frac{1}{4\pi\varepsilon_0 c^2}\left(\frac{1}{r^3}\dot{\mathbf{d}}_{\text{ret}} \wedge \mathbf{r} + \frac{1}{cr^2}\ddot{\mathbf{d}}_{\text{ret}} \wedge \mathbf{r}\right) \tag{15.22}$$

The first term is the retarded Biot-Savart law.

The second term can be seen to fall off as r^{-1}, and therefore dominates in the far field region. We take θ to be the angle between $\hat{\mathbf{z}}$ and $\mathbf{r}$, so in the far field

$$\mathbf{B} = \frac{1}{4\pi\varepsilon_0 c^3}q\omega^2\sin(\omega t - kr)\frac{\sin\theta}{r}|\mathbf{x}_0|\hat{\phi} \tag{15.23}$$

To calculate the electric field of the radiation part, we could go back to the potentials, but one needs higher-order corrections to get it right. Instead, we will assume that the radiation part behaves as we saw earlier: **E** will be perpendicular to both **B** and the direction of propagation (in this case **r**), and the magnitude will be Bc.

$$\mathbf{E} = c\frac{\mathbf{B} \wedge \mathbf{r}}{r} \tag{15.24}$$

$$= \frac{(\ddot{\mathbf{d}}_{\text{ret}} \wedge \mathbf{r}) \wedge \mathbf{r}}{4\pi\varepsilon_0 c^2 r^3} \tag{15.25}$$

It is possible to derive the general, exact solution for fields of an accelerated point charge from a similar procedure: start with the Liénard-Weichert potential and compute the fields, though in this case it's easier to work from $F^{\mu\nu}$. One needs to keep in mind that the derivatives in $F^{\mu\nu}$ refer to the observation point rather than at the source. The charge's motion is then evaluated in the source rest frame (*i.e.*, the charge might be moving within a source device). More details can be found in the next section. For the purpose of what we're doing, we'll just look at the solutions:

$$\mathbf{E} = \frac{q}{4\pi\varepsilon_0\kappa^2}\left(\frac{\mathbf{n} - \mathbf{v}/c}{\gamma^2 r^2} + \frac{\mathbf{n} \wedge [(\mathbf{n} - \mathbf{v}/c) \wedge \mathbf{a}]}{c^2 r}\right) \tag{15.26}$$

$$\mathbf{B} = \mathbf{n} \wedge \mathbf{E}/c \tag{15.27}$$

$$\mathbf{n} = \mathbf{r}/r \tag{15.28}$$

$$\kappa = 1 - \frac{v_r}{c} = 1 - \frac{\mathbf{n} \cdot \mathbf{v}}{c} \tag{15.29}$$

We see again the division into two terms: a first term which falls off as r^{-2}, and a far field term which falls off as r^{-1}.

The first term describes a "non-radiative" or "bound" field which remains connected with the source. If the source stops accelerating, the field settles to a static solution (though not zero).

The second term describes the "radiative" field. As we will see, because it falls off only as r^{-1}, its energy content does not decrease at large r, unlike the non-radiative term. This "unbound" radiation only occurs when there is an accelerating charge.

It is worth noting that it is the sum of the two fields which is a solution to the Maxwell equations; either one by itself is not, so the two are not completely independent.

15.2 *Field of an accelerated charge (details)

This section fills out some details from Steane's Appendix D.

Start from the Liénard-Weichert potential:

$$\mathsf{A} = \frac{q}{4\pi\varepsilon_0}\left.\frac{\mathsf{U}/c}{(-\mathsf{R} \cdot \mathsf{U})}\right|_{\text{ret}} \tag{15.30}$$

where U is the 4-velocity of the source event, and

$$\mathsf{R} = \mathsf{R}_f - \mathsf{R}_s \tag{15.31}$$

is the 4-vector difference of the field observation to the source event. Note that $\mathsf{R} \cdot \mathsf{R} = 0$, since it is reflects the propagation of the field from the source event to the observation event. (In radiation problems, remember there are two times involved: the time at the source where/when the radiation is emitted, and the time where/when it's observed.)

Define

$$k \equiv -\frac{q}{4\pi\varepsilon_0 c} \tag{15.32}$$

$$s \equiv \mathsf{R} \cdot \mathsf{U} \tag{15.33}$$

so we can write

$$\mathsf{A} = k\mathsf{U}/s \tag{15.34}$$

Since we're going to try to find the fields themselves,

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{15.35}$$

let's try to evaluate $\partial_\mu A^\nu$ first. Note that we've taken the index on the differential operator down so we're only dealing with contravariant coordinates (the one's we're most used to). Also note that ∂_μ differentiates with respect to R_f, the field observation point, since $F^{\mu\nu}$ comes from the variation in A^μ where it's observed.

$$\partial_\mu A^\nu = k\frac{s\partial_\mu U^\nu - U^\nu \partial_\mu s}{s^2} \tag{15.36}$$

Define the acceleration of the source event in terms of the source's proper time,

$$a^\nu = \frac{dU^\nu}{d\tau} \tag{15.37}$$

which allows us to evaluate the derivative of the source's 4-velocity:

$$\partial_\mu U^\nu = \frac{dU^\nu}{d\tau}\frac{\partial\tau}{\partial x^\mu} = a^\nu \partial_\mu \tau \tag{15.38}$$

More generally,

$$\partial_\mu = (\partial_\mu \tau)\frac{d}{d\tau} \tag{15.39}$$

We now observe that since it is always true that $\mathsf{R} \cdot \mathsf{R} = 0$, then R is always orthogonal to its gradient. Therefore

$$\begin{aligned}
0 &= R_\nu \partial_\mu R^\nu && (15.40)\\
&= R_\nu \partial_\mu (R_f^\nu - R_s^\nu) && (15.41)\\
&= R_\nu (\delta_\mu^\nu - (\partial_\mu \tau)\frac{dR_s^\nu}{d\tau}) && (15.42)\\
&= R_\nu (\delta_\mu^\nu - U^\nu \partial_\mu \tau) && (15.43)\\
&= R_\mu - R_\nu U^\nu (\partial_\mu \tau) && (15.44)\\
\partial_\mu \tau &= R_\mu / s && (15.45)\\
\partial_\mu U^\nu = a^\nu \partial_\mu \tau &= a^\nu R_\mu / s && (15.46)\\
s\partial_\mu U^\nu &= a^\nu R_\mu && (15.47)
\end{aligned}$$

For the second term, we need

$$\begin{aligned}
\partial_\mu s &= \partial_\mu(R_\nu U^\nu) && (15.48)\\
&= (\partial_\mu R^\nu)U_\nu + R_\nu(\partial_\mu U^\nu) && (15.49)\\
&= (\delta^\nu_\mu - U^\nu\partial_\mu\tau)U_\nu + R_\nu(a^\nu\partial_\mu\tau) && (15.50)\\
&= U_\mu - U^\nu U_\nu\frac{R_\mu}{s} + R_\nu a^\nu\frac{R_\mu}{s} && (15.51)\\
&= U_\mu + \frac{R_\mu}{s}(c^2 + \mathsf{R}\cdot\mathsf{a}) && (15.52)\\
U^\nu\partial_\mu s &= U^\nu U_\mu + \frac{R_\mu}{s}(c^2 + \mathsf{R}\cdot\mathsf{a})U^\nu && (15.53)
\end{aligned}$$

Combining,

$$\begin{aligned}
\partial_\mu A^\nu &= \frac{k}{s^2}(a^\nu R_\mu - U^\nu U_\mu - \frac{R_\mu}{s}(c^2 + \mathsf{R}\cdot\mathsf{a})U^\nu) && (15.54)\\
&= -\frac{k}{s^2}U^\nu U_\mu - \frac{kc^2R_\mu}{s^3}(-\frac{s}{c^2}a^\nu + U^\nu + \frac{1}{c^2}(\mathsf{R}\cdot\mathsf{a})U^\nu) && (15.55)\\
&= -\frac{k}{s^2}U^\nu U_\mu - \frac{kc^2}{s^3}R_\mu\tilde{U}^\nu && (15.56)\\
\partial^\mu A^\nu &= -\frac{k}{s^2}U^\nu U^\mu - \frac{kc^2}{s^3}R^\mu\tilde{U}^\nu && (15.57)
\end{aligned}$$

where

$$\tilde{U}^\nu \equiv U^\nu - \frac{1}{c^2}(sa^\nu - U^\nu(\mathsf{R}\cdot\mathsf{a})) \qquad (15.58)$$

Combining into $F^{\mu\nu}$, the first term drops out, so

$$\begin{aligned}
F^{\mu\nu} &= \partial^\mu A^\nu - \partial^\nu A^\mu && (15.59)\\
&= -\frac{kc^2}{s^3}(R^\mu\tilde{U}^\nu - R^\nu\tilde{U}^\mu) && (15.60)\\
&= \frac{qc}{4\pi\varepsilon_0}\frac{R^\mu\tilde{U}^\nu - R^\nu\tilde{U}^\mu}{(\mathsf{R}\cdot\mathsf{U})^3} && (15.61)
\end{aligned}$$

To evaluate the field itself, we choose a frame in which the source moves (*i.e.*, a frame of a device in which the charges move):

$$\begin{aligned}
\mathsf{R} &= (r, \mathbf{r}) && (15.62)\\
\mathsf{U} &= \gamma(c, \mathbf{u}) && (15.63)\\
s &= g_{\mu\nu}R^\mu U^\nu = \gamma(-cr + \mathbf{r}\cdot\mathbf{u}) = -c\gamma(r - \frac{\mathbf{r}\cdot\mathbf{u}}{c}) && (15.64)
\end{aligned}$$

(Remember that we're using the metric $-+++$)

We will also need the acceleration. Because we're evaluating in the device's frame, the proper time is the same as the time at the source.

$$a^\mu = \frac{dU^\mu}{d\tau} = \left(\frac{d\gamma}{d\tau}c, \frac{d\gamma}{d\tau}\mathbf{u} + \gamma\frac{d\mathbf{u}}{d\tau}\right) = \gamma(\dot{\gamma}c, \dot{\gamma}\mathbf{u} + \gamma\mathbf{a}) \qquad (15.65)$$

The expression for the electric field is

$$E^i = cF^{0i} = \frac{-kc^3}{s^3}(R^0\tilde{U}^i - R^i\tilde{U}^0) \tag{15.66}$$

So we evaluate the modified velocities:

$$\begin{aligned}
\mathsf{R}\cdot\mathsf{a} &= g_{\mu\nu}R^\mu a^\nu = \gamma(-\dot\gamma rc + \dot\gamma\mathbf{u}\cdot\mathbf{r} + \gamma\mathbf{a}\cdot\mathbf{r}) && (15.67)\\
\tilde{U}^0 &= U^0 - \frac{1}{c^2}(sa^0 - U^0(\mathsf{R}\cdot\mathsf{a})) && (15.68)\\
&= \gamma c - \frac{1}{c^2}(\gamma(\mathbf{r}\cdot\mathbf{u})\gamma\dot\gamma c - \gamma cr\gamma\dot\gamma c - \gamma^2 c(\gamma\mathbf{a}\cdot\mathbf{r} + \dot\gamma\mathbf{u}\cdot\mathbf{r} - \dot\gamma rc)) && (15.69)\\
&= \gamma c + \gamma^3\frac{\mathbf{a}\cdot\mathbf{r}}{c} && (15.70)\\
\tilde{U}^i &= U^i - \frac{1}{c^2}(sa^i - U^i(\mathsf{R}\cdot\mathsf{a})) && (15.71)\\
&= \gamma\mathbf{u} - \frac{1}{c^2}(\gamma(\mathbf{u}\cdot\mathbf{u} - cr)(\gamma\dot\gamma\mathbf{u} - \gamma^2\mathbf{a}) - \gamma\mathbf{u}(\gamma^2\mathbf{a}\cdot\mathbf{r} + \gamma\dot\gamma\mathbf{u}\cdot\mathbf{r} - \gamma\dot\gamma rc)) && (15.72)\\
&= \gamma\mathbf{u} - \frac{1}{c^2}(\gamma\mathbf{u}(\gamma\dot\gamma\mathbf{r}\cdot\mathbf{u} - \gamma\dot\gamma cr - \gamma^2\mathbf{a}\cdot\mathbf{r} - \gamma\dot\gamma\mathbf{u}\cdot\mathbf{r} + \gamma\dot\gamma rc) + \gamma^3\mathbf{a}(\mathbf{r}\cdot\mathbf{u} - cr)) && (15.73)\\
&= \gamma\mathbf{u} + \frac{\gamma^3}{c^2}(\mathbf{a}\cdot\mathbf{r})\mathbf{u} - \frac{\gamma^3}{c^2}(\mathbf{r}\cdot\mathbf{u} - cr)\mathbf{a} && (15.74)\\
&= \gamma\mathbf{u} + \frac{\gamma^3}{c^2}(\mathbf{a}\cdot\mathbf{r})\mathbf{u} + \frac{\gamma^2}{c^2}s\mathbf{a} && (15.75)
\end{aligned}$$

Then after a fair amount of tedious algebra (as if this wasn't already), and taking advantage of the vector algebra rule

$$\mathbf{a}\wedge(\mathbf{b}\wedge\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c} \tag{15.76}$$

we get the formula

$$\mathbf{E} = \frac{q}{4\pi\varepsilon_0(r - \mathbf{r}\cdot\mathbf{u}/c)^3}\left[\frac{\mathbf{r} - \mathbf{u}r/c}{\gamma^2} - \frac{\mathbf{r}\wedge((\mathbf{r} - \mathbf{u}r/c)\wedge\mathbf{a})}{c^2}\right] \tag{15.77}$$

In summary (Steane),

$$\begin{aligned}
\mathbf{E} &= \frac{q}{4\pi\varepsilon_0\kappa^2}\left(\frac{\mathbf{n} - \mathbf{v}/c}{\gamma^2 r^2} + \frac{\mathbf{n}\wedge[(\mathbf{n} - \mathbf{v}/c)\wedge\mathbf{a}]}{c^2 r}\right) && (15.78)\\
\mathbf{B} &= \mathbf{n}\wedge\mathbf{E}/c && (15.79)\\
\mathbf{n} &= \mathbf{r}/r && (15.80)\\
\kappa &= 1 - \frac{v_r}{c} = 1 - \frac{\mathbf{n}\cdot\mathbf{v}}{c} && (15.81)
\end{aligned}$$

15.3 *Half-wave electric dipole antenna

For an antenna, we have

$$\omega d_0 = \omega qL = IL \tag{15.82}$$

where I is the ac current in a short wire of length L. Then

$$E = cB = \frac{\omega^2 d_0}{4\pi\varepsilon_0 c^2}\sin(\omega t - kr)\frac{\sin\theta}{r} \tag{15.83}$$

$$= -\frac{iI}{2\varepsilon_0 c}\frac{L}{\lambda}\frac{\sin\theta}{r}e^{i(kr-\omega t)} \tag{15.84}$$

where one takes the real part to get the field value. Note that the current goes as $\dot{d} \sim \cos\omega t$, so is 1/4 out of phase.

A half-wave dipole antenna is designed with $L = \lambda/2$, so the maximum current oscillation is at the center, and zero at the ends; one feeds in power at the center. This maximizes the power output for the antenna. The current as a function of position in the antenna is

$$I = I_0 \cos kz \tag{15.85}$$

so that, integrating,

$$\int I dz = I_0\frac{\lambda}{\pi} \tag{15.86}$$

which leads to fields (replacing IL with the integral)

$$E = cB \approx -\frac{iI_0}{2\pi\varepsilon_0 c}\frac{\sin\theta}{r}e^{i(kr-\omega t)} \tag{15.87}$$

15.3.1 *Radiated power

To calculate the power, we work from the radiation field, ignoring the near Coulomb field

$$\mathbf{E}_{rad} = \frac{q}{4\pi\varepsilon_0\kappa^3}\frac{\hat{\mathbf{n}}\wedge[(\hat{\mathbf{n}}-\mathbf{v}/c)\wedge\mathbf{a}]}{c^2 r} \tag{15.88}$$

$$\mathbf{B}_{rad} = \frac{\hat{\mathbf{n}}\wedge\mathbf{E}_{rad}}{c} \tag{15.89}$$

$$\hat{\mathbf{n}} \equiv \mathbf{r}/r \tag{15.90}$$

$$\kappa \equiv 1 - v_r/c = 1 - \hat{\mathbf{n}}\cdot\mathbf{v}/c \tag{15.91}$$

For $v/c \ll 1$,

$$\mathbf{E}_{rad} \approx \frac{q}{4\pi\varepsilon_0 c^2}\left(\frac{\hat{\mathbf{n}}\wedge(\hat{\mathbf{n}}\wedge\mathbf{a})}{r}\right) \tag{15.92}$$

We also ignore diffraction effects, and the energy required to power the antenna itself.

The Poynting vector is

$$\mathbf{N} \equiv \varepsilon_0 c^2 \mathbf{E}\wedge\mathbf{B} \tag{15.93}$$

$$= \varepsilon_0 c\mathbf{E}_{rad}\wedge(\hat{\mathbf{n}}\wedge\mathbf{E}_{rad}) \tag{15.94}$$

$$= \varepsilon_0 c[(\mathbf{E}_{rad}\cdot\mathbf{E}_{rad})\hat{\mathbf{n}} - (\mathbf{E}_{rad}\cdot\hat{\mathbf{n}})\mathbf{E}_{rad}] \tag{15.95}$$

We see that the second term is zero because we can define $\mathbf{b} = \hat{\mathbf{n}} \wedge \mathbf{a}$ and

$$\begin{aligned} \mathbf{E}_{rad} \cdot \hat{\mathbf{n}} &\propto (\hat{\mathbf{n}} \wedge \mathbf{b}) \cdot \hat{\mathbf{n}} && (15.96)\\ &= (\hat{\mathbf{n}} \wedge \hat{\mathbf{n}}) \cdot \mathbf{b} && (15.97)\\ &= 0 && (15.98) \end{aligned}$$

Therefore

$$\mathbf{N} = \varepsilon_0 c E_{rad}^2 \hat{\mathbf{n}} \tag{15.99}$$

An easy way to evaluate E_{rad}^2 is

$$\begin{aligned} E_{rad}^2 &\propto (\hat{\mathbf{n}} \wedge \mathbf{b}) \cdot (\hat{\mathbf{n}} \wedge \mathbf{b}) && (15.100)\\ &= \hat{\mathbf{n}} \cdot [\mathbf{b} \wedge (\hat{\mathbf{n}} \wedge \mathbf{b})] && (15.101)\\ &= \hat{\mathbf{n}} \cdot [b^2 \hat{\mathbf{n}} - (\mathbf{b} \cdot \hat{\mathbf{n}})\mathbf{b}] && (15.102)\\ &= b^2 - (\hat{\mathbf{n}} \cdot \mathbf{b})^2 && (15.103) \end{aligned}$$

We know that $\hat{\mathbf{n}} \cdot \mathbf{b} = \hat{\mathbf{n}} \cdot (\hat{\mathbf{n}} \wedge \mathbf{a}) = 0$ and $b^2 = a^2 \sin^2\theta$, where θ is the angle between $\mathbf{a}$ (parallel to the orientation of the dipole) and the radial direction $\hat{\mathbf{n}}$.

$$\begin{aligned} E_{rad}^2 &= \left(\frac{q}{4\pi\varepsilon_0 c^2}\right)^2 \frac{a^2 \sin^2\theta}{r^2} && (15.104)\\ \mathbf{N} &= \varepsilon_0 c \left(\frac{q}{4\pi\varepsilon_0 c^2}\right)^2 \frac{a^2 \sin^2\theta}{r^2} \hat{\mathbf{n}} && (15.105)\\ & && (15.106) \end{aligned}$$

We see that the energy flux is purely radial. The differential power is

$$\frac{dP}{d\Omega} = Nr^2 = \frac{q^2}{4\pi\varepsilon_0 c^3} \frac{1}{4\pi} a^2 \sin^2\theta \tag{15.107}$$

This is the power flux into an area $r^2 d\Omega$. The $\sin^2\theta$ distribution is characteristic of a dipole.

The total power is then

$$\begin{aligned} P_L &= \int d\Omega \frac{dP}{d\Omega} && (15.108)\\ &= \frac{q^2}{4\pi\varepsilon_0 c^3} \frac{1}{4\pi} a^2 2\pi \int_{-1}^{1} (1 - \cos^2\theta) d\cos\theta && (15.109)\\ &= \frac{q^2}{4\pi\varepsilon_0 c^3} \frac{a^2}{2} (2 - \frac{2}{3}) && (15.110)\\ &= \frac{2}{3} \frac{q^2}{4\pi\varepsilon_0} \frac{a^2}{c^3} && (15.111) \end{aligned}$$

which as we see does not diminish with r. This is the Larmor formula for a non-relativistic accelerating charge.

For the relativistic case, we note that P must be an invariant, and we're looking for a formula which reduces to P_L when $\beta \to 0$. Previous expressions for $\mathbf{E}$ and $\mathbf{B}$ indicate that the only

(3-)vectors we can use are $\mathbf{v}$ and $\mathbf{a}$. We can rewrite also the non-relativistic Larmor formula in a suggestive form

$$P_L = \frac{2}{3}\frac{q^2}{4\pi\varepsilon_0 c^3}|\dot{\mathbf{v}}|^2 = \frac{2}{3}\frac{q^2}{4\pi\varepsilon_0 m^2 c^3}\left(\frac{d\mathbf{p}}{dt}\cdot\frac{d\mathbf{p}}{dt}\right) \tag{15.112}$$

in terms of non-relativistic momenta. We infer a covariant form

$$P = \frac{2}{3}\frac{q^2}{4\pi\varepsilon_0 m^2 c^3}\frac{dp_\mu}{d\tau}\frac{dp^\mu}{d\tau} \tag{15.113}$$

where

$$d\tau = \frac{dt}{\gamma} \tag{15.114}$$

The 4-acceleration can be worked out from

$$\begin{aligned}
\mathsf{U} &= (\gamma c, \gamma\mathbf{v}) && (15.115)\\
\frac{d\mathsf{U}}{d\tau} &= \gamma\dot{\mathsf{U}} && (15.116)\\
&= \gamma(\dot{\gamma}c, \dot{\gamma}\mathbf{v} + \gamma\mathbf{a}) && (15.117)
\end{aligned}$$

In terms of $\mathbf{u}$ and $\mathbf{a}$,

$$\dot{\gamma} = \frac{d}{dt}\left(1 - \frac{\mathbf{v}\cdot\mathbf{v}}{c^2}\right)^{-1/2} = \gamma^3\frac{\mathbf{v}\cdot\mathbf{a}}{c^2} \tag{15.118}$$

Therefore we have

$$\begin{aligned}
\frac{d\mathsf{U}}{d\tau} &= \gamma\left(\gamma^3\frac{\mathbf{v}\cdot\mathbf{a}}{c}, \gamma^3\frac{\mathbf{v}\cdot\mathbf{a}}{c^2}\mathbf{v} + \gamma\mathbf{a}\right) && (15.119)\\
&= \gamma^2\left(\gamma^2\frac{\mathbf{v}\cdot\mathbf{a}}{c}, \gamma^2\frac{\mathbf{v}\cdot\mathbf{a}}{c^2}\mathbf{v} + \mathbf{a}\right) && (15.120)\\
\frac{d\mathsf{U}}{d\tau}\cdot\frac{d\mathsf{U}}{d\tau} &= \gamma^4\left[-\left(\frac{\mathbf{v}\cdot\mathbf{a}}{c}\right)^2\gamma^4 + \left(\frac{\mathbf{v}\cdot\mathbf{a}}{c^2}\gamma^2\mathbf{v} + \mathbf{a}\right)^2\right] && (15.121)\\
&= \gamma^4\left[-\left(\frac{\mathbf{v}\cdot\mathbf{a}}{c}\right)^2\gamma^4 + \left(\frac{\mathbf{v}\cdot\mathbf{a}}{c^2}\right)^2\gamma^4 v^2 + a^2 + 2\gamma^2\frac{(\mathbf{v}\cdot\mathbf{a})^2}{c^2}\right] && (15.122)\\
&= \gamma^4\left[-\left(\frac{\mathbf{v}\cdot\mathbf{a}}{c}\right)^2\gamma^4(-1+\beta^2) + a^2 + 2\gamma^2\frac{(\mathbf{v}\cdot\mathbf{a})^2}{c^2}\right] && (15.123)\\
&= \gamma^4\left[-\left(\frac{\mathbf{v}\cdot\mathbf{a}}{c}\right)^2\gamma^2 + a^2 + 2\gamma^2\frac{(\mathbf{v}\cdot\mathbf{a})^2}{c^2}\right] && (15.124)\\
&= \gamma^4\left[a^2 + \gamma^2\frac{(\mathbf{v}\cdot\mathbf{a})^2}{c^2}\right] && (15.125)\\
&= \gamma^4\left[a^2\gamma^2(1-\beta^2) + \gamma^2\frac{(\mathbf{v}\cdot\mathbf{a})^2}{c^2}\right] && (15.126)\\
&= \gamma^6\left[a^2 - \left(\frac{av}{c}\right)^2 + \left(\frac{av}{c}\right)^2\cos^2\alpha\right] && (15.127)\\
&= \gamma^6\left[a^2 - \left(\frac{av}{c}\right)^2\sin^2\alpha\right] && (15.128)\\
&= \gamma^6\left[a^2 - \left(\frac{(\mathbf{v}\wedge\mathbf{a})^2}{c^2}\right)\right] && (15.129)
\end{aligned}$$

Combining, we obtain the Liénard result

$$P = \frac{2}{3}\frac{q^2}{4\pi\varepsilon_0 c^3}\gamma^6\left[a^2 - \frac{(\mathbf{v}\wedge\mathbf{a})^2}{c^2}\right] \tag{15.130}$$

15.3.2 Energy loss in accelerators

Particle accelerators generally come in two varieties, linear and circular. Linear accelerators add energy to a beam only once as it passes along the machine, so important limitations come from how long they can be built, and how strong their accelerating elements can be made.

Circular accelerators have the advantage of being able to add energy to a beam many times as it recirculates in the machine. The drawback is one then has to bring the beam back (thus the circle). This means that for most of the distance around the accelerator, one redirects the beam without actually adding any energy to it.

The acceleration is

$$a = \frac{v^2}{\rho} = \frac{\beta^2 c^2}{\rho} \tag{15.131}$$

where ρ is the orbit radius. The radiated power (loss) from Larmor's formula is

$$P_L = \frac{2}{3}\frac{q^2}{4\pi\varepsilon_0 c^3}\gamma^6\frac{1}{\gamma^2}\frac{\beta^4 c^4}{\rho^2} = \frac{2}{3}\frac{q^2 c}{4\pi\varepsilon_0\rho^2}\beta^4\gamma^4 \tag{15.132}$$

The energy loss per revolution is

$$\Delta E = \frac{P}{f} = \frac{2\pi\rho}{\beta c}P = \frac{4\pi}{3}\frac{q^2}{4\pi\varepsilon_0\rho}\beta^3\gamma^4 = \frac{4\pi}{3}\frac{\alpha\hbar c}{\rho}\beta^3\left(\frac{E}{mc^2}\right)^4 \tag{15.133}$$

where $\alpha = 1/137$ is the fine structure constant, $\hbar c = 197$ MeV fm, E is the beam energy, and m is the mass of the beam particle.

In some of the first electron synchrotrons,

$$\rho \sim 1\text{ m} \tag{15.134}$$

$$E_{max} \sim 0.3\text{ GeV} \tag{15.135}$$

which ends up around 1 keV per revolution. This is less than the energy gain of a few keV per turn, but not negligible. On the other hand, for LEP this amounted to $O(0.1\%)$ energy loss per revolution and was measured in megawatts, so a lot of energy was expended just keeping the beam at the same energy. (And one more "on the other hand": this power loss, in the form of "synchrotron radiation", is actually useful in other branches of physics, chemistry, and biology because it is a source of high-intensity light—and thus, for example, specially built facilities such as the Diamond Light Source. One field's annoyance is another field's opportunity.)

For the LHC, this effect is much suppressed because protons are 2000 times more massive than electrons. The LHC is the first proton accelerator where the synchrotron radiation is even noticeable.

Chapter 16

Energy-momentum tensor

Earlier, we obtained for the energy density and flux

$$u = \frac{1}{2\mu_0}\left(B^2 + \frac{E^2}{c^2}\right) \tag{16.1}$$

$$\mathbf{N} = \frac{1}{\mu_0}\mathbf{E}\wedge\mathbf{B} \tag{16.2}$$

We also had a continuity equation

$$\frac{\partial u}{\partial t} + \nabla\cdot\mathbf{N} = -\mathbf{j}\cdot\mathbf{E} \tag{16.3}$$

One might be tempted to put u and $\mathbf{N}$ into a 4-vector N and write the equation in the form

$$\partial_\mu N^\nu = ? \tag{16.4}$$

but $\mathbf{j}\cdot\mathbf{E}$ isn't a proper 4-vector. If anything, it appears to be a time-like component of $F_{0\mu}J^\mu$. This suggests that u and $\mathbf{N}$ are really part of another object. This is the "stress-energy" tensor, which describes momentum flows within a body. It was originally used to describe mechanical stresses and how forces change directions, but it applies to any system which can be described with a field. In the case of the electromagnetic field, it is often divorced from its mechanical origin and called the "energy-momentum" tensor.

The elements of an arbitrary stress-energy tensor can be interpreted as follows

T^{00}	energy density
T^{0j}	energy flux
T^{j0}	momentum density
T^{jj}	pressure
T^{ij}, $i \neq j$	shear stress

The space elements of the tensor represent the momentum flux. There's not a guarantee that this tensor is symmetric, but it's also ambiguous up to a 4-divergence, *i.e.*, one can add a 4-divergence term which will not affect any physics.

The units of the tensor elements are of force per unit area, or pressure.

We would like to assign the energy density to T^{00}. And since the Poynting vector represents energy flow, it should find its place in T^{0i}. But to get a better picture of the tensor, let's look at some fluid examples first.

(The symmetric rank-2 tensor field is a member of the $(1,1)$ representation space. It shows up as a spin-2 graviton.)

16.1 Fluid examples

The simplest fluid we can envisage is something like a dust cloud: in its rest frame, there is no stress, and the only energy comes from the rest masses of the particles themselves. So in the cloud's rest frame, we have

$$T^{\mu\nu} = \begin{pmatrix} \rho_0 c^2 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix} \tag{16.5}$$

where ρ_0 is the mass density.

Now let's give the cloud an overall motion 4-vector U. If we choose a direction $\hat{\mathbf{x}}$, then the amount of dust crossing a plane of constant x is clearly proportional to u_x. The amount of momentum crossing with each dust particle is also proportional to u_x. Similarly, the amount of y momentum crossing with each dust particle is proportional to u_y.

This leads us to a simple form of the stress-energy tensor:

$$T^{\mu\nu} = \rho_0 U^\mu U^\nu \tag{16.6}$$

which we also might have expected from the fact that we only had ρ_0 and U out of which to fashion the tensor.

For an ideal fluid, there is no heat conduction, and no viscosity. These conditions imply that $T^{0i} = T^{i0} = 0$ and $T^{ij} = 0$ for $i \neq j$. In its rest frame, there should be an energy density $\rho_0 c^2$ and pressure p.

$$T^{\mu\nu} = \begin{pmatrix} \rho_0 c^2 & & & \\ & p & & \\ & & p & \\ & & & p \end{pmatrix} \tag{16.7}$$

If we build the tensor out of 4-velocities as before, we expect it to have the form

$$T^{\mu\nu} = (\rho_0 + \frac{p}{c^2}) U^\mu U^\nu + p g^{\mu\nu} \tag{16.8}$$

(If we use a metric with a different signature, the sign of the second term can change.)

16.2 *Energy-momentum tensor of the EM field

Returning to the electromagnetic field, how can we fit the energy density and Poynting vector into the tensor and still have $T^{\mu\nu}$ constructed out of other tensors? Let's start with T^{0i}. Since it is a cross product between $\mathbf{E}$ and $\mathbf{B}$, it suggests a contraction within $F^{\mu\nu}$

$$\frac{1}{c}(\mathbf{E} \wedge \mathbf{B})^i = g_{\mu\nu} F^{0\mu} F^{\nu i} \tag{16.9}$$

We assign this to T^{0i}. But if we extend this to the time-time component, we get

$$g_{\mu\nu} F^{0\mu} F^{\nu 0} = -E^2/c^2 \tag{16.10}$$

so we "fix this up" by adding an invariant term for diagonal terms.

$$F_{\mu\nu} F^{\mu\nu} = 2\left(B^2 - \frac{E^2}{c^2}\right) \tag{16.11}$$

In the end, we work with the following:

$$T^{\mu\nu} = \frac{1}{\mu_0}\left[-\frac{1}{4}(F_{\alpha\beta}F^{\alpha\beta})g^{\mu\nu} - F^{\mu}{}_{\gamma}F^{\gamma\nu}\right] \tag{16.12}$$

With any luck, we'll be able to look at some more methodical derivations in the last few lectures.

When we look at the components, we find

$$T^{\mu\nu} = \begin{pmatrix} \frac{1}{2\mu_0}(B^2 + \frac{E^2}{c^2}) & N_x/c & N_y/c & N_z/c \\ N_x/c & P_{11} & P_{12} & P_{13} \\ N_y/c & P_{21} & P_{22} & P_{23} \\ N_z/c & P_{31} & P_{32} & P_{33} \end{pmatrix} \tag{16.13}$$

The spatial elements

$$P_{ij} \equiv \frac{1}{\mu_0}\left[\frac{1}{2}\delta_{ij}(B^2 + \frac{E^2}{c^2}) - (B_i B_j + \frac{E_i E_j}{c^2})\right] \tag{16.14}$$

then form the momentum flux tensor.

Now the form-invariant generalization of Poynting's theorem is

$$\partial_\alpha T^{\alpha\beta} = 0 \tag{16.15}$$

It is also worth looking at the angular momentum of the electromagnetic field,

$$\mathbf{L}_{\text{field}} = \int \mathbf{x} \wedge (\mathbf{E} \wedge \mathbf{B}) d^3x \tag{16.16}$$

The generalization is a rank-3 tensor

$$M^{\alpha\beta\gamma} = T^{\alpha\beta}x^\gamma - T^{\alpha\gamma}x^\beta \tag{16.17}$$

Angular momentum conservation is

$$\begin{aligned} \partial_\alpha M^{\alpha\beta\gamma} &= \partial_\alpha T^{\alpha\beta} x^\gamma - \partial_\alpha T^{\alpha\gamma} x^\beta && (16.18)\\ &= (\partial_\alpha T^{\alpha\beta}) x^\gamma + T^{\gamma\beta} - (\partial_\alpha T^{\alpha\gamma}) x^\beta - T^{\beta\gamma} && (16.19)\\ &= T^{\gamma\beta} - T^{\beta\gamma} && (16.20)\\ &= 0 && (16.21) \end{aligned}$$

which is zero because of the symmetry of the stress-energy tensor.

16.3 *Applications with simple geometries

16.3.1 *Parallel-plate capacitor

A parallel-plate capacitor with area A oriented with its gap along the x direction has a constant electric field $E\hat{\mathbf{x}}$. The charge on each plate is

$$Q = A\epsilon_0 E \tag{16.22}$$

so the force pulling the plates together is

$$f = \frac{QE}{2} = \epsilon_0 A \frac{E^2}{2} \tag{16.23}$$

The stress tensor is

$$T^{\mu\nu} = \frac{\epsilon_0 E^2}{2} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \tag{16.24}$$

The energy density and pressure are as expected, the pressure being negative since the plates are being pulled together. There is also an outward pressure which reflects the tension among the field lines.

16.3.2 *Long straight solenoid

This case is similar to the parallel-plate capacitor, with a similar layout of field lines. The stress tensor is

$$T^{\mu\nu} = \frac{\epsilon_0 c^2 B^2}{2} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \tag{16.25}$$

16.3.3 *Plane waves

Let's examine a plane wave travelling in the x direction, polarized along the y direction:

$$\mathbf{E} = (0, E, 0)\cos(\omega t - kx) \tag{16.26}$$
$$\mathbf{B} = (0, 0, B)\cos(\omega t - kx) \tag{16.27}$$
$$B = \frac{E}{c} = \frac{kE}{\omega} \tag{16.28}$$

The energy density is

$$B^2 + \frac{E^2}{c^2} = \frac{2E^2}{c^2} \tag{16.29}$$

The Poynting vector is

$$\mathbf{N} = \frac{1}{\mu_0}\mathbf{E} \wedge \mathbf{B} = \frac{1}{\mu_0}\hat{\mathbf{x}}EB\cos^2(\omega t - kx) \tag{16.30}$$
$$= \hat{\mathbf{x}}\frac{E^2}{\mu_0 c}\cos^2(\omega t - kx) \tag{16.31}$$

The plane wave therefore carries energy in the x direction (as expected), and we should see that it exerts a pressure (momentum flow) in the x direction as well:

$$P_{11} = \frac{1}{2\mu_0}\frac{2E^2}{c^2} - \frac{1}{\mu_0}(B_x^2 + \frac{E_x^2}{c^2}) = \frac{E^2}{\mu_0 c^2} \tag{16.32}$$
$$P_{22} = \frac{1}{2\mu_0}\frac{2E^2}{c^2} - \frac{1}{\mu_0}(B_y^2 + \frac{E_y^2}{c^2}) = 0 \tag{16.33}$$
$$P_{33} = \frac{1}{2\mu_0}\frac{2E^2}{c^2} - \frac{1}{\mu_0}(B_z^2 + \frac{E_z^2}{c^2}) = 0 \tag{16.34}$$

The stress tensor is therefore

$$T^{\mu\nu} = \frac{E^2}{\mu_0 c^2}\cos^2(\omega t - kx)\begin{pmatrix} 1 & 1 & & \\ 1 & 1 & & \\ & & 0 & \\ & & & 0 \end{pmatrix} \tag{16.35}$$

A more general way to write this tensor admits any direction:

$$T^{\mu\nu} = \frac{E^2\cos^2(\mathsf{K}\cdot\mathsf{X})}{\mu_0 c^2\omega^2}\mathsf{K}^\mu\mathsf{K}^\nu \tag{16.36}$$

where

$$\mathsf{K} = (\omega/c, k, 0, 0) \tag{16.37}$$

is the 4-wavenumber vector. An interesting side-note on this way of writing the tensor is that the quotient rule, which states that an object which contracts with a tensor to produce another tensor must itself be a tensor, implies $(E/\omega)^2$ is a (rank-0) tensor, *i.e.*, a scalar.

Chapter 17

Noether's theorem

In this chapter, we'll develop the connection between a symmetry and a conservation law as shown in Noether's Theorem.

Emmy Noether actually proved two theorems in her 1918 paper. The first, which links global symmetries with conservation laws, is the most widely used and taught. As a result (and as usual for important physics), it has gone through a number of mutations and simplifications as it has been taught and re-taught. What is presented here is thus not the original form of Noether's Theorem, but a simpler "modern" form.

The second theorem has recently gotten some more attention: it shows dependencies between equations of motion imposed by "local" symmetries. See KA Barding (*Studies in History and Philosophy of Modern Physics*, 33 (2002) 3-22). It is beyond the scope of this course to discuss these local symmetries in detail, but we'll take a peek at one of the most important, that of "local gauge invariance", a principle embedded deeply into how we understand the physics of elementary particles. Which is, when you think of it, pretty remarkable for a theorem which was proven in the context of classical, non-quantum physics.

17.1 Discrete systems

(based on Bañados and Reyes, arxiv 1601.03616v2)

Noether's (First) Theorem applies to any system that can be derived from an action and possesses some continuous (non-gauge) symmetry.

There are two kinds of symmetry at play. First, there's a loose form of *invariance* in the form of the action. Second, there's an invariance in the action with respect to variations around the classical path. Noether's Theorem involves equating the two.

17.1.1 Action invariance

The action is a functional of a path, $I[q^k(t)]$, for k generalized coordinates which are a function of the variable t (which looks like time, but is in reality a dummy variable).

$$I[q^k(t)] = \int L(q^k(t), \dot{q}^k(t))dt \tag{17.1}$$

The path is now modified by a (small) function $f^k(t)$. The change in the action is

$$\delta I[q^k(t), f^k(t)] = I[q^k(t) + f^k(t)] - I[q^k(t)] \tag{17.2}$$

In its strictest form, $f^k(t)$ is a symmetry if $\delta I = 0$ for all paths $q^k(t)$. We can be a bit looser, however, and let the difference include a "boundary term":

$$\delta I[q^k(t), f^k(t)] = I[q^k(t) + f^k(t)] - I[q^k(t)] = \int dt \frac{dK}{dt} \tag{17.3}$$

This is the same as modifying the L by adding a term dK/dt. It should be emphasized that $f^k(t)$ is a symmetry if this is true *for arbitrary* paths $q^k(t)$.

In the literature (and subsequently in these lectures), $f^k(t)$ is usually written as $\delta q^k(t)$, which does remind us that it's supposed to be small. On the other hand, it tends to make us think it's somehow related to the $q^k(t)$. It's not. The functions $f^k(t)$ implement the transformation over which the action is invariant, whereas $q^k(t)$ represent any path (among which is the classical path which solves the Euler-Lagrange equations).

17.1.2 On-shell variation

In this case, we start from the classical path $q^k(t) = \bar{q}^k(t)$ which already solves the Euler-Lagrange equations.

$$0 = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right) - \frac{\partial L}{\partial q^k} \tag{17.4}$$

Then we take the variation $\delta q^k(t)$ to be arbitrary but small. (In this sense it's the opposite of the previous type of variation.)

The change in action is then

$$\delta I[q^k, \delta q^k] = I[q^k + \delta q^k] - I[q^k] \tag{17.5}$$

$$= \int dt \left(L(q^k, \dot{q}^k) + \frac{\partial L}{\partial q^k}\delta q^k + \frac{\partial L}{\partial \dot{q}^k}\delta \dot{q}^k - L(q^k, \dot{q}^k) \right) \tag{17.6}$$

We use integration by parts

$$\int dt \frac{d}{dt}(a(t)b(t)) = \int dt \frac{da}{dt} b + \int dt a \frac{db}{dt} \tag{17.7}$$

$$\int dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\delta q^k\right) = \int dt \frac{\partial L}{\partial \dot{q}^k}\frac{d}{dt}(\delta q^k) + \int dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right)\delta q^k \tag{17.8}$$

$$= \int dt \frac{\partial L}{\partial \dot{q}^k}\delta \dot{q}^k + \int dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right)\delta q^k \tag{17.9}$$

$$\int dt \frac{\partial L}{\partial \dot{q}^k}\delta \dot{q}^k = \int dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\delta q^k\right) - \int dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right)\delta q^k \tag{17.10}$$

to remove the time derivative in $\delta \dot{q}^k$

$$\delta I[q^k, \delta q^k] = \int dt \left[\frac{\partial L}{\partial q^k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right)\right]\delta q^k + \int dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\delta q^k\right) \tag{17.11}$$

$$= \int dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\delta q^k\right) \tag{17.12}$$

where we've used the fact that $q^k(t)$ solves the Euler-Lagrange equations.

17.1.3 Noether's Theorem

If we denote $f^k(t)$ as $\delta_s q^k(t)$ to indicate the action symmetry, and $\bar{q}^k(t)$ to denote the classical path,

$$\delta I[q^k(t), \delta_s q^k(t)] = \int dt \frac{dK}{dt} \tag{17.13}$$

$$\delta I[\bar{q}^k(t), \delta q^k(t)] = \int dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\delta q^k\right) \tag{17.14}$$

We now relate the form invariance to the on-shell variation by setting $q^k(t) = \bar{q}^k(t)$ and $\delta q^k(t) = \delta_s q^k(t)$ in the two formulae and equating the δI's.

$$\delta I[\bar{q}^k(t), \delta_s q^k(t)] = \int dt \left(\frac{dK}{dt}\right) = \int dt \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\delta_s q^k\right) \tag{17.15}$$

Setting the two δI's equal leads to this simplified version of Noether's First Theorem, that given a symmetry $\delta_s q^k(t)$, there is a quantity

$$Q = K - \frac{\partial L}{\partial \dot{q}^k}\delta_s q^k \tag{17.16}$$

which is conserved.

In practice, the Q calculated here will not be the conserved current as usually conceived; it will be a conserved current multipled by infinitesimal parameters of the symmetry. Since these are arbitrary parameters, we drop them and extract the conserved current itself. Some

treatments of the theorem (including Noether's own) carry these parameters (which we call ϵ_j for illustration) through by writing

$$K = \sum_j \frac{\partial K}{\partial \epsilon_j}\epsilon_j \tag{17.17}$$

$$\delta q^k = \sum_j \frac{\partial(\delta q^k)}{\partial \epsilon_j}\epsilon_j \tag{17.18}$$

up to the point where the two δI's are equated: at this point, the "arbitrary parameters" argument is invoked to equate each ϵ_j term separately and thus cancel the ϵ_j's. Nowadays, it is more common to drop the parameters at the very last stage.

17.1.4 Examples

We'll use examples from non-relativistic problems with an distance-dependent interaction potential. The Lagrangian is

$$L = \sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^{\,2} - \frac{1}{2}\sum_{j\neq i} V(|\mathbf{r}_i - \mathbf{r}_j|) \tag{17.19}$$

and its action

$$I[\mathbf{r}_i(t)] = \int dt \left(\sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^{\,2} - \frac{1}{2}\sum_{j\neq i} V(|\mathbf{r}_i - \mathbf{r}_j|) \right) \tag{17.20}$$

The second sum is taken over all pairs of i and j where $j \neq i$. Since the potential is symmetric with respect to each pair of bodies, the sum includes two identical terms for each pair, and thus the 1/2 factor in front.

Rotations

First, we note that the Lagrangian and action are invariant with respect to simultaneous rotations of all the $\mathbf{r}_i$,

$$\mathbf{r}_i(t) \to \mathbf{r}'_i(t) = \mathbf{R}\mathbf{r}_i(t) \tag{17.21}$$

where $\mathbf{R}$ is an orthogonal matrix. Clearly the potential is invariant with respect to such a rotation of all the bodies:

$$V(|\mathbf{r}'_i - \mathbf{r}'_j|) = V(|\mathbf{R}\mathbf{r}_i - \mathbf{R}\mathbf{r}_j|) \tag{17.22}$$

$$= V(|\mathbf{R}(\mathbf{r}_i - \mathbf{r}_j)|) \tag{17.23}$$

$$= V(|\mathbf{r}_i - \mathbf{r}_j|) \tag{17.24}$$

For the kinetic energy part, we need to rephrase the rotation as an infinitesimal change (since we're concerned with continuous symmetries). For small rotations,

$$\mathbf{R}\mathbf{r}_i = \mathbf{r}_i + \boldsymbol{\alpha} \wedge \mathbf{r}_i + O(\alpha^2) \tag{17.25}$$

where $\boldsymbol{\alpha}$ is a small, constant vector. The small change is then

$$\mathbf{f}_i(t) = \boldsymbol{\alpha} \wedge \mathbf{r}_i(t) \tag{17.26}$$

so that

$$\begin{aligned}
\dot{\mathbf{r}}_i'(t) &= \dot{\mathbf{r}}_i + \boldsymbol{\alpha} \wedge \dot{\mathbf{r}}_i && (17.27)\\
\dot{\mathbf{r}}_i'^2 &= \dot{\mathbf{r}}_i^2 + 2\dot{\mathbf{r}}_i \cdot (\boldsymbol{\alpha} \wedge \dot{\mathbf{r}}_i) + O(\alpha^2) && (17.28)\\
&= \dot{\mathbf{r}}_i^2 + 2\boldsymbol{\alpha} \cdot (\dot{\mathbf{r}}_i \wedge \dot{\mathbf{r}}_i) + O(\alpha^2) && (17.29)\\
&= \dot{\mathbf{r}}_i^2 + O(\alpha^2) && (17.30)
\end{aligned}$$

We see from this that the kinetic energy also remains unchanged to first order. The action is invariant, with the boundary term $K = 0$.

Since the boundary term vanishes, the conserved current comes from the on-shell variation,

$$\begin{aligned}
Q &= -\sum_{i,k} \frac{\partial L}{\partial \dot{r}_i^k} \delta_s r_i^k && (17.31)\\
&= -\sum_i m_i \dot{\mathbf{r}}_i \cdot (\boldsymbol{\alpha} \wedge \mathbf{r}_i) && (17.32)\\
&= \sum_i m_i \boldsymbol{\alpha} \cdot (\mathbf{r}_i \wedge \dot{\mathbf{r}}_i) && (17.33)\\
&= \boldsymbol{\alpha} \cdot \left(\sum_i m_i (\mathbf{r}_i \wedge \dot{\mathbf{r}}_i) \right) && (17.34)
\end{aligned}$$

Since $\boldsymbol{\alpha}$ is really three arbitrary constants, we find we have three conserved currents

$$\frac{d\mathbf{L}}{dt} = \frac{d}{dt} \left(\sum_i m_i \mathbf{r}_i \wedge \dot{\mathbf{r}}_i \right) = 0 \tag{17.35}$$

which is of course total angular momentum conservation for a distance-dependent interaction potential.

Translations

Using the same action as before, let

$$\begin{aligned}
\mathbf{r}_i(t) \to \mathbf{r}_i'(t) &= \mathbf{r}(t) - \epsilon \dot{\mathbf{r}}_i(t) && (17.36)\\
\mathbf{f}_i(t) &= -\epsilon \dot{\mathbf{r}}_i(t) && (17.37)
\end{aligned}$$

with ϵ a constant.

$$I[\mathbf{r}_i'] = \int dt \left(\sum_i \frac{1}{2} m_i (\dot{\mathbf{r}}_i - \epsilon \ddot{\mathbf{r}}_i)^2 - \frac{1}{2} \sum_{j \neq i} V(|\mathbf{r}_i - \epsilon \dot{\mathbf{r}}_i - \mathbf{r}_j + \epsilon \dot{\mathbf{r}}_j|) \right) \tag{17.38}$$

The first sum, to first order, is

$$\int dt \sum_i (\frac{1}{2} m_i \dot{\mathbf{r}}_i^2 - \epsilon m_i \dot{\mathbf{r}}_i \cdot \ddot{\mathbf{r}}_i) = \int dt \sum_i (\frac{1}{2} m_i \dot{\mathbf{r}}_i^2 - \frac{1}{2} \epsilon m_i \frac{d\dot{\mathbf{r}}_i}{dt}) \tag{17.39}$$

For the second (potential) term, it is useful to think of the potential as a function of $r_{ij}^2 = (\mathbf{r}_i - \mathbf{r}_j)^2$. For each pair i and j, we have to first order

$$\begin{aligned} V((\mathbf{r}_i - \mathbf{r}_j - \epsilon\dot{\mathbf{r}}_i + \epsilon\dot{\mathbf{r}}_j)^2) &= V((\mathbf{r}_{ij} - \epsilon\dot{\mathbf{r}}_{ij})^2) && (17.40)\\ &= V(\mathbf{r}_{ij}^2 - 2\epsilon \mathbf{r}_{ij} \cdot \dot{\mathbf{r}}_{ij} + O(\epsilon^2)) && (17.41)\\ &= V(r_{ij}^2) - 2\epsilon \mathbf{r}_{ij} \cdot \dot{\mathbf{r}}_{ij} \left(\frac{dV}{d(r_{ij}^2)} \right) && (17.42)\\ &= V(r_{ij}^2) - \frac{d(r_{ij}^2)}{dt} \left(\frac{dV}{d(r_{ij}^2)} \right) && (17.43)\\ &= V(r_{ij}^2) - \frac{dV}{dt} && (17.44) \end{aligned}$$

Combining and taking the difference,

$$\delta I[\mathbf{r}, \mathbf{f}] = -\epsilon \int dx \frac{d}{dt} \left(\sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 - \frac{1}{2} \sum_{j \neq i} V(r_{ij}^2) \right) \tag{17.45}$$

so this is a symmetry with boundary term

$$K = -\epsilon \left(\sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 - \frac{1}{2} \sum_{j \neq i} V(r_{ij}^2) \right) = -\epsilon L \tag{17.46}$$

The physical interpretation of this symmetry can be seen by considering a path which is displaced wholesale in time:

$$q'(t + \epsilon) = q(t) \tag{17.47}$$

In order to evaluate δI consistently, we need to find q' as a function of t, not $t + \epsilon$.

$$q'(t) = q(t - \epsilon) = q(t) - \epsilon \dot{q}(t) + O(\epsilon^2) \tag{17.48}$$

so we see that the $\mathbf{f}(t)$ given above is the infinitesimal version of a time translation.

For the on-shell variation, we have for each body i (and implicitly summing over the components k)

$$\frac{\partial L}{\partial \dot{r}_i^k} \delta_s r_i^k = m_i \dot{r}_i^k (-\epsilon \dot{r}_i^k) = -\epsilon m_i \dot{\mathbf{r}}_i^2 \tag{17.49}$$

from which follows the conserved current

$$\begin{aligned} Q &= -\epsilon \left(\sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 - \frac{1}{2} \sum_{j \neq i} V(r_{ij}^2) - \sum_i m_i \dot{\mathbf{r}}_i^2 \right) && (17.50)\\ &= \epsilon \left(\sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 + \frac{1}{2} \sum_{j \neq i} V(r_{ij}^2) \right) = \epsilon E && (17.51) \end{aligned}$$

from which we see that time invariance leads to the conservation of total energy in the central force problem.

Translations in Special Relativity

We can write down a relativistic Lagrangian with an interacting potential of the form

$$L = -\sum_i m_i c(-U_\mu U^\mu)^{1/2} - \sum_{j\neq i} V((x_i - x_j)_\mu (x_i - x_j)^\mu) \tag{17.52}$$

The first term is one of the usual "kinetic" terms. The second potential is a function of the invariant interval between two bodies. For instance, the potential could be a delta function which selects only those pairs which could be causally connected via a null vector. All of our earlier collision problems would fall in this category.

The translation takes the form

$$x_i'^\mu = x_i^\mu - \epsilon^\mu \tag{17.53}$$

where the ϵ^μ are four constant displacements. From this, it is easy to see that

$$U_i' = U_i \tag{17.54}$$
$$(x_i' - x_j') = (x_i - x_j) \tag{17.55}$$

and therefore the Lagrangian is form-invariant, so $K = 0$.

To evaluate the on-shell variation,

$$\frac{\partial(-U^\mu U_\mu)}{\partial U^\alpha} = \frac{\partial}{\partial U^\alpha}(-g_{\mu\nu}U^\mu U^\nu) \tag{17.56}$$
$$= -g_{\mu\nu}(\delta^\mu_\alpha U^\nu + U^\mu \delta^\nu_\alpha) \tag{17.57}$$
$$= -g_{\alpha\nu}U^\nu - g_{\mu\alpha}U^\mu \tag{17.58}$$
$$= -2g_{\alpha\mu}U^\mu \tag{17.59}$$

so we get for each U_i

$$\frac{\partial L}{\partial(U^\alpha)} = -m_i c(-U^\mu U_\mu)^{-1/2}\frac{1}{2}(-2g_{\alpha\nu}U^\nu) \tag{17.60}$$
$$= m_i g_{\alpha\nu}U^\nu \tag{17.61}$$
$$= P_\alpha \tag{17.62}$$

The conserved current is then

$$Q = K - \sum_i \frac{\partial L}{\partial(U_i^\alpha)}\delta_s(x_i)_\alpha \tag{17.63}$$

but since $\delta_s(x_i)_\alpha$ is simply ϵ_α, and (as in the case of the rotation example above) these are arbitrary constants, what we end up with are four conserved currents

$$Q_\alpha = \sum_i (P_i)_\alpha \tag{17.64}$$

Thus we see that from translational symmetry we get the 4-momentum conservation we mentioned at the very beginning of the course.

17.2 Noether's Theorem for classical fields

The change from discrete particles to fields is not difficult.

Some discussions of Noether's Theorem also discuss change of coordinates, but it's easiest to restrict ourselves to changes in fields. We've seen that a change in coordinates can be implemented as a change in fields, so we'll restrict ourselves to the latter category of variations. This also helps us distinguish from trivial changes which aren't connected to symmetries at all, such as changing or substituting for dummy variables in integrals.

A set of symmetries is a set of infinitesimal functions $\delta_s\phi(x)$ such that

$$\delta I[\phi, \delta_s\phi] = I[\phi(x) + \delta_s\phi(x)] - I[\phi(x)] = \int d^4x \partial_\mu K^\mu \tag{17.65}$$

for all $\phi(x)$. The last term corresponds to the "boundary term". Note that the $\delta_s\phi(x)$ don't involve any changes to coordinates (which in fields function as labels).

The on-shell variation starts from the action

$$I[\phi(x)] = \int d^4x \mathcal{L}(\phi, \partial_\mu\phi) \tag{17.66}$$

and the solution to the Euler-Lagrange equation:

$$0 = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} \right) - \frac{\partial \mathcal{L}}{\partial\phi} \tag{17.67}$$

We call the solution $\bar{\phi}(x)$, and then consider the arbitrary but small variation $\delta\phi(x)$.

$$\begin{aligned}
\delta I[\bar{\phi}, \delta\phi] &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi) \right) && (17.68)\\
&= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial\phi}\delta\phi - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} \right) \delta\phi \right) + \int d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi \right) && (17.69)\\
&= \int d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi \right) && (17.70)
\end{aligned}$$

where the last step takes advantage of the fact that $\bar{\phi}$ solves the Euler-Lagrange equation.

Equating the δI's as before, we get

$$0 = \int d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)}\delta_s\phi - K^\mu \right) = \int d^4x \partial_\mu J^\mu \tag{17.71}$$

where

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi(x) - K^\mu \tag{17.72}$$

is the conserved current. The derivative ∂_μ is, as before, a "total partial derivative". In practice, since J^μ is usually written as a function of the coordinates only, this equation is often seen in the form

$$0 = \partial_\mu J^\mu \tag{17.73}$$

In the case that J drops to zero at large distances (one needs to check this, as it isn't always true), we can write

$$0 = \partial_\mu J^\mu = \frac{\partial J^0}{\partial t} + \nabla \cdot \mathbf{J} \tag{17.74}$$

If we take the spatial volume integral, we have for the first term

$$\int_V d^3x \frac{\partial J^0}{\partial t} = \frac{d}{dt}\int_V d^3x J^0 = \frac{dQ}{dt} \tag{17.75}$$

In the first equality, the "total partial derivative" has become a full derivative, because we've integrated out the other coordinates. For the second term

$$\int_V d^3x \nabla \cdot \mathbf{J} = \int_S \mathbf{J} \cdot d\mathbf{a} = 0 \tag{17.76}$$

so we get

$$\frac{dQ}{dt} = 0 \tag{17.77}$$

with Q the conserved quantity.

17.2.1 Translations

Let's look at a scalar field with the Klein-Gordon Lagrangian:

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m^2\phi^2 \tag{17.78}$$

Consider the translation

$$x^\mu \to x^\mu + \epsilon^\mu \tag{17.79}$$

which becomes the field transformation

$$\phi(x^\mu) \to \phi(x^\mu - \epsilon^\mu) \approx \phi(x^\mu) - \epsilon^\mu\partial_\mu\phi(x^\mu) \tag{17.80}$$

The change in the Lagrangian density is then

$$\begin{aligned}
\delta\mathcal{L} &= -\frac{1}{2}(\partial_\mu(\phi - \epsilon^\nu\partial_\nu\phi))(\partial^\mu(\phi - \epsilon^\lambda\partial_\lambda\phi)) - \frac{1}{2}m^2(\phi - \epsilon^\nu\partial_\nu\phi)^2 && (17.81)\\
&\quad + \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) + \frac{1}{2}m^2\phi^2 && (17.82)\\
&= -\frac{1}{2}(\partial_\mu\phi - \epsilon^\nu\partial_\nu\partial_\mu\phi)(\partial^\mu\phi - \epsilon^\lambda\partial_\lambda\partial_\mu\phi) - \frac{1}{2}m^2(\phi - \epsilon^\nu\partial_\nu\phi)^2 && (17.83)\\
&\quad + \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) + \frac{1}{2}m^2\phi^2 && (17.84)\\
&= \epsilon^\nu(\partial_\nu\partial_\mu\phi)(\partial^\mu\phi) + m^2\epsilon^\nu\phi\partial_\nu\phi && (17.85)\\
&= \epsilon^\nu[(\partial_\nu\partial_\mu\phi)(\partial^\mu\phi) + m^2\phi\partial_\nu\phi] && (17.86)\\
&= \partial_\nu[\frac{1}{2}\epsilon^\nu(\partial_\mu\phi)(\partial^\mu\phi) + \frac{1}{2}m^2\epsilon^\nu\phi^2] && (17.87)\\
&= \partial_\nu[-\epsilon^\nu\mathcal{L}] && (17.88)
\end{aligned}$$

Which gives a boundary term

$$K^\lambda = -\epsilon^\lambda \mathcal{L} \tag{17.89}$$

and therefore a conserved current

$$J^\lambda = \frac{\partial \mathcal{L}}{\partial(\partial_\lambda \phi)}\delta\phi - K^\lambda = (-\partial^\lambda \phi)(-\epsilon^\mu \partial_\mu \phi) + \epsilon^\lambda \mathcal{L} = \epsilon^\mu T^\lambda{}_\mu \tag{17.90}$$

where

$$\begin{aligned} T^\lambda{}_\mu &= (\partial^\lambda \phi)(\partial_\mu \phi) + \delta^\lambda_\mu \mathcal{L} && (17.91)\\ T^{\lambda\mu} &= (\partial^\lambda \phi)(\partial^\mu \phi) + g^{\lambda\mu}\mathcal{L} && (17.92) \end{aligned}$$

which is the "canonical" energy-momentum (stress-energy) tensor. The conservation law takes the form

$$0 = \partial_\mu T^{\mu\nu} \tag{17.93}$$

which is actually four conservation laws (conserved currents) at once.

17.2.2 Complex fields

We saw earlier that the Klein-Gorden equation can be used to describe two independent fields with identical mass. We can look at them from the perspective of Noether's Theorem to see more of their relationship.

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu \phi^*)(\partial^\mu \phi) - \frac{1}{2}m^2 \phi^* \phi \tag{17.94}$$

A unitary transformation of the fields gives for the two fields

$$\begin{aligned} \phi' &= e^{i\lambda}\phi && (17.95)\\ \phi^{*\prime} &= e^{-i\lambda}\phi^* && (17.96) \end{aligned}$$

(I happen to be giving priority to the asterisk rather than the ', since we're really consider ϕ^* as an independent field which we happening to be transforming.)

If we turn these into infinitesimal changes,

$$\begin{aligned} \delta\phi &= i\lambda\phi && (17.97)\\ \delta\phi^* &= -i\lambda\phi^* && (17.98) \end{aligned}$$

Pauli called this a "gauge transformation of the first kind".

The form variation in the Lagrangian density is

$$\begin{aligned} \delta\mathcal{L}[\phi, \phi^*, \phi', \phi^{*\prime}] &= \mathcal{L}[\phi + \delta\phi, \phi^* + \delta\phi^*] - \mathcal{L}[\phi, \phi^*] && (17.99)\\ &= \mathcal{L}[\phi + i\lambda\phi, \phi^* - i\lambda\phi^*] - \mathcal{L}[\phi, \phi^*] && (17.100)\\ &= -\frac{1}{2}(\partial_\mu(\phi^* - i\lambda\phi^*))(\partial^\mu(\phi + i\lambda\phi)) && (17.101)\\ &\quad -\frac{1}{2}m^2(\phi^* - i\lambda\phi^*)(\phi + i\lambda\phi) && (17.102)\\ &= \mathcal{L}\cdot(1 + \lambda^2) && (17.103)\\ &= \mathcal{L} + O(\lambda^2) && (17.104) \end{aligned}$$

The on-shell variation in the Lagrangian density is

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi) + \frac{\partial\mathcal{L}}{\partial\phi^*}\delta\phi^* + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}\delta(\partial_\mu\phi^*) \tag{17.105}$$

We note that

$$\delta\left(\frac{\partial\phi}{\partial x^\mu}\right) = \frac{\partial(\delta\phi)}{\partial x^\mu} \tag{17.106}$$

and thus the second term is

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\frac{\partial(\delta\phi)}{\partial x^\mu} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right) - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\delta\phi \tag{17.107}$$

and similarly for the fourth term. Assuming then that ϕ and ϕ^* already satisfy the Euler-Lagrange equations,

$$\begin{aligned}
\delta\mathcal{L} &= \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right) + \left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\right]\delta\phi && (17.108)\\
&\quad + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}\delta\phi^*\right) + \left[\frac{\partial\mathcal{L}}{\partial\phi^*} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}\right)\right]\delta\phi^* && (17.109)\\
&= \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}\delta\phi^*\right) && (17.110)\\
&= -\frac{i}{2}\lambda\partial_\mu[(\partial^\mu\phi^*)\phi - \phi^*(\partial^\mu\phi)] && (17.111)
\end{aligned}$$

Setting $\delta\mathcal{L} = 0$ (since we showed form invariance without boundary terms above), we have a conserved 4-current

$$\begin{aligned}
0 &= \partial_\mu s^\mu && (17.112)\\
s^\mu &= i\left(\frac{\partial\phi^*}{\partial x^\mu}\phi - \phi^*\frac{\partial\phi}{\partial x^\mu}\right) && (17.113)
\end{aligned}$$

We also see that if we swap ϕ and ϕ^*, s^μ changes sign. So we see that the fields have not only identical mass but also opposite "charge", something we'd expect for fields of particles and anti-particles, and that relativistic field theory readily accommodates them. (Note that this argument isn't limited to electric charge; any conserved internal attribute with complex field will do. It also only works for complex fields; real-valued fields would have no such conserved current.)

We can do a similar operation with the Dirac Lagrangian:

$$\mathcal{L} = \bar{\psi}i\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \tag{17.114}$$

The transformation takes the same form

$$\begin{aligned}
\delta\psi &= i\lambda\psi && (17.115)\\
\delta\bar{\psi} &= -i\lambda\bar{\psi} && (17.116)
\end{aligned}$$

The Lagrangian is clearly form-invariant. The on-shell variation follows

$$\begin{aligned}\delta\mathcal{L} &= \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\delta\psi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\bar{\psi})}\delta\bar{\psi}\right) && (17.117)\\ &= \partial_\mu(\bar{\psi}i\gamma^\mu i\lambda\psi) && (17.118)\\ &= -\lambda\partial_\mu(\bar{\psi}\gamma^\mu\psi) && (17.119)\end{aligned}$$

so the conserved current is proportional to

$$s^\mu = \bar{\psi}\gamma^\mu\psi \tag{17.120}$$

If in a future course this Dirac current form pops out at you, now you know where it comes from.

17.2.3 Maxwell's equations

We can also obtain energy-momentum tensor of the electromagnetic field by considering their translational symmetry. The Lagrangian density is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{17.121}$$

We could then vary the 4-potential field A_μ to find the change in the action directly, and many traditional derivations do this. But in light of the need to preserve gauge invariance, we calculate the form variation in terms of the field tensor.

As before, we consider the translation of field component values from one spacetime point to another, displayed by a constant 4-vector ϵ^α.

$$\begin{aligned}F'_{\mu\nu}(x) &= F_{\mu\nu}(x-\epsilon) && (17.122)\\ &= \partial_\mu A_\nu(x-\epsilon) - \partial_\nu A_\mu(x-\epsilon) && (17.123)\end{aligned}$$

The (small) variation in A_μ is

$$\begin{aligned}A'_\mu(x) &= A_\mu(x-\epsilon) = A_\mu(x) - \epsilon^\alpha\partial_\alpha A_\mu(x) && (17.124)\\ \delta A_\mu(x) &= A'_\mu(x) - A_\mu(x) = -\epsilon^\alpha\partial_\alpha A_\mu(x) && (17.125)\end{aligned}$$

At this point, we notice the old problem of gauge invariance: since the change in A_μ above isn't invariant, we might be concerned that the conserved current we obtain from it won't be, either. Indeed, most textbooks will push ahead to get a "canonical stress-energy tensor" which, among other undesirable qualities, is asymmetric. Then they'll show that the physics of the tensor admits the addition of a divergence term (rather like a gauge transform), and that a symmetric tensor, yielding identical physics, can thus always be derived.

We'll take a simpler approach by defining an "improved" variation which combines the spacetime translation with a gauge transformation:

$$\delta A_\mu = -\epsilon^\alpha(\partial_\alpha A_\mu) + \partial_\mu(\partial^\alpha A_\alpha) = F_{\mu\alpha}\epsilon^\alpha \tag{17.126}$$

This is now gauge invariant. The transformed field tensor is then (taking all fields as evaluated at x)

$$\begin{aligned} F'_{\mu\nu} &= \partial_\mu(A_\nu + F_{\nu\alpha}\epsilon^\alpha) - \partial_\nu(A_\mu + F_{\mu\alpha}\epsilon^\alpha) && (17.127)\\ &= F_{\mu\nu} + \epsilon^\alpha(\partial_\mu F_{\nu\alpha} - \partial_\nu F_{\mu\alpha}) && (17.128)\\ \delta F_{\mu\nu} &= \epsilon^\alpha(\partial_\mu F_{\nu\alpha} - \partial_\nu F_{\mu\alpha}) && (17.129)\\ &= \epsilon^\alpha(\partial_\mu F_{\nu\alpha} + \partial_\nu F_{\alpha\mu}) && (17.130)\\ & && (17.131)\end{aligned}$$

We can pull in Bianchi's identity

$$0 = \partial_\mu F_{\nu\alpha} + \partial_\nu F_{\alpha\mu} + \partial_\alpha F_{\mu\nu} \tag{17.132}$$

to get

$$\begin{aligned} \delta F_{\mu\nu} &= -\epsilon^\alpha \partial_\alpha F_{\mu\nu} && (17.133)\\ \delta(F_{\mu\nu}F^{\mu\nu}) &= 2F^{\mu\nu}(\delta F_{\mu\nu}) && (17.134)\\ &= -2\epsilon^\alpha F^{\mu\nu}\partial_\alpha F_{\mu\nu} && (17.135)\\ &= -\epsilon^\alpha \partial_\alpha(F^{\mu\nu}F_{\mu\nu}) && (17.136)\end{aligned}$$

The change in the form of the action is then

$$\begin{aligned} \delta I &= -\frac{1}{4}\int d^4x \delta(F_{\mu\nu}F^{\mu\nu}) && (17.137)\\ &= \frac{1}{4}\int d^4x \partial_\alpha(\epsilon^\alpha F_{\mu\nu}F^{\mu\nu}) && (17.138)\end{aligned}$$

The divergence term is then

$$K^\alpha = \frac{1}{4}\epsilon^\alpha F_{\mu\nu}F^{\mu\nu} = -\epsilon^\alpha \mathcal{L} \tag{17.139}$$

The on-shell variation is calculated with respect to the stationary configuration of A_μ.

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial(\partial_\alpha A_\beta)} &= -\frac{1}{2}F^{\mu\nu}\frac{\partial F_{\mu\nu}}{\partial(\partial_\alpha A_\beta)} && (17.140)\\ &= -\frac{1}{2}F^{\mu\nu}\frac{\partial}{\partial(\partial_\alpha A_\beta)}(\partial_\mu A_\nu - \partial_\nu A_\mu) && (17.141)\\ &= -\frac{1}{2}F^{\mu\nu}(\delta^\alpha_\mu\delta^\beta_\nu - \delta^\alpha_\nu\delta^\beta_\mu) && (17.142)\\ &= \frac{1}{2}(F^{\beta\alpha} - F^{\alpha\beta}) && (17.143)\\ &= F^{\beta\alpha} && (17.144)\end{aligned}$$

The conserved current is therefore

$$\begin{aligned} J^\alpha &= \frac{\partial\mathcal{L}}{\partial(\partial_\alpha A_\beta)}\delta A_\beta - K^\alpha && (17.145)\\ &= F^{\beta\alpha}F_{\beta\gamma}\epsilon^\gamma - \frac{1}{4}\epsilon^\alpha F_{\mu\nu}F^{\mu\nu} && (17.146)\end{aligned}$$

where we've used the "improved" transformation of A_μ. We then gather the terms of each independent variation parameter ϵ^γ:

$$J^\alpha = -F^{\alpha\beta}F_{\beta\gamma}\epsilon^\gamma - \frac{1}{4}\epsilon^\gamma\delta^\alpha_\gamma F_{\mu\nu}F^{\mu\nu} = T^\alpha{}_\gamma\epsilon^\gamma \tag{17.147}$$

We can recast this in more familiar form as follows:

$$\begin{aligned} T^{\alpha\kappa} &= T^\alpha{}_\gamma g^{\gamma\kappa} && (17.148)\\ &= (-F^{\alpha\beta}F_{\beta\gamma} - \frac{1}{4}\delta^\alpha_\gamma F_{\mu\nu}F^{\mu\nu})g^{\gamma\kappa} && (17.149)\\ &= -F^{\alpha\beta}F_\beta{}^\kappa - \frac{1}{4}g^{\alpha\kappa}F_{\mu\nu}F^{\mu\nu} && (17.150)\\ &= -F^\alpha{}_\beta F^{\beta\kappa} - \frac{1}{4}g^{\alpha\kappa}F_{\mu\nu}F^{\mu\nu} && (17.151) \end{aligned}$$

which is the symmetric stress-energy tensor we saw before.

17.3 Local gauge invariance

Finally, to close with one loose end. This is related to Noether's second theorem. We don't have time to go into it, but here's an illustration of the role of what we might call an "internal" symmetry.

We saw earlier how, when we added the electromagnetic interaction to the Lagrangian density, we ended up with a density which wasn't gauge invariant:

$$\mathcal{L} = -\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu} + J^\mu A_\mu \tag{17.152}$$

At the time, we appealed to the idea that gauge invariance was a "weak" condition which could be fixed up later. And indeed when we derived the field equations of motion, the problem disappeared because the field derivative removed the offending A_μ.

$$\partial_\beta F^{\alpha\beta} = \mu_0 J^\alpha \tag{17.153}$$

We aren't quite so lucky when we try to write down interactions with other fields. When we added the interaction to the single-particle Lagrangian, we added the term $qU^\mu A_\mu$, which then entered the equation of motion as a term qA_μ added to the momentum. Similarly, if we add the electromagnetic interaction to the Dirac equation, we end up with

$$(i\gamma^\mu\partial_\mu + q\gamma^\mu A_\mu - m)\psi = 0 \tag{17.154}$$

which would appear to change if we apply a gauge transformation

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu\chi \tag{17.155}$$

Now let's take a step into quantum mechanics. Recall that with quantum states (and this applies to quantized fields), we can multiply the quantum state by a phase, and it shouldn't affect the physics.

One may then argue that there's no particular reason the phase has to be same everywhere.

$$\psi(x) \to \psi'(x) = e^{iq\alpha(x)}\psi(x) \tag{17.156}$$

Let's see how this transformation affects the Dirac equation:

$$\begin{aligned}
(i\gamma^\mu\partial_\mu - m)\psi' &= i\gamma^\mu\partial_\mu(e^{iq\alpha}\psi) - me^{iq\alpha}\psi && (17.157)\\
&= -q\gamma^\mu(\partial_\mu\alpha)e^{iq\alpha}\psi + ie^{iq\alpha}\gamma^\mu(\partial_\mu\psi) - me^{iq\alpha}\psi && (17.158)\\
&= e^{iq\alpha}(i\gamma^\mu\partial_\mu - q\gamma^\mu(\partial_\mu\alpha) - m)\psi && (17.159)
\end{aligned}$$

The added term looks rather like the term you get with a gauge transformation of A_μ. Let's add that interaction back in and transform it at the same time:

$$\begin{aligned}
(i\gamma^\mu\partial_\mu + q\gamma^\mu A'_\mu - m)\psi' &= -q\gamma^\mu(\partial_\mu\alpha)e^{iq\alpha}\psi + ie^{iq\alpha}\gamma^\mu(\partial_\mu\psi) - me^{iq\alpha}\psi && (17.160)\\
&\quad +q\gamma^\mu A_\mu e^{iq\alpha}\psi + q\gamma^\mu(\partial_\mu\chi)e^{iq\alpha}\psi && (17.161)\\
&= e^{iq\alpha}(i\gamma^\mu\partial_\mu + q\gamma^\mu A_\mu - m)\psi && (17.162)\\
&\quad +q\gamma^\mu(\partial_\mu\chi - \partial_\mu\alpha)e^{iq\alpha}\psi && (17.163)
\end{aligned}$$

So we find that we can absorb the effect of the gauge transformation into the local phase of the quantum state if we simply set $\alpha(x)$ to $\chi(x)$.

For this reason it is sometimes said that the internal phase symmetry underlies the electromagnetic field, which is carried by the photon field A_μ. The symmetry group is that of 1×1 unitary "matrices", $U(1)$, applied to the quantum states. And indeed this seems to work not just for $U(1)$, but independently for $SU(2)$ and $SU(3)$. The current Standard Model is thus described in terms of the direct product of three Lie groups, $U(1) \times SU(2) \times SU(3)$.

Some claim that this direct product structure is too complicated, and that there must be a covering group. The smallest such covering group is $SU(5)$. This argument then becomes the basis of what are called "Grand Unified Theories". Which is another theory with as yet no experimental evidence. Nevertheless, it remains a compelling idea, inspired by symmetries for which there is ample evidence.

B2: Symmetry and Relativity

J Tseng

December 29, 2019

Contents

1 Vectors and tensors **4**

1.1 Books and links . . . 4

1.2 Postulates of Special Relativity . . . 4

1.3 Vectors . . . 4

1.3.1 Summation convention . . . 6

1.3.2 Coordinate transformation matrices . . . 6

1.4 Tensors . . . 6

1.4.1 Tensor algebra . . . 7

1.4.2 Special tensors . . . 7

1.4.3 The metric . . . 7

1.4.4 Levi-Civita tensor . . . 8

1.4.5 Caveats . . . 8

1.5 3D Physics tensors . . . 9

1.5.1 Angular momentum . . . 9

1.5.2 Moment of inertia . . . 9

1.5.3 Stress . . . 9

2 Special Relativity **10**

2.1 Definitions . . . 10

2.2 Units . . . 10

2.3 Minkowski space . . . 10

2.4 4-vectors . . . 12

2.5 4-momentum conservation . . . 13

2.5.1 Particle decay . . . 13

2.5.2 Particle annihilation and formation . . . 14

2.5.3 Compton scattering . . . 14
2.6 Mechanics 4-tensors . . . 14
2.6.1 Angular momentum . . . 14
2.6.2 Stress-energy tensor . . . 14
2.6.3 Angular momentum density . . . 15

3 Forces and fields **16**
3.1 Force . . . 16
3.1.1 Transformation of force . . . 16
3.1.2 Motion with a given force . . . 17
3.1.3 Uniform circular motion . . . 17
3.1.4 Linear motion . . . 17
3.2 Electromagnetic fields . . . 18
3.2.1 Field tensor . . . 18
3.2.2 4-potential . . . 19
3.2.3 Energy and momentum in vacuum . . . 19
3.2.4 Plane waves in vacuum . . . 20
3.2.5 Stress-energy tensor . . . 20

4 Radiation **22**
4.1 Uniformly moving charge . . . 22
4.2 Retarded potentials . . . 22
4.3 Accelerated charge . . . 23
4.3.1 Slowly oscillating dipole . . . 24
4.3.2 Radiated power . . . 24
4.4 Additional material . . . 25

5 Groups **26**
5.1 Permutation group . . . 26
5.2 Representation . . . 27
5.2.1 Reducibility . . . 27
5.3 Rotations . . . 28
5.3.1 SU(2) . . . 28
5.4 Lorentz group . . . 29

5.4.1 Generators of Lorentz boosts 29
5.4.2 Poincaré group 30
5.5 Gauge invariance 31

Chapter 1

Vectors and tensors

1.1 Books and links

The main text for the course is AM Steane, *Relativity Made Relatively Easy* (OUP, 2012).

Additional details can be found in JD Jackson's *Classical Electrodynamics*.

The webpage for the course can be found at
`http://www-pnp.physics.ox.ac.uk/ tseng/teaching/b2/`.

The page includes problem sets, links to related courses and materials, and old lecture notes with detailed derivations. The old lecture notes follow the older syllabus, so in some areas contain material which is beyond the current course

1.2 Postulates of Special Relativity

1. Principle of Relativity (two versions):
   - Laws of physics take the same mathematical form in all inertial frames of reference.
   - Motion of bodies in a given space are the same among themselves, whether the space is at rest or moving uniformly.
2. Light speed postulate: there is a maximum speed for signals.
3. Space dimensions are flat ("Euclidean")

1.3 Vectors

A *point* is used in two ways:

- geometrically, it's a location, independent of how the location is described;

- a set of N variables (*coordinates* which describe a location, conventionally

$$x^1, x^2, \cdots x^N$$

  The superscript is also by convention (not an exponent).

A *coordinate system* is a mapping between sets of coordinates and the geometric locations. The sets form an N-dimensional vector space.

We typically expect that the laws of physics shouldn't depend on the choice of coordinate system.

Note that we are using the word "location" in a broad sense: it can be a spatial location, but if combined with a time coordinate, it's a spacetime point, sometimes called an *event*.

A *coordinate transformation* is a set of N functions of x^i

$$\overline{x}^i = \overline{x}^i(x^1, x^2, \cdots x^N)$$

The transformation is usually invertible.

A *contravariant vector field* is a set of N functions of x^i

$$A^i = A^i(x^1, x^2, \cdots x^N)$$

which transforms as follows under a coordinate transformation:

$$\overline{A}^i = \sum_{j=1}^{N} \frac{\partial \overline{x}^i}{\partial x^j} A^j(x^1, x^2, \cdots x^N)$$

A *covariant vector field* is a set of N functions of x^i

$$B_i = B_i(x^1, x^2, \cdots x^N)$$

which transforms as follows under a coordinate transformation:

$$\overline{B}_i = \sum_{j=1}^{N} \frac{\partial x^j}{\partial \overline{x}^i} B_j(x^1, x^2, \cdots x^N)$$

The scalar product of A^i and B_i is

$$\sum_{i=1}^{N} A^i B_i$$

which is invariant. Contravariant and covariant vectors form dual linear spaces.

Length (squared) is defined as the scalar product of a vector with itself

$$L^2 = \sum_{i=1}^{N} A^i A_i$$

1.3.1 Summation convention

1. Sum over an index which appears as both an "upper" (contravariant) and "lower" (covariant) index, from 1 to N.
2. A differential with respect to a contravariant coordinate ($\partial/\partial x^i$) counts as a covariant index, and vice versa.

$$\overline{A}^i = \sum_{j=1}^{N} \frac{\partial \overline{x}^i}{\partial x^j} A^j \quad \Rightarrow \quad \overline{A}^i = \frac{\partial \overline{x}^i}{\partial x^j} A^j$$
$$\sum_{i=1}^{N} A^i B_i \quad \Rightarrow \quad A^i B_i$$

When we begin to work explicitly with Minkowski space for Special Relativity, we will use different index types as a shorthand for their range:

1. Latin indices will range from 1 to 3, encompassing the spatial dimensions.
2. Greek indices will range from 0 to 3, with 0 corresponding to the time dimension.

1.3.2 Coordinate transformation matrices

The coordinate transformation is often written in terms of a matrix

$$\begin{aligned} \Lambda^i{}_j &= \frac{\partial \overline{x}^i}{\partial x^j} \\ \overline{A}^i &= \Lambda^i{}_j A^j \\ \overline{A}_i &= A_j (\Lambda^{-1})^j{}_i \end{aligned}$$

(The order of the last equation is simply to write it more obviously as a matrix multiplication.)

1.4 Tensors

A *rank-2 tensor* is a set of N^2 functions with the following transformation rules:

- contravariant tensor
$$\overline{C}^{ij} = \frac{\partial \overline{x}^i}{\partial x^a} \frac{\partial \overline{x}^j}{\partial x^b} C^{ab}$$
- covariant tensor
$$\overline{C}_{ij} = \frac{\partial x^a}{\partial \overline{x}^i} \frac{\partial x^b}{\partial \overline{x}^j} C_{ab}$$

- mixed tensor

$$\overline{C}^{i}{}_{j} = \frac{\partial \overline{x}^{i}}{\partial x^{a}} \frac{\partial x^{b}}{\partial \overline{x}^{j}} C^{a}{}_{b}$$

Higher rank tensors generalize these transformations. Vectors are rank-1 tensors. Tensors are said to be of the *same kind* when they have the same number and order (and type) of indices.

1.4.1 Tensor algebra

Tensors of the same kind form a linear space.

The *outer product* of two tensors of rank s and r is another tensor of rank $s + r$:

$$T^{i}{}_{j} S^{k} = C^{i}{}_{j}{}^{k}$$

A tensor of rank s can be *contracted* by summing over a pair of upper/lower indices, producing a tensor of rank $s - 2$:

$$A^{i}{}_{ji} = D_{j}$$

Quotient rule: if the inner product of N^p functions A (with appropriate upper and lower indices) with an *arbitrary* tensor B is another tensor C then A is a rank-p tensor.

1.4.2 Special tensors

A *symmetric tensor* is one where two indices can be swapped, *e.g.*, $A^{i}{}_{jk} = A^{i}{}_{kj}$ if A is symmetric in the jk indices.

A *skew-symmetric tensor*, sometimes called an *anti-symmetric tensor*, flips sign when the specified indices are swapped, *e.g.*, $A^{i}{}_{jk} = -A^{i}{}_{kj}$.

The Kronecker delta δ^{i}_{j} is a rank-2 mixed tensor. It is also *isotropic* in that it doesn't change under coordinate transformations.

The related Kronecker delta δ_{ij} is isotropic under linear transformations in a Euclidean space, but is not a tensor in general.

1.4.3 The metric

The *metric* or *fundamental tensor* g_{ij} defines the line element

$$ds^2 = g_{ij} dx^{i} dx^{j}$$

It is also isotropic under linear transformations.

The *reciprocal fundamental tensor* is g^{ij} such that $g_{ij} g^{jk} = \delta^{k}_{i}$. (In general, $A_{ij} A^{jk} \neq \delta^{k}_{i}$.)

A *Cartesian coordinate system* is one in which the components g_{ij} of the metric tensor are constants.

A *Euclidean space* is a space for which there exists a coordinate transformation to a system where $ds^2 = \delta_{ij} dx^i dx^j$.

Index manipulation

The fundamental tensor allows one to raise and lower indices.

$$\begin{aligned} g_{ij} T^{abjc}{}_{d} &= T^{ab}{}_{i}{}^{c}{}_{d} \\ g^{ij} T^{ab}{}_{j}{}^{c}{}_{d} &= T^{abic}{}_{d} \end{aligned}$$

One can also flip the upper/lower indices in a contraction:

$$T^{i}{}_{i} = T_{i}{}^{i}$$

1.4.4 Levi-Civita tensor

The *Levi-Civita symbol* e_{ijk} is defined to be skew-symmetric in all pairs of its indices, with $e_{123} = +1$. It is not a tensor in general.

The *Levi-Civita tensor* is defined to be $\epsilon_{ijk} = \sqrt{|g|} e_{ijk}$, where $|g|$ is the determinant of the metric, considered as a matrix.

For Cartesian coordinate systems in Euclidean space, $|g| = 1$.

1.4.5 Caveats

The following caveats apply to more general coordinate systems and transformations:

1. In non-Cartesian coordinate systems, the components g_{ij} are in general not equal to the components g^{ij}.
2. Even though the gradient of a scalar function $\partial\phi/\partial x^i$ (which we can write more succinctly as $\partial_i \phi$) transforms as a covariant vector, the gradient of a vector $\partial_i A^j$ is not in general a tensor.

However, in this course we are mostly concerned with a Cartesian coordinate system of a Minkowski space in which the metric is known

$$g_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

and we limit ourselves to linear transformations. In this case, the components of $g^{\mu\nu}$ are identical to $g_{\mu\nu}$, and $\partial_\mu A^\nu$ transforms as a rank-2 mixed tensor.

1.5 3D Physics tensors

Examples of 3D tensors in physics. Note that since these reside in Euclidean space, contravariant and covariant components transform in the same way, so the upper and lower indices here are used here mostly to imply summation.

1.5.1 Angular momentum

Angular momentum can be written as a skew-symmetric outer product of position and momentum.

$$L^{ij} = x^i p^j - x^j p^i$$

The angular momentum vector is related to the tensor:

$$L_i = \frac{1}{2}\epsilon_{ijk}L^{jk}.$$

In matrix form,

$$L^{ij} = \begin{pmatrix} 0 & L_z & -L_y \\ -L_z & 0 & L_x \\ L_y & -L_x & 0 \end{pmatrix}$$

1.5.2 Moment of inertia

Moment of inertia for a rigid body:

$$I^{ij} = \int_V d^3x\rho(x)[(x^k x_k)\delta^{ij} - x^i x^j]$$

Angular momentum:

$$L^i = I^{ij}\omega_j$$

where ω_j is the angular velocity vector.

1.5.3 Stress

The stress tensor σ^{ij} can be thought of as indicating the flow of the momentum component p^i across face j of a volume. "Face j" refers to the face which is normal to the j direction.

Stress tensor for an ideal fluid:

$$\sigma^{ij} = \begin{pmatrix} p & & \\ & p & \\ & & p \end{pmatrix}$$

Euler equation for incompressible fluids:

$$\frac{\partial u^i}{\partial t} + u^j\partial_j u^i = \frac{1}{\rho}\partial_j\sigma^{ij}$$

where $u^i(x^j)$ is the velocity of each point in the fluid.

Chapter 2

Special Relativity

2.1 Definitions

A *frame* is a coordinate system with time and space.

An *inertial frame* is a frame moving at constant velocity relative to another frame. A special case (for convenience) is the "standard configuration", in which the inertial frame moves in the x direction relative to the other frame.

An *event* or *spacetime point* is a point specified with space and time. Note there is a distinction between an event and an *observation event*: observation also happens at a location and time, from which one can then infer the coordinates of the original event.

A *spacetime diagram* is a diagram with time (or ct to have units of length) as one axis, conventionally vertical.

If t, x, y, and z are the usual component values which specify an event, the contravariant 4-vector with Cartesian coordinates is

$$(x^0, x^1, x^2, x^3) = (ct, x, y, z) = (ct, \mathbf{x})$$

2.2 Units

Components of a vector should have the same units, *e.g.*, $x^0 = ct$.

"Natural units" with $c = 1$ (and $\hbar = 1$ for that matter) are often used. One then reintroduces the units at the end by multiplying by as many factors of c and $\hbar$ to get the units you find useful.

2.3 Minkowski space

Metric tensor, in Cartesian coordinates, of the *Minkowski space* of special relativity is $g_{00} = -1$, $g_{11} = g_{22} = g_{33} = 1$. This is sometimes written diag$(-+++)$. In matrix form, as shown

before,

$$g_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

Spacetime length is then

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$$

A *time-like* interval has $ds^2 < 0$. A *space-like* interval has $ds^2 > 0$. A *null* interval has $ds^2 = 0$.

Warning: the subatomic physics paper B4 uses a different sign convention for the metric, $g_{\mu\nu} = \text{diag}(- + ++)$. The difference persists because when we want to focus on geometry, we like space-like intervals to have positive length (as they have in Euclidean space). If our focus is on momentum and energy (see Section 2.4), we may prefer the magnitude of the vector to reflect positive mass and energy.

In a spacetime diagram, a *light cone* separates time-like from space-like intervals. A *worldline* is a trajectory $x^\mu(\tau)$, where τ is a real parameter, and each point on the worldline is causally connected, *i.e.*, separated by a time-like interval.

The Minkowski metric is isotropic for

1. spatial rotations, *e.g.*,

$$\begin{aligned} \overline{x}^1 &= x^1\cos\theta + x^2\sin\theta \\ \overline{x}^2 &= -x^1\sin\theta + x^2\cos\theta \end{aligned}$$

2. Lorentz transformations (*boosts*), *e.g.*,

$$\begin{aligned} \overline{x}^0 &= x^0\cosh\eta - x^1\sinh\eta \\ \overline{x}^1 &= -x^0\sinh\eta + x^1\cosh\eta \end{aligned}$$

   where η is the *rapidity* of the transformation.

3. spatial inversion (parity transformation): $\overline{x}^i = -x^i$ for each $i = 1, 2, 3$.

4. time reversal: $\overline{x}^0 = -x^0$.

The relationship between rapidity and the usual parameters of the Lorentz transformation are $\gamma = \cosh\eta$ and $\gamma\beta = \sinh\eta$.

Two successive, aligned Lorentz transformations η_1 and η_2 give a combined Lorentz transformation with rapidity $\eta_1 + \eta_2$. The velocities therefore add

$$\beta_3 = \tanh(\eta_1 + \eta_2) = \frac{\tanh\eta_1 + \tanh\eta_2}{1 + \tanh\eta_1\tanh\eta_2} = \frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2}$$

Since lengths are invariants, boosts do not change the time-like, space-like, or null characteristic of an interval.

Proper time is the time experienced by a body in its own rest frame. An interval in proper time is always the minimum amount of time passing; in any other frame, a larger amount of time will pass. For a body along its worldline, $dt = \gamma d\tau$, where $d\tau$ is the interval in proper time, and dt is the interval in another frame.

2.4 4-vectors

1. Displacement:
$$x^\mu = (ct, \mathbf{x})$$

2. 4-velocity:
$$U^\mu = \frac{dx^\mu}{d\tau} = \gamma(c, \mathbf{u})$$
where $\mathbf{u}$ is the 3-velocity.
$$U_\mu U^\mu = -c^2$$

3. 4-momentum:
$$P^\mu = mU^\mu = \gamma m(c, \mathbf{u}) = (E/c, \mathbf{p})$$
where E is energy and $\mathbf{p}$ the 3-momentum.
$$P_\mu P^\mu = -m^2c^2$$

4. 4-wave vector:
$$K^\mu = (\omega/c, \mathbf{k})$$
where ω is the frequency and $\mathbf{k}$ is the usual wave vector.

   - For a plane wave, the phase $K_\mu x^\mu = \mathbf{k} \cdot \mathbf{x} - \omega t$ is invariant.
   - Doppler effect along line of motion
   $$\frac{\omega}{\omega_0} = \sqrt{\frac{1-\beta}{1+\beta}}$$
   where ω_0 is the source frequency and βc is the speed, and the source is receding from the observer.
   - Doppler effect off line of motion
   $$\frac{\omega}{\omega_0} = \frac{1}{\gamma(1 - \frac{u}{v_p}\cos\theta)}$$
   where u is the speed of the source, $v_p = \omega/k$ is the phase velocity in the observer's frame, and θ is the angle between the source's velocity and direction to the observer.

5. 4-current density:
$$J^\mu = (c\rho, \mathbf{j})$$
where ρ is charge density and $\mathbf{j}$ is the 3-current.
$$J_\mu J^\mu = -c^2\rho_0^2$$
where ρ_0 is the proper (rest frame) charge density. Current conservation takes the form
$$\partial_\mu J^\mu = 0$$

6. 4-acceleration:
$$A^\mu = \frac{dU^\mu}{d\tau} = \gamma(\dot{\gamma}c, \dot{\gamma}\mathbf{u} + \gamma\mathbf{a})$$
where $\dot{\gamma} = d\gamma/dt$ and $\mathbf{a} = d\mathbf{u}/dt$ is the usual 3-acceleration.
$$A_\mu A^\mu = a_0^2$$
where a_0 is the proper acceleration (acceleration in the body's rest frame).

7. 4-force:
$$F^\mu = \frac{dP^\mu}{d\tau} = \left(\frac{\gamma}{c}\frac{dE}{dt}, \gamma\mathbf{f}\right)$$
where $\mathbf{f} = d\mathbf{p}/dt$ is the usual 3-force.

2.5 4-momentum conservation

Students should know how to exploit 4-momentum conservation in elastic collisions to obtain results in this section.

2.5.1 Particle decay

Two-body decay of parent with mass M into daughters with masses m_1 and m_2, in parent rest frame:
$$E_1 = \frac{M^2c^4 + m_1^2c^4 - m_2^2c^4}{2Mc^2}$$
For general decay in flight, can calculate the mass of the parent if one knows the 4-momenta of all the daughters:
$$\begin{aligned} P^\mu &= \sum_i P_i^\mu \\ (P^\mu)^2 &= \left(\sum_i P_i^\mu\right)^2 \end{aligned}$$

2.5.2 Particle annihilation and formation

Particle formation in a fixed-target experiment: minimum incident particle energy E_{in} required to produce a final state with particle masses m_i:

$$E_{in} = \frac{1}{2Mc^2}\left[\left(\sum_i m_i\right)^2 - M^2 - m^2\right]c^4$$

where m is the mass of the incoming particle and M is the mass of the target.

In a symmetric collider experiment with two beams of equal but opposite momenta, the beam energy required is

$$E_{in} = \frac{1}{2}\left(\sum_i m_i c^2\right)$$

2.5.3 Compton scattering

Scattering of a photon off a particle (usually an electron) initially at rest.

$$\lambda' - \lambda = \frac{h}{mc}(1 - \cos\theta)$$

where λ and λ' are the initial and final photon wavelengths, m is the target mass, and $\cos\theta$ is the scattering angle.

2.6 Mechanics 4-tensors

2.6.1 Angular momentum

The tensor form of 4-angular momentum is

$$L^{\mu\nu} = x^\mu P^\nu - P^\mu x^\nu$$

which is conserved for a free particle, and for a collection of particles with internal elastic interactions. For such systems, consideration of the time-space part of $L^\mu\nu$ leads to the energy-weighted centroid velocity

$$\mathbf{v}_C = \frac{\sum_i E_i \mathbf{x}_i}{\sum_i E_i}$$

being the same as the velocity of the center-of-momentum frame.

2.6.2 Stress-energy tensor

Interpretation of elements:

T^{00}	energy density
T^{0j}	energy flux
T^{j0}	momentum density
T^{jj}	pressure
$T^{ij}, i \neq j$	shear stress

The tensor can be made symmetric by adding a 4-divergence term with no physical consequence, so we generally consider the tensor to be symmetric.

Conservation of momentum: $\partial_\nu T^{\mu\nu} = 0$

Dust

In rest frame, dust particles are all at rest.

$$T^{\mu\nu} = \begin{pmatrix} \rho_0 c^2 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}$$

where ρ_0 is the density. In an arbitrary frame, can write

$$T^{\mu\nu} = \rho_0 U^\mu U^\nu$$

where U^μ is the collective 4-velocity of the dust particles

Ideal fluid

In rest frame,

$$T^{\mu\nu} = \begin{pmatrix} \rho_0 c^2 & & & \\ & p & & \\ & & p & \\ & & & p \end{pmatrix}$$

where ρ_0 is the density and p is the (isotropic) pressure. In an arbitrary frame,

$$T^{\mu\nu} = (\rho_0 + \frac{p}{c^2})U^\mu U^\nu + p g^{\mu\nu}$$

2.6.3 Angular momentum density

This is analogous to angular momentum as defined for a flow of momentum:

$$M^{\lambda\mu\nu} = T^{\lambda\mu} x^\nu - T^{\lambda\nu} x^\mu$$

Conservation of angular momentum: $\partial_\lambda M^{\lambda\mu\nu} = 0$

Chapter 3

Forces and fields

An *instantaneous rest frame* is one of a set which move at the same velocity as a body at a given time. In each instantaneous rest frame, the body is at rest but is accelerating.

3.1 Force

Recall the 4-force:

$$F^\mu = \frac{dP^\mu}{d\tau} = \left(\frac{\gamma}{c}\frac{dE}{dt}, \gamma\mathbf{f}\right)$$

where $\mathbf{f} = d\mathbf{p}/dt$ is the usual 3-force.

Invariant for a body with 4-velocity U^μ and 4-force F^μ on it:

$$U^\mu F_\mu = -\gamma^2\frac{dE}{dt} + \gamma^2\mathbf{u}\cdot\mathbf{f}$$

which in the body's instantaneous rest frame is

$$U^\mu F_\mu = -\frac{dm}{d\tau}c^2$$

A *pure force* is one which does not change the mass of the system, but only kinetic energies. Then $U_\mu F_\mu = 0$ and

$$\frac{dE}{dt} = \mathbf{u}\cdot\mathbf{f}$$

3.1.1 Transformation of force

Consider a particle with 4-velocity U^μ in S, subject to 4-force F^μ. Frame S' travels with velocity $\mathbf{v}$ relative to S. The parallel component of the 3-force transforms as

$$f'_\parallel = \frac{f_\parallel - \frac{\beta_v}{c}\frac{dE}{dt}}{1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}}$$

which for a pure force simplifies to

$$f'_{\parallel} = \frac{f_{\parallel} - \frac{\mathbf{v}(\mathbf{u}\cdot\mathbf{f})}{c^2}}{1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}}$$

If $\mathbf{f}$, $\mathbf{u}$, and $\mathbf{v}$ are all aligned, this simplifies to $f'_{\parallel} = f_{\parallel}$.

The transverse part of the 3-force transforms as

$$\mathbf{f}'_{\perp} = \frac{\mathbf{f}_{\perp}}{\gamma_v(1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2})}$$

3.1.2 Motion with a given force

Consider a pure force applied to a particle with 3-velocity $\mathbf{u}$.

$$\begin{aligned} \mathbf{f} &= \gamma_u m\mathbf{a} + \frac{1}{c^2}(\mathbf{f}\cdot\mathbf{u})\mathbf{u} \\ f_{\parallel} &= \gamma^3 m a_{\parallel} \\ \mathbf{f}_{\perp} &= \gamma m \mathbf{a}_{\perp} \end{aligned}$$

3.1.3 Uniform circular motion

Larmor radius for a particle in a uniform magnetic field B

$$r = p/qB$$

where q is the charge of the particle, and p the momentum of the particle transverse to the field.

3.1.4 Linear motion

Arbitrary proper acceleration a_0 as a function of time:

$$\sinh\eta = \frac{1}{c}\int a_0(t)dt + C$$

If a_0 is a constant in any instantaneous rest frame,

$$\beta(t) = \frac{(a_0 t/c)}{[1 + (a_0 t/c)^2]^{1/2}}$$

which approaches 1 as $t \to \infty$. The distance travelled $x(t)$ is given by the equation

$$(x - b)^2 - c^2t^2 = (c^2/a_0)^2$$

where b is a constant. This is sometimes called "hyperbolic motion".

3.2 Electromagnetic fields

Lorentz force: $\mathbf{f} = q(\mathbf{E} + \mathbf{v} \wedge \mathbf{B})$

Maxwell equations:

$$\begin{aligned} \nabla \cdot \mathbf{E} &= \frac{\rho}{\varepsilon_0} \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \wedge \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0 \\ \nabla \wedge \mathbf{B} - \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} &= \mu_0 \mathbf{j} \end{aligned}$$

Transformation:

$$\begin{aligned} \mathbf{E}'_{\parallel} &= \mathbf{E}_{\parallel} \\ \mathbf{E}'_{\perp} &= \gamma(\mathbf{E}_{\perp} + \mathbf{v} \wedge \mathbf{B}) \\ \mathbf{B}'_{\parallel} &= \mathbf{B}_{\parallel} \\ \mathbf{B}'_{\perp} &= \gamma(\mathbf{B}_{\perp} - \frac{\mathbf{v} \wedge \mathbf{E}}{c^2}) \end{aligned}$$

3.2.1 Field tensor

Maxwell field tensor:

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{pmatrix}$$

"Dual" field tensor:

$$\tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\kappa\lambda}F_{\kappa\lambda} = \begin{pmatrix} 0 & B_x & B_y & B_z \\ -B_x & 0 & -E_z/c & E_y/c \\ -B_y & E_z/c & 0 & -E_x/c \\ -B_z & -E_y/c & E_x/c & 0 \end{pmatrix}$$

where $\epsilon^{\mu\nu\kappa\lambda} = 1$ for even permutations, -1 for odd permutations, and 0 for other combinations of indices.

Lorentz force: $F^{\mu} = qF^{\mu\nu}U_{\nu}$

Maxwell equations:

$$\begin{aligned} \partial_{\nu}F^{\mu\nu} &= \mu_0 J^{\mu} \\ \partial_{\nu}\tilde{F}^{\mu\nu} &= 0 \end{aligned}$$

Invariants:

$$\begin{aligned} F^{\mu\nu}F_{\mu\nu} &= 2(B^2 - \frac{E^2}{c^2}) \\ F^{\mu\nu}\tilde{F}_{\mu\nu} &= -\frac{4\mathbf{E} \cdot \mathbf{B}}{c} \end{aligned}$$

3.2.2 4-potential

The 4-potential

$$A^\mu = (\phi/c, \mathbf{A})$$

leads to fields

$$\begin{aligned} \mathbf{B} &= \nabla \wedge \mathbf{A} \\ \mathbf{E} &= -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \end{aligned}$$

The field tensor can then be written

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$$

and we have a wave equation

$$\partial_\mu \partial^\mu A^\nu = -\mu_0 J^\nu$$

A gauge transformation

$$A^\mu \to A^\mu + \partial^\mu \chi$$

with χ a scalar function of x^μ, leaves the fields unchanged.

Common gauge conditions:

- Lorenz: $\partial_\mu A^\mu = 0$
- Coulomb: $\nabla \cdot \mathbf{A} = 0$
  (often combined with ϕ independent of time)

3.2.3 Energy and momentum in vacuum

The energy density of the field is

$$u = \frac{1}{2\mu_0}\left(B^2 + \frac{E^2}{c^2}\right)$$

Poynting's vector

$$\mathbf{N} = \frac{1}{\mu_0}\mathbf{E} \wedge \mathbf{B}$$

gives the energy flow carried by the field.

Continuity:

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{N} = -\mathbf{j} \cdot \mathbf{E}$$

3.2.4 Plane waves in vacuum

Complex form of plane waves:

$$\begin{aligned} \mathbf{E} &= \mathbf{E}_0 e^{ik^\mu x_\mu} \\ \mathbf{B} &= \mathbf{B}_0 e^{ik^\mu x_\mu} \end{aligned}$$

The real parts are the physical fields. $\mathbf{E}$, $\mathbf{B}$, and the wave direction $\mathbf{k}$ are mutually orthogonal. Also, $E_0 = cB_0$ as long as $\omega = kc$ (true in vacuum).

Solutions to the vacuum wave equation

$$\partial_\mu \partial^\mu A^\nu = 0$$

take the form

$$A^\mu = a\epsilon^\mu e^{ik_\nu x^\nu}$$

with a a constant and ϵ_μ the polarization vector, which can always be chosen with $\epsilon^0 = 0$ and $k_i \epsilon^i = 0$, *i.e.*, polarization transverse to $\mathbf{k}$. The fields are then

$$\begin{aligned} \mathbf{B} &= i\mathbf{k} \wedge \mathbf{A} \\ \mathbf{E} &= -i\frac{\omega}{c^2}\mathbf{A} \end{aligned}$$

3.2.5 Stress-energy tensor

The stress-energy tensor is also known as the "energy-momentum tensor".

$$T^{\mu\nu} = \frac{1}{\mu_0}\left[-\frac{1}{4}(F_{\alpha\beta}F^{\alpha\beta})g^{\mu\nu} - F^\mu{}_\gamma F^{\gamma\nu}\right]$$

Components:

$$T^{\mu\nu} = \begin{pmatrix} \frac{1}{2\mu_0}(B^2 + \frac{E^2}{c^2}) & N^1/c & N^2/c & N^3/c \\ N^1/c & \sigma^{11} & \sigma^{12} & \sigma^{13} \\ N^2/c & \sigma^{21} & \sigma^{22} & \sigma^{23} \\ N^3/c & \sigma^{31} & \sigma^{32} & \sigma^{33} \end{pmatrix}$$

with

$$\sigma^{ij} \equiv \frac{1}{\mu_0}\left[\frac{1}{2}\delta^{ij}(B^2 + \frac{E^2}{c^2}) - (B^i B^j + \frac{E^i E^j}{c^2})\right]$$

Parallel-plate capacitor with area A, with gap along x direction:

$$T^{\mu\nu} = \frac{\epsilon_0 E^2}{2}\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

Long straight solenoid along x:

$$T^{\mu\nu} = \frac{\epsilon_0 c^2 B^2}{2} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

Plane wave travelling in x direction, polarized along y direction

$$\begin{aligned} \mathbf{E} &= (0, E, 0)\cos(\omega t - kx) \\ \mathbf{B} &= (0, 0, B)\cos(\omega t - kx) \end{aligned}$$

leads to a stress-energy tensor

$$T^{\mu\nu} = \frac{E^2}{\mu_0 c^2}\cos^2(\omega t - kx) \begin{pmatrix} 1 & 1 & & \\ 1 & 1 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}$$

More general form:

$$T^{\mu\nu} = \frac{E^2 \cos^2(k_\lambda x^\lambda)}{\mu_0 \omega^2} k^\mu k^\nu$$

Chapter 4

Radiation

4.1 Uniformly moving charge

Field in S' in which charge q is at rest:

$$\begin{aligned} \mathbf{E}' &= \frac{q}{4\pi\varepsilon_0 r'^3}\mathbf{r}' \\ \mathbf{B}' &= 0 \end{aligned}$$

Field in a frame S in standard configuration:

$$\begin{aligned} \mathbf{E} &= \frac{\gamma q}{4\pi\varepsilon_0[\gamma^2(x-vt)^2+y^2+z^2]^{3/2}}(x-vt,y,z) \\ \mathbf{B}_{\parallel} &= 0 \\ \mathbf{B}_{\perp} &= \frac{\mathbf{v}\wedge\mathbf{E}}{c^2} \end{aligned}$$

With vector potentials:

$$\begin{aligned} \phi' &= \frac{q}{4\pi\varepsilon_0 r'} \\ \mathbf{A}' &= 0 \\ \phi &= \frac{\gamma q}{4\pi\varepsilon_0[\gamma^2(x-vt)^2+y^2+z^2]^{1/2}} \\ A_x &= v\phi \\ A_y = A_z &= 0 \end{aligned}$$

4.2 Retarded potentials

Green function method on wave equation:

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) G(t,\mathbf{x};t_s,\mathbf{x}_s) = -\delta(\mathbf{x}-\mathbf{x}_s)\delta(t-t_s)$$

where $(t_s, \mathbf{x}_s)$ is the source point and $(t, \mathbf{x})$ is the observation point.

Fourier transform of the equation and solution:

$$\begin{aligned}\left(\nabla^2 + (\frac{\omega}{c})^2\right) G_\omega(\mathbf{x};\mathbf{x}_s) &= -\delta(\mathbf{x}-\mathbf{x}_s) \\ G(t,\mathbf{x};t_s,\mathbf{x}_s) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega e^{-i\omega(t-t_s)} G_\omega(\mathbf{x};\mathbf{x}_s)\end{aligned}$$

Spherical symmetric solution:

$$\begin{aligned}G_\omega(\mathbf{x};\mathbf{x}_s) &= G_\omega(|\mathbf{x}-\mathbf{x}_s|) = G_\omega(r) \\ G_\omega^{\pm}(r) &= \frac{e^{\pm i\omega r/c}}{4\pi r}\end{aligned}$$

where $(+)$ is the "retarded" Green function (behavior after the source event), and $(-)$ is the "advanced" Green function (behavior before the source event). Then

$$G(t,\mathbf{x};t_s,\mathbf{x}_s) = \frac{1}{4\pi|\mathbf{x}-\mathbf{x}_s|}\delta\left(t_s - (t \mp \frac{|\mathbf{x}-\mathbf{x}_s|}{c})\right)$$

Applied to wave equation for the 4-potential:

$$A^\mu(t,\mathbf{x}) = \frac{1}{4\pi\varepsilon_0 c^2}\int d\mathbf{x}_s \frac{J^\mu(t - r_{\rm sf}/c, \mathbf{x}_s)}{r_{\rm sf}}$$

where $r_{\rm sf} \equiv |\mathbf{x}-\mathbf{x}_s|$.

Liénard-Wiechert potential for an arbitrarily moving charge:

$$A^\mu = \frac{q}{4\pi\varepsilon_0}\frac{U^\mu/c}{(-R^\nu_{\rm sf}U_\nu)}$$

where $R^\nu_{\rm sf} = R^\nu - R^\nu_s$.

4.3 Accelerated charge

Field at $\mathbf{r}_f$ of an accelerated charge

$$\begin{aligned}\mathbf{E} &= \frac{q}{4\pi\varepsilon_0\kappa^2}\left(\frac{\mathbf{n}-\mathbf{v}/c}{\gamma^2 r^2} + \frac{\mathbf{n}\wedge[(\mathbf{n}-\mathbf{v}/c)\wedge\mathbf{a}]}{c^2 r}\right) \\ \mathbf{B} &= \mathbf{n}\wedge\mathbf{E}/c \\ \mathbf{n} &= \mathbf{r}/r \\ \kappa &= 1 - \frac{v_r}{c} = 1 - \frac{\mathbf{n}\cdot\mathbf{v}}{c} \\ \mathbf{r} &= \mathbf{r}_f - \mathbf{r}_s\end{aligned}$$

where $\mathbf{r}_s$ is the position of the source at the retarded time $t_s = t - r/c$. $\mathbf{v}$ and $\mathbf{a}$ are the velocity and acceleration at the source event $(t_s, \mathbf{r}_s)$.

The second term in the above fields, which varies as $1/r$, is the "radiative" field.

4.3.1 Slowly oscillating dipole

For a dipole consisting of a stationary charge $-q$ and oscillating charge q with amplitude d_0 and frequency ω, the far field magnitudes are

$$E_{rad} = cB = \frac{q\omega^2 d_0}{4\pi\varepsilon_0 c^2}\frac{\sin\theta}{r}\sin(kr - \omega t)$$

In the case of a centre-fed half-wave dipole antenna with length $L = \lambda/2$,

$$E_{rad} = \frac{I_0}{2\pi\varepsilon_0 c}\frac{\sin\theta}{r}\sin(kr - \omega t)$$

4.3.2 Radiated power

For an accelerated charge at non-relativistic velocities, the radiative field is

$$\mathbf{E}_{rad} \approx \frac{q}{4\pi\varepsilon_0 c^2}\left(\frac{\hat{\mathbf{n}} \wedge (\hat{\mathbf{n}} \wedge \mathbf{a})}{r}\right)$$

The Poynting vector is $\mathbf{N} = \varepsilon_0 c E_{rad}^2 \hat{\mathbf{n}}$, which gives a differential power distribution

$$\frac{dP}{d\Omega} = Nr^2 = \frac{q^2}{4\pi\varepsilon_0 c^3}\frac{1}{4\pi}a^2\sin^2\theta$$

where θ is the angle between the $\mathbf{a}$ and the radial direction $\hat{\mathbf{n}}$. The total power gives *Larmor's formula*,

$$P_L = \frac{2}{3}\frac{q^2}{4\pi\varepsilon_0}\frac{a^2}{c^3}$$

Relativistic case (from Liénard):

$$P = \frac{2}{3}\frac{q^2}{4\pi\varepsilon_0 c^3}\gamma^6\left[a^2 - \frac{(\mathbf{v} \wedge \mathbf{a})^2}{c^2}\right]$$

Circular accelerator

Power loss from a circular accelerator of radius ρ:

$$P_L = \frac{2}{3}\frac{q^2 c}{4\pi\varepsilon_0\rho^2}\beta^4\gamma^4$$

Energy loss per revolution is then

$$\Delta E = \frac{P}{f} = \frac{4\pi}{3}\frac{\alpha\hbar c}{\rho}\beta^3\left(\frac{E}{mc^2}\right)^4$$

where α is the electromagnetic fine structure constant, E is the beam energy, and m is the mass of the beam particle.

4.4 Additional material

The following material is to be shown on slides.

Derivation of fields for accelerated charge (Section 4.3: start from the Liénard-Wiechert potential. Define $k \equiv -q/4\pi\varepsilon_0 c$ and $s \equiv R^\kappa U_\kappa$.

$$\begin{aligned} F^{\mu\nu} &= \partial^\mu A^\nu - \partial^\nu A^\mu \\ \partial_\mu A^\nu &= k\frac{s\partial_\mu U^\nu - U^\nu \partial_\mu s}{s^2} \end{aligned}$$

Define the acceleration of the source event in terms of the source's proper time,

$$a^\nu = \frac{dU^\nu}{d\tau}$$

which allows us to evaluate the derivative of the source's 4-velocity:

$$\partial_\mu U^\nu = \frac{dU^\nu}{d\tau}\frac{\partial\tau}{\partial x^\mu} = a^\nu \partial_\mu \tau$$

More generally,

$$\partial_\mu = (\partial_\mu \tau)\frac{d}{d\tau}$$

Since it is always true that $R^\nu R_\nu = 0$, then R^ν is always orthogonal to its gradient. Therefore

$$\begin{aligned} 0 &= R_\nu \partial_\mu R^\nu \\ &= R_\nu \partial_\mu (R_f^\nu - R_s^\nu) \\ &= R_\nu (\delta_\mu^\nu - (\partial_\mu \tau)\frac{dR_s^\nu}{d\tau}) \\ &= R_\nu (\delta_\mu^\nu - U^\nu \partial_\mu \tau) \\ &= R_\mu - R_\nu U^\nu (\partial_\mu \tau) \\ \partial_\mu \tau &= R_\mu / s \\ \partial_\mu U^\nu = a^\nu \partial_\mu \tau &= a^\nu R_\mu / s \end{aligned}$$

For the second term,

$$\begin{aligned} \partial_\mu s &= \partial_\mu (R_\nu U^\nu) \\ &= (\partial_\mu R^\nu) U_\nu + R_\nu (\partial_\mu U^\nu) \\ &= (\delta_\mu^\nu - U^\nu \partial_\mu \tau) U_\nu + R_\nu (a^\nu \partial_\mu \tau) \\ &= U_\mu - U^\nu U_\nu \frac{R_\mu}{s} + R_\nu a^\nu \frac{R_\mu}{s} \end{aligned}$$

Combining,

$$\partial^\mu A^\nu = -\frac{k}{s^2} U^\nu U^\mu - \frac{kc^2}{s^3} R^\mu V^\nu$$

where

$$V^\nu \equiv U^\nu - \frac{1}{c^2}[s a^\nu - U^\nu (R_\kappa a^\kappa)]$$

Then

$$F^{\mu\nu} = \frac{qc}{4\pi\varepsilon_0}\frac{R^\mu V^\nu - R^\nu V^\mu}{(R_\kappa U^\kappa)^3}$$

Chapter 5

Groups

A *group* is a set G and an operator $\cdot$ for which

1. Closure: for all $a, b \in G$,
$$a \cdot b \in G$$
2. Associativity: for all $a, b, c \in G$,
$$(a \cdot b) \cdot c = a \cdot (b \cdot c)$$
3. Identity element: there exists an element e in G such that $e \cdot a = a \cdot e = a$ for any $a \in G$
4. Inverse: for every a in G, there exists an element a^{-1} such that $a^{-1} \cdot a = a \cdot a^{-1} = e$

A *subgroup* is a subset of a group for which the group axioms hold.

5.1 Permutation group

Consider N slots, labelled from 1 to N. The string $p_1 p_2 \cdots p_N$, with p_i an integer from 1 to N, denotes a permutation by which the original contents of slot p_1 have moved to slot 1, slot p_2 to slot 2, *etc.* The identity element $12 \cdots N$ is also called e.

Multiplication table for $N = 3$:

	e	231	312	213	132	321
e	e	231	312	213	132	321
231	231	312	e	321	213	132
312	312	e	231	132	321	213
213	213	132	321	e	231	312
132	132	321	213	312	e	231
321	321	213	132	231	312	e

The upper left quadrant forms a subgroup of cyclic permutations.

The elements e, 231, and 213 can be taken as *generators* of the group in that all elements can be written as products of these three.

5.2 Representation

A *homomorphism* from a group G to another group H is a mapping which preserves the multiplication table, *i.e.*, for elements $g_1, g_2, g_3 \in G$ and corresponding elements $h_1, h_2, h_3 \in H$, then $g_1 g_2 = g_3$ implies $h_1 h_2 = h_3$.

An *isomorphism* is a homomorphism in which the mapping of elements is one-to-one.

A *group representation* is homomorphism between two groups, usually an abstract group and another group composed of more concrete elements such as a number or matrix. If the mapping is isomorphic, the representation is *faithful*.

The *representation space* is the space on which the elements of the group representation acts.

Physicists often use abstract and representation groups interchangeably. A common pattern is that the representation takes the form of a group of matrices, and the representation space a set of column vectors.

5.2.1 Reducibility

An *invariant subspace* U of a representation space V is one in which the action of any element D of the representation is closed within U, *i.e.*, for any vector $u \in U$ and any D in the representation, $Du \in U$.

A representation is *reducible* if it has an invariant subspace. It is *irreducible* if does not.

Two representations are *equivalent* if they are related through a similarity transformation, *i.e.*, they differ only in the basis chosen.

In matrix form, a reducible representation is equivalent to a representation in which the matrices have the same block-diagonal form

$$\begin{pmatrix} D_1 & 0 & \cdots \\ 0 & D_2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

The reducible representation is then said to be the *direct sum* of sub-representations

$$D_1 \oplus D_2 \oplus \cdots$$

5.3 Rotations

The group of rotations are an example of a *Lie group*, which, practically speaking, is an infinite group whose elements can be parameterized smoothly.

The group $O(3)$ consists of all linear transformations which preserve the length of a 3D vector. This includes rotations as well as space inversion. Elements can be (faithfully) represented by 3×3 orthogonal matrices with determinant ± 1.

The "special subgroup" $SO(3)$ consists only of those rotations smoothly connected to the identity, *i.e.*, excluding space inversion. Matrix representations of elements of $SO(3)$ have determinant $+1$.

Elements of $SO(3)$ can be written in the form

$$R(\hat{\mathbf{n}}, \theta) = e^{-i\theta n^k J_k}$$

where $\hat{\mathbf{n}}$ is the axis of rotation, θ the angle, and J_k the *generators* of the group.

$$(J_1)^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad (J_2)^\mu{}_\nu = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad (J_3)^\mu{}_\nu = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(Note that these generators are defined to give *active* rotations using the exponentiation rule above, *i.e.*, rotating the object rather than the coordinate system.) The generators are related through a *Lie algebra*

$$[J_j, J_k] = i \sum_m \epsilon_{jkm} J_m$$

A *Casimir operator* commutes with all the generators. For $SO(3)$, this is $J^2 = J_1^2 + J_2^2 + J_3^2$.

Using the raising and lowering matrices $J_\pm = J_1 \pm iJ_2$, we find that J_3 is represented by a diagonal matrix. We denote its eigenvectors $|jm\rangle$ where

$$\begin{aligned} J^2|jm\rangle &= j(j+1)|jm\rangle \\ J_3|jm\rangle &= m|jm\rangle \\ J_\pm|jm\rangle &= \sqrt{j(j+1) - m(m \pm 1)}|j, m \pm 1\rangle \\ j &= 0, \frac{1}{2}, 1, \frac{3}{2}, 2, ... \\ m &= -j, -j+1, ..., j-1, j \end{aligned}$$

The rotation matrices generated by the J_k for a given value of j form an irreducible representation of $SO(3)$. The $|jm\rangle$ form the basis of the corresponding representation space. Only $j = 1$ is a faithful representation of $SO(3)$.

5.3.1 SU(2)

(Non-syllabus material)

The $j = 1/2$ case is a double-valued representation which requires a rotation through 4π to get back to the identity. The generators $J_k = \sigma_k/2$, where σ_k are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The rotation matrices thus generated are 2×2 unitary matrices with determinant $+1$, representing the group $SU(2)$. This group is a simply-connected double cover of $SO(3)$.

5.4 Lorentz group

The full Lorentz group is $SO(3,1)$.

There are four divisions of the full Lorentz group given by $\det\Lambda = \pm 1$, and ${\Lambda^0}_0 \geq 1$ or ≤ -1. The division which contains the identity element is $\det\Lambda = 1$ and ${\Lambda^0}_0 \geq 1$ forms the *orthochronous* (preserving the normal time direction) subgroup $SO(3,1)^{\uparrow}_{+}$.

5.4.1 Generators of Lorentz boosts

A finite (active) boost along the direction η^j is

$${L^\mu}_\nu = e^{-i\eta^j K_j}$$

with the generators

$$({K_1})^\mu{}_\nu = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad ({K_2})^\mu{}_\nu = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad ({K_3})^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}$$

Combined with rotations (adding an empty 0-column and 0-row to the J_k matrices above),

$${\Lambda^\mu}_\nu = e^{-i\theta^j J_j \pm i\eta^j K_j}$$

where the two signs indicate the chirality of the combined boost and rotation. The two chiralities are two inequivalent representations of the Lorentz group.

Commutators

$$\begin{aligned} [J_a, J_b] &= i\sum_c \epsilon_{abc} J_c \\ [K_a, J_b] &= i\sum_c \epsilon_{abc} K_c \\ [K_a, K_b] &= -i\sum_c \epsilon_{abc} J_c \end{aligned}$$

Complexification (non-syllabus):

$$\begin{aligned} M_a &\equiv (J_a + iK_a)/2 \\ N_a &\equiv (J_a - iK_a)/2 \end{aligned}$$

which leads to commutators

$$\begin{aligned} [M_a, M_b] &= i\sum_c \epsilon_{abc} M_c \\ [N_a, N_b] &= i\sum_c \epsilon_{abc} N_c \\ [M_a, N_b] &= 0 \end{aligned}$$

which indicates a direct product structure $SU(2) \times SU(2)$. The Casimir operators are M^2 and N^2, with eigenvalues $m(m+1)$ and $n(n+1)$, with m and n non-negative integers or half-integers. Representations can be labelled by (m, n). Space inversion turns an (m, n) representation into (n, m).

Physical representations with parity symmetry come in form (n, n) or $(m, n) \oplus (n, m)$ with $m \neq n$. Examples include scalars $(0, 0)$ and 4-vectors $(\frac{1}{2}, \frac{1}{2})$. The electromagnetic field tensor comes from the $(0, 1) \oplus (1, 0)$ representation, and Dirac spinors representing fermions are from $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$.

5.4.2 Poincaré group

(Non-syllabus material)

Casimir operators are $P_\mu P^\mu$ and $W_\mu W^\mu$, where W^μ is the Pauli-Lubanski vector:

$$W^\mu \equiv \frac{1}{2}\epsilon^{\mu\nu\kappa\lambda} P_\nu J_{\kappa\lambda}$$

where P_ν is the 4-momentum and $J_{\kappa\lambda}$ is the angular momentum 4-tensor. The W^ν is zero for orbital angular momentum, but non-zero for spin (half-integer j) angular momenta.

The unitary representations of the Poincaré group are as follows:

1. $P^\mu P_\mu = -m^2$ where m is a real number. These are finite-mass particles, with spin values $s = 0, \frac{1}{2}, 1, \frac{3}{2}$, *etc.* The eigenvalue of $W^\mu W_\mu$ is $m^2 s(s+1)$. States are labelled by the spin z-component s_3 and the (continuous) 3-momentum $\mathbf{p}$.

2. $P^\mu P_\mu = 0$ and $W^\mu W_\mu = 0$. These are massless particles with two helicity states. Physically, these are photons with helicity ± 1, gravitons with ± 2, and (in the massless approximation) neutrinos with $\pm\frac{1}{2}$.

3. $P^\mu P_\mu = 0$ but $W^\mu W_\mu > 0$. These are massless particles but with continuous spin. Particles of this type have not been found.

4. $P^\mu P_\mu > 0$. Particles in this category would be tachyons. Such particles have not been found per se, but feature in virtual contributions to scattering amplitudes.

5.5 Gauge invariance

Quantum physics is invariant with respect to transformations

$$\psi(x) \to \psi'(x) = e^{iq\alpha}\psi(x)$$

If α is a constant, this is a *global gauge transformation*. If α is also a function of spacetime, the transformation is a *local gauge transformation*. Such transformations are elements of the group $U(1)$.

Dirac equation with a 4-vector potential field (non-syllabus for B2):

$$(i\gamma^\mu\partial_\mu + q\gamma^\mu A_\mu - m)\psi = 0$$

where γ^μ are Dirac matrices. The equation by itself is not gauge invariant with respect to A_μ, but invariance can be restored by also transforming ψ. Similar arguments can be made for further unitary transformations $SU(2)$ and $SU(3)$, resulting in further "gauge fields".