Notes for Electricity and Magnetism 3, Phys 442, University of Waterloo

J. D. D. Martin

`last commit hash: cf238f7        compiled: 2022-02-24, 09:58 EST`

What's new?

2022-02-23:

- added: Electromagnetic unit systems demystified

2022-02-05:

- added: Cyclotron motion

2022-02-03:

- corrected some typos in: The Minkowski force

I update these notes frequently. For the most recent version, please see:

`https://www.dropbox.com/s/yk989rgahy2p4no/phys442_notes_generated.pdf?dl=0`

If you are viewing this file through the dropbox link, you may download it using the icon. You do not need to sign up for Dropbox.

Contents

1 Topics to be covered

We will study the following topics in order:[1]

(1) (a) Maxwell's equations
  (b) An example of magnetostatics:
    (i) The magnetic field of a current loop along its symmetry axis
    (ii) Helmholtz coils
  (c) Maxwell's equations in matter

(2) (a) Relativistic velocity addition
  (b) An experimental test of relativistic velocity addition
  (c) The relativity of simultaneity
  (d) Relativistic time dilation
  (e) Experimental observation of time dilation
  (f) Relativistic length contraction
  (g) Length contraction in electromagnetism
  (h) Lorentz transformations
  (i) Four-vectors and tensors
  (j) Proper time and the terrible twins
  (k) Momenergy from an invariant action
  (l) The Lorentz force law from an invariant action
  (m) The Minkowski force
  (n) Transformations of electric and magnetic fields between inertial frames
  (o) Fields due to a uniformly moving point charge
  (p) Maxwell's equations in a manifestly covariant form

(3) (a) Poynting's theorem
  (b) Momentum conservation and Maxwell's stress tensor
  (c) Wave equations
  (d) Electromagnetic plane waves
  (e) Energy and momentum transport in electromagnetic waves

[1] The corresponding sections of GITE4 are indicated in the Section titles in the main text; e.g., in the title for Section 5 we find "(GITE4 5.2.2)", indicating that the corresponding section of Griffiths' book is Section 5.2.2.

(4) (a) Solving Maxwell's equations in the Lorenz gauge
    (b) Introduction to radiation
    (c) Hertzian dipole
    (d) Intuitive treatment of point charge radiation

These descriptions will be refined slightly as the course proceeds. At this point they give an idea of the scope of the course. Further clarification will be provided on the alignment with the tests and quizzes.

2 Introduction

The primary reference for this course is Griffiths, *Introduction to electrodynamics*, 4th ed. [1], which will be abbreviated as GITE4 throughout these notes.

However, I will order the material differently from GITE4. More specifically, we will start by reviewing special relativity. Then, where possible, we will incorporate special relativity into our arguments as the course proceeds.

As an example of the benefits of studying special relativity first, consider determining the electric and magnetic fields due to a uniformly *moving* point charge. There are two different — but equally valid — approaches to determining these fields:

(1) use the general formulae for the fields due to moving point charges, or

(2) take the Coulomb field in the rest frame of the electron, where there is no magnetic field, and then use the rules for transforming fields from one inertial frame to another, to determine the magnetic and electric fields in the frame where the electron is moving; i.e., the frame that we are interested in.

The second approach, based on special relativity, is more efficient. It is also more elegant and the relevant formulae are easily remembered.

The "relativity first" approach to electrodynamics is taken by one of the classic books of physics: Landau and Lifshitz's *The classical theory of fields* [2], which is great to look at for inspiration and insight, but utterly unsuitable as a textbook.

In contrast, it is hard to imagine improving on the overall style, substance, and attention to pedagogy of GITE4.

I will mainly reproduce GITE4's arguments, but try to add material that is different or complementary, particularly relating to the reordered treatment of this course.

3 Some abbreviations, symbols, and housekeeping matters

symbol	meaning
GITE4	Griffiths's *Introduction to electrodynamics*, 4th ed.
L&L	Landau and Lifshitz (possibly followed by specification of a volume and section number)
DLMF	NIST Digital Library of Mathematical Functions
A&S	Abramowitz and Stegun's Handbook of Mathematical Functions
SR	special relativity
wrt	with respect to
op. cit.	in the work previously mentioned or quoted
cf.	compare
XX	placeholder for material to be corrected or added later
	link to a video that I have made (Youtube version)
	link to a video that I have made (Microsoft Streams version, no ads, UW only)
	video pages and (optionally) comments/corrections on presentation (incomplete/experimental)
$:=$	definition (symbol on LHS is being defined)
$=:$	definition (symbol on RHS is being defined)
$\equiv$	an equivalence
$\square$	end of proof or demonstration
$\otimes$	into page (vector)
$\odot$	out of page (vector)
$\stackrel{?}{=}$	an incorrect or dubious result

New terminology is introduced in these notes using **bold-face**. When brackets are put around a part of the terminology, it is to be understood that for brevity the bracketed part may be dropped in what follows, but is implicit. For example, after I introduce **(Lorentz) tensors**, I normally drop "Lorentz", since I do not talk about any other sort of tensor.

I try to follow the notation and conventions of GITE4, with some exceptions.[2]

In handwriting, I signify vector quantities with $\vec{}$, as in $\vec{E}$. I omit $\vec{}$ from ∇ as it only makes sense as "vector". Otherwise, here and in tests and quizzes I use bold face, as in $\mathbf{E}$.

There are a variety of electromagnetic "unit systems" in use. I will use the the same system as in GITE4: the **International System of Units**, abbreviated as "SI".[3] For further comments on differing unit systems see Section E.

These notes contain references to journal articles. Some are to acknowledge the sources that I have used and may not be particularly useful to look at — I have tried to explicitly indicate which ones

[2] Instead of $:=$, GITE4 uses the physics style $\equiv$ for definitions. But a definition is an asymmetric relation, so I side with the mathematicians (!) and use $:=$ and $=:$ for definitions and (occasionally) $\equiv$ for equivalences, not definitions.

[3] In an earlier version of these notes, I referred to the SI system using the equivalent but outdated terminology: "rationalized metre kilogram second ampere" (RMKSA) system.

might be useful. These references will also normally be links that take you directly to the "official" journal page for the articles. These pages will offer the option to purchase the article; however, if you are a UW student **it is not necessary to pay for these articles**, as they may be freely accessed using the library: `https://login.proxy.lib.uwaterloo.ca/login` (Please contact me if you have problems accessing an article freely.)

4 Maxwell's equations

From an extremely reductionist viewpoint our subject is that of the Lorentz force law and Maxwell's equations.

The origin of Maxwell's equations will be familiar from Phys 242 and 342, especially Coulomb's law and the Biot-Savart law. In the accompanying video, I give a brief overview.

5 The magnetic field of a current loop along its symmetry axis (GITE4 5.2.2)

Here we show an application of the Biot-Savart law in the form specialized for "loops of current":

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi}\, I \oint \frac{d\boldsymbol{\ell}' \times \hat{\boldsymbol{\mathscr{r}}}}{\mathscr{r}^2} \tag{1}$$

6 Helmholtz coils

Often one wants to make a (relatively) homogeneous field over a region of space; i.e., we have an extended sample, and we want all parts of the example to be exposed to the same magnetic field. One of the simplest ways to obtain good homogeneity is with so-called **Helmholtz coil** configuration. Looking at the details of this configuration is a nice example of the results of the previous section, where we derived the **B** field due to a circular current ring along its axis.

7 Maxwell's equations in matter

- Maxwell's equations are sometimes written as:

$$\nabla \cdot \mathbf{D} = \rho_f \tag{2}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{4}$$

$$\nabla \times \mathbf{H} = \mathbf{J}_f + \frac{\partial \mathbf{D}}{\partial t} \tag{5}$$

where

$$\mathbf{D} \coloneqq \epsilon_0 \mathbf{E} + \mathbf{P} \tag{6}$$

and

$$\mathbf{H} := \frac{\mathbf{B}}{\mu_0} - \mathbf{M} \tag{7}$$

which are known as the **auxiliary fields**. We will need to define $\mathbf{P}$ and $\mathbf{M}$. For now, note that they are related to the amount of polarization and magnetization of a material respectively. In vacuum, or if the material is not polarized or magnetized, then $\mathbf{P}$ and $\mathbf{M}$ are $\mathbf{0}$, and then Eq. 2 reduces to the normal $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$ and Eq. 5 reduces to $\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \partial \mathbf{E}/\partial t$. The subscript f on ρ_f and $\mathbf{J}_f$ signify that these are "free" charge density and currents respectively. Again, we will define what "free" means shortly.

Equations 2 to 5 are often referred to as the **macroscopic Maxwell's equations**. That might (wrongly) seem to imply that somehow the regular Maxwell's equations are not valid in matter. In fact, as we shall see, these equations are just a different way to write Maxwell's equations, that is *sometimes* more convenient when considering materials that can be polarized and/or magnetized.

- I will start by explaining how and why $\mathbf{P}$ and ρ_f are introduced. Recall that in electrostatics ($\nabla \times \mathbf{E} = \mathbf{0}$) we can represent $\mathbf{E}$ with a potential field:

$$\mathbf{E} = -\nabla V, \tag{8}$$

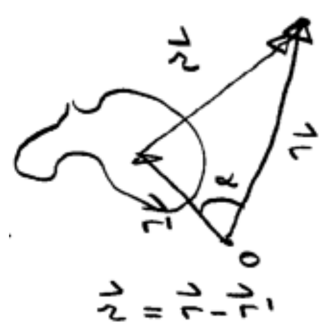
and that for a point charge at the origin

$$V(\mathbf{r}) = \frac{q}{4\pi\epsilon_0} \frac{1}{r}, \tag{9}$$

which may be generalized — using superposition — to determine the potential anywhere due an arbitrary charge distribution:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int d\tau' \, \frac{\rho(\mathbf{r}')}{\mathscr{r}}. \tag{10}$$

We will now manipulate this expression into a series in increasing powers of $1/r$. Start by rewriting $\mathscr{r}$. From geometry:



we see that:

$$\begin{aligned} \mathscr{r} &= \left[r^2 + r'^2 - 2rr' \cos\alpha \right]^{1/2} && (11) \\ &= r \left[1 + \left(\frac{r'}{r} \right)^2 - 2\frac{r'}{r} \cos\alpha \right]^{1/2}. && (12) \end{aligned}$$

Now we need a purely mathematical result, a **generating function** for the Legendre polynomials. Specifically, it can be shown (pg. 569 of Ref. [3]), that:

$$\frac{1}{\left[1-2xt+t^2\right]^{1/2}} = \sum_{\ell=0}^{\infty} P_\ell(x) t^\ell \tag{13}$$

where the P_ℓ are the **Legendre Polynomials**; i.e., $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = (3x^2-1)/2$, etc...

Notice the similarity between the generating function of Eq. 13 and our geometrical result of Eq. 12. Identifying $t \leftrightarrow r'/r$ and $x \leftrightarrow \cos\alpha$, we may rewrite Eq.'s 12 and 10 as:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{1}{r}\int d\tau'\ \rho(\mathbf{r}')\sum_{\ell=0}^{\infty} P_\ell(\cos\alpha)\left(\frac{r'}{r}\right)^\ell \tag{14}$$

$$= \frac{1}{4\pi\epsilon_0}\sum_{\ell=0}^{\infty}\frac{1}{r^{\ell+1}}\int d\tau'\ \rho(\mathbf{r}')P_\ell(\cos\alpha)r'^\ell \tag{15}$$

Note that the integrals now *only* depend on ℓ, the source charge distribution, and the angle of the observation point — not the distance to the observation point. The distance dependence between the observation point and source is in the $1/r^{\ell+1}$ prefactor, which only depends on ℓ. This hierarchy is great: at larger distances higher ℓ terms become less significant than those of lower ℓ.

Now let us consider the value of the integral for different values of ℓ. For $\ell = 0$:

$$\int d\tau'\rho(\mathbf{r}')P_\ell(\cos\theta)r'^\ell = \int d\tau'\rho(\mathbf{r}') \tag{16}$$

$$= Q_{\text{tot}} \tag{17}$$

i.e., just the total charge. So, as we might expect, the dominant contribution to the potential at large distances is the total charge, and the potential is independent of the angle of observation. For $\ell = 1$:

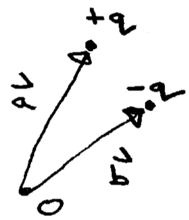
$$\int d\tau'\rho(\mathbf{r}')P_\ell(\cos\theta)r'^\ell = \int d\tau'\rho(\mathbf{r}')r'\cos\alpha. \tag{18}$$

The $r'\cos\alpha$ factor can be rewritten as $r'\,\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}' = \hat{\mathbf{r}}\cdot\mathbf{r}'$, so that:

$$= \hat{\mathbf{r}}\cdot\underbrace{\int d\tau'\rho(\mathbf{r}')\mathbf{r}'}_{=:\mathbf{p}} \tag{19}$$

Like the total charge Q_{tot}, the **electric dipole moment p** is *only* a property of the charge distribution, and does not depend on the observation location.

Why do we use the term *dipole*? Consider two charges of opposite sign at different locations (the two poles):



with the charge density distribution:

$$\rho(\mathbf{r}) = q\delta^3(\mathbf{r}-\mathbf{a}) - q\delta^3(\mathbf{r}-\mathbf{b}) \tag{20}$$

so that the electric dipole moment is:

$$\begin{align} \mathbf{p} &= \int d\tau\ \rho(\mathbf{r})\mathbf{r} \tag{21}\\ &= q\mathbf{a} - q\mathbf{b} \tag{22}\\ &= q(\mathbf{a}-\mathbf{b}). \tag{23} \end{align}$$

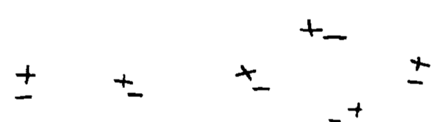
This result is probably familiar to you.

We could proceed with the $\ell = 2$ term of Eq. 15 (the "quadrupole moment") and so-on, but we do not need these terms right now. Summarizing:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{Q_{\text{tot}}}{r} + \frac{1}{4\pi\epsilon_0}\frac{1}{r^2}\,\hat{\mathbf{r}}\cdot\mathbf{p} + O\left(\frac{1}{r^3}\right) \tag{24}$$

Like we might expect, the total charge of the distribution is the most important as we move to larger distances. But since most matter is electrically neutral overall, the dipole moment term is often the leading non-zero term in the expansion of V.

The importance of the electric dipole moment suggests that we define a second field. Charge density is a scalar field that we integrate to get the total charge. We now define the **dielectric polarization**[4] $\mathbf{P}$ as the vector field that we may integrate to get the electric dipole moment of a charge distribution. You can view $\mathbf{P}$ as the average dipole moment per unit volume; i.e., imagine a gas of dipoles:



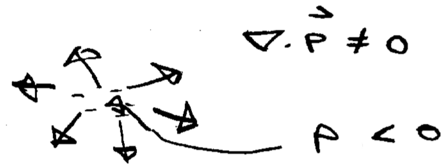
Then

$$\mathbf{P} = \langle\mathbf{p}\rangle\,\frac{N}{V} \tag{25}$$

where $\langle\mathbf{p}\rangle$ is the average moment of each dipole and N/V is the number of dipoles per unit volume (the "number density" of dipoles).

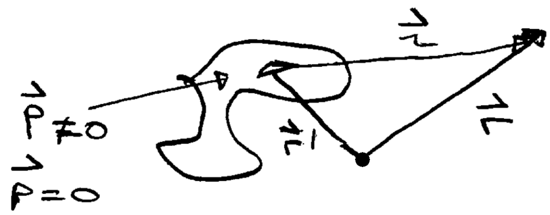
[4] Sometimes "dielectric" is dropped, and we just say "polarization". Depending on the context, that invites confusion with the (related but not the same) concept of the polarization of an electromagnetic wave; e.g., circular, linear, etc...

- it turns out that we can identify $\nabla \cdot \mathbf{P}$ as a *charge* density. A hand-wavy way to see that this identification is plausible is to imagine a $\mathbf{P}$ field that points out from a certain point:



  There is an excess of negative charge at the source of the divergence. In fact $\rho = -\nabla \cdot P$, which at least makes dimensional sense (check it!).

- I will now justify $\rho = -\nabla \cdot P$ using electrostatics. Let us imagine a body that has no total charge but has a dielectric polarization:



  Generalizing the result for the electrostatic potential due to a single dipole (the second term of Eq. 24) to a dipole moment distribution characterized by the $\mathbf{P}$ field:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int d\tau' \, \mathbf{P}(\mathbf{r}') \cdot \frac{\hat{\mathscr{r}}}{\mathscr{r}^2} \tag{26}$$

  Now let us rewrite the $\hat{\mathscr{r}}/\mathscr{r}^2$ factor in different form. Remember that:

$$\nabla \left(\frac{1}{r} \right) = -\frac{\hat{\mathbf{r}}}{r^2} \tag{27}$$

  and thus

$$\nabla' \frac{1}{\mathscr{r}} = \nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \tag{28}$$

$$= \frac{\hat{\mathscr{r}}}{\mathscr{r}^2} \tag{29}$$

  which allows Eq. 30 to be rewritten as:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int d\tau' \, \mathbf{P}(\mathbf{r}') \cdot \nabla' \left(\frac{1}{\mathscr{r}} \right). \tag{30}$$

  The form of the integrand suggests an "integration by parts" approach — generalized for vector calculus. More specifically, for an arb. scalar field f and vector field $\mathbf{v}$:

$$\nabla \cdot (f\mathbf{g}) = f\nabla \cdot \mathbf{g} + \nabla f \cdot \mathbf{g}. \tag{31}$$

Rearranging gives:

$$\nabla f \cdot \mathbf{g} = \nabla \cdot (f\mathbf{g}) - f\nabla \cdot \mathbf{g} \tag{32}$$

which is a form that appears in Eq. 30, and thus allows us to rewrite it as:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\int d\tau' \left[\nabla' \cdot \left(\frac{1}{\mathscr{r}}\mathbf{P}(\mathbf{r}')\right) - \frac{1}{\mathscr{r}}\nabla' \cdot \mathbf{P}(\mathbf{r}')\right]. \tag{33}$$

The divergence theorem may be used for the first term, so that:

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\underbrace{\int}_{\text{surface}} \frac{1}{\mathscr{r}}\mathbf{P}(\mathbf{r}') \cdot \hat{\mathbf{n}} da - \frac{1}{4\pi\epsilon_0}\int d\tau' \frac{1}{\mathscr{r}}\nabla' \cdot \mathbf{P}(\mathbf{r}') \tag{34}$$

As written, the computation of the potential due to the dielectric field $\mathbf{P}$ over a body may be interpreted as being the computation of the potential due to a surface charge density on the body:

$$\sigma_b \coloneqq \mathbf{P} \cdot \hat{\mathbf{n}} \tag{35}$$

and a volume charge density throughout the body:

$$\rho_b \coloneqq -\nabla \cdot \mathbf{P} \tag{36}$$

The subscript b is used for both densities to signify that these are *bound* charge densities — they arise from the dielectric nature of the material (the separation of atoms into dipoles). These are charges that are not free to move about.

- we define the **free charge** density as being the difference between the total charge density and the bound charge density:

$$\rho_f \coloneqq \rho - \rho_b. \tag{37}$$

- we now can show how the first of Maxwell's macroscopic equations arises. Start with

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0 \tag{38}$$

and substitute $\rho = \rho_b + \rho_f$, so that:

$$\nabla \cdot \mathbf{E} = \frac{\rho_b + \rho_f}{\epsilon_0}. \tag{39}$$

Using the definition of ρ_b, this becomes:

$$\nabla \cdot \mathbf{E} = \frac{-\nabla \cdot \mathbf{P}}{\epsilon_0} + \frac{\rho_f}{\epsilon_0} \tag{40}$$

which we may rearrange to obtain:

$$\nabla \cdot \left(\mathbf{E} + \frac{\mathbf{P}}{\epsilon_0}\right) = \frac{\rho_f}{\epsilon_0} \tag{41}$$

$$\nabla \cdot (\epsilon_0\mathbf{E} + \mathbf{P}) = \rho_f \tag{42}$$

which suggests defining a new vector field $\mathbf{D} \coloneqq \epsilon_0\mathbf{E} + \mathbf{P}$, the **electric displacement**, so that:

$$\boxed{\nabla \cdot \mathbf{D} = \rho_f} \tag{43}$$

- for magnetization, the introduction of $\mathbf{M}$ and $\mathbf{H}$ is quite analogous to that for $\mathbf{P}$ and $\mathbf{D}$. I will not go into $\mathbf{M}$ and $\mathbf{H}$ in detail, but rather summarize, pointing out the parallels with the dielectric polarization case.

- in the same way that there is dielectric polarization $\mathbf{P}$ that is the electric dipole moment per unit volume, there is a **magnetization** $\mathbf{M}$ which the magnetic dipole moment per unit volume. Equations 5.86 and 5.90 of GITE4 show how magnetic dipole moments $\mathbf{m}$ can be computed from arbitrary current distributions (the magnetic version of the definition of $\mathbf{p}$ in Eq. 19).

  In the same way that there is a simple model of a dipole (two separated point charge), the canonical "elementary" magnetic dipole is a circular loop of current:

   $\vec{m} = aI\hat{n}$

  If we consider a magnetized body, the magnetic fields may be computed by using superimposing the fields for an elementary dipole, based on the $\mathbf{M}$ field defined within the body (analogous to Eq. 34).

  In the same way as for the dielectric case, the required integration may be converted into the sum of a "surface integral" involving a bound volume current:

$$\mathbf{J}_b = \nabla \times \mathbf{M} \tag{44}$$

  and a volume integral involving bound surface current:

$$\mathbf{K}_b = \mathbf{M} \times \hat{\mathbf{n}} \tag{45}$$

  in a similar manner as Eq. 34.

  If the magnetic case was entirely analogous to the dielectric case, then we might *expect* to be able to define a free current density $\mathbf{J}_f \stackrel{?}{=} \mathbf{J} - \mathbf{J}_b$. However, there is an additional current density that we must consider due to changing dielectric polarization. Consider polarizing a sample; i.e., $\mathbf{P} = 0$ initially, but then $\mathbf{P} \neq 0$ after some time. The rate "separation of the charges" to produce dipoles is actually a current.[5] In fact $\partial \mathbf{P}/\partial t$ has the dimensions as volume current density $\mathrm{A/m^2}$. We call it the **polarization current** $\mathbf{J}_p$ and thus define $\mathbf{J_f} = \mathbf{J} - \mathbf{J}_b - \mathbf{J}_p$, so that:

$$\nabla \times \mathbf{B} = \mu_0(\mathbf{J}_b + \mathbf{J}_f + \mathbf{J}_p) + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \tag{46}$$

  Substituting $\mathbf{J}_p = \partial \mathbf{P}/\partial t$ and $\mathbf{J}_b = \nabla \times \mathbf{M}$ we obtain:

$$\nabla \times \mathbf{B} = \mu_0(\nabla \times \mathbf{M} + \mathbf{J}_f + \partial \mathbf{P}/\partial t) + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \tag{47}$$

[5] Recall the charge continuity equation: $\nabla \cdot \mathbf{J} = -\partial \rho/\partial t$, necessary for local charge conservation. Given the bound charge density $\sigma_b = -\nabla \cdot \mathbf{P}$, introduction of a polarization current $\partial \mathbf{P}/\partial t$, ensures that $\nabla \cdot \mathbf{J}_p = -\partial \rho_b/\partial t$; i.e., introduction of the polarization current ensures charge conservation.

Divide both sides by μ_0 and then slightly rearrange to obtain:

$$\nabla \times \left(\frac{\mathbf{B}}{\mu_0} - \mathbf{M} \right) = \mathbf{J}_f + \frac{\partial(\epsilon_0 \mathbf{E} + \mathbf{P})}{\partial t}. \tag{48}$$

Recognizing $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$, and defining $\mathbf{H} \coloneqq \mathbf{B}/\mu_0 - \mathbf{M}$, we have finally:

$$\boxed{\nabla \times \mathbf{H} = \mathbf{J}_f + \frac{\partial \mathbf{D}}{\partial t}} \tag{49}$$

The appearance of the time derivative of the electric displacement field reveals why we call Maxwell's addition the "displacement current", even though it has the form $\mu_0 \epsilon_0 \partial \mathbf{E}/\partial t$.

- now that we have looked at the introduction of **P** and **D** in detail, and **M** and **H** in a bit less detail (skipping some steps), we see that our two "new" Maxwell's equations (Eq. 43 and 49), are just different versions of the "plain old" microscopic Maxwell's equations.[6]

- one might be a bit bothered by the preceding electrostatic (and implicit magnetostatic) arguments, since it is claimed that these new macroscopic Maxwell's equations are general. More specifically, in the dielectric case, we computed the potential due to the **P** field using the electrostatic result for the potential due to a dipole, and then introduced equivalent volume and surface charge densities based on the resemblance of Eq. 34 to Coulomb's law (for the potential) — a form which is only valid in electrostatics.

  From this perspective, you may view the entire preceding argument as motivational rather than rigorous. For rigour you may consider that the introduction of the two fields **P** and **M** are simply ways to describe the parts of the "source" fields ρ and **J** that correspond to the dielectric polarization and magnetization of materials.

8 Interlude 1

That concludes our brief overview of both the "microscopic" and "macroscopic" forms of Maxwell's equations. Or to put it slightly differently: "Maxwell's equations" and "Maxwell's equations in a form sometimes useful in matter".

Notice that I have not discussed how **P** might depend on **E**, or how **M** might depend on **H**: the so-called **constitutive relations**. It is common to assume that **P** and **E** are related by a proportionality constant depending on the material in equation (and likewise for **M** and **H**); in other words, that the dielectric and magnetic responses of the material are linear, isotropic, and homogeneous.

[6] Stating that the macroscopic and microscopic versions of Maxwell's equations are equivalent requires qualification: they are equivalent *if* one accepts a continuous model of matter; i.e., there are no atoms. Otherwise, one must consider that there has been some sort of averaging over microscopic dipoles that has taken place, and that our fields are now spatially *averaged* fields. In that sense, the macroscopic Maxwell's equations are a phenomenological theory — their accuracy and utility must be considered based on how we are applying them.

There is another important case: a material may be magnetized $\mathbf{M} \neq 0$, even in the absence of an applied field — so called **permanent magnetism**. There are also “permanent dielectrics”, but they are not as common as permanent magnets.

You have almost certainly considered both permanent magnetism and the linear dielectric and magnetic responses before. But for now, we put these ideas away, and will ignore $\mathbf{P}$ and $\mathbf{M}$, only returning to them at the end of the course, when we consider the propagation of electromagnetic waves in matter. At that point, we will discuss how the “dielectric and magnetic susceptiblities” (essentially the proportionality constants) may be considered to be frequency dependent, and some simple physics based models for that dependence.

Now we move on to special relativity.

9 Relativistic velocity addition (GITE4 12.1.1)

- in mechanics problems it is common to use the velocity addition rule:

$$\mathbf{v}_{AC} = \mathbf{v}_{AB} + \mathbf{v}_{BC}, \tag{50}$$

The velocity of A observed wrt to reference frame C

The velocity of A observed wrt to reference frame B

The velocity of reference frame B observed wrt to reference frame C

  a relation that seems almost unquestionable, as if it is a tautology.

  For example, consider a ping-pong ball (B) hit by a moving paddle (P). Assuming that the paddle’s motion is unaffected by the collision, the ball’s motion is most easily analyzed, not in the earth’s frame (E), but in the frame in which the paddle is stationary. In a one-dimensional model we determine the ball’s velocity in the paddle frame, before the collision, using:[7]

$$v_{BP,x,\text{before}} = v_{BE,x,\text{before}} - v_{PE,x}. \tag{51}$$

  Then in the straightforward case that the collision with the bat is elastic, the ball’s velocity in the paddle frame just reverses, so that $v_{BP,x,\text{after}} = -v_{VE,x,\text{before}}$, and then we use

$$v_{BE,x,\text{after}} = v_{BP,x,\text{after}} + v_{PE,x} \tag{52}$$

  to get the final velocity of the ball in the earth’s frame, which is what we are interested in.

- despite the “obviousness” of Eq. 50, it is not true in special relativity. Instead, in one spatial dimension we have the **relativistic velocity addition formula**:

$$v_{AC,x} = \frac{v_{AB,x} + v_{BC,x}}{1 + \dfrac{v_{AB,x} v_{BC,x}}{c^2}}, \tag{53}$$

[7] In this section, all motion is in one spatial dimension, which I take to be x, and indicate the projections by $v_{BP,x} \coloneqq \mathbf{v}_{BP} \cdot \hat{\mathbf{x}}$. In an earlier version of these notes (and my videos), I indicated that in this section omitting the vector sign referred to the projection along the axis, not the magnitude. What a terrible idea ☹!

where c is the speed of light. (In cases where the velocity vectors are non-collinear the formula is a bit more complicated; we will just discuss one spatial dimension in this section.)

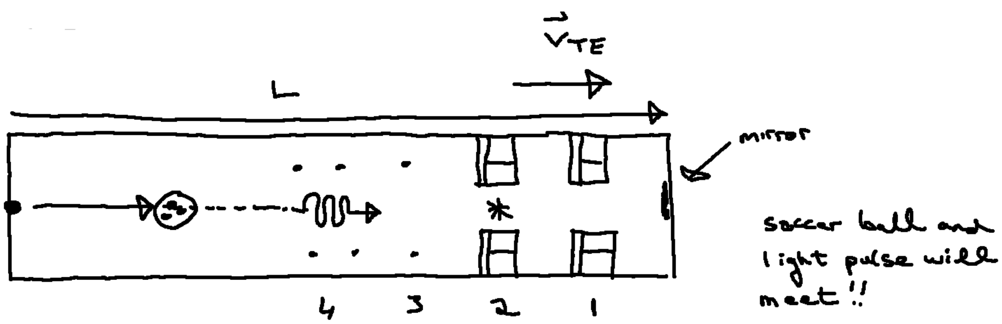
Note that Eq. 53 *does* reduce to the familiar Eq. 50 in the non-relativistic limit: $|v_{AB,x}|/c \ll 1$ and $|v_{BC,x}|/c \ll 1$.

- GITE4 calls attention to the relativistic velocity addition rule in his introductory section for special relativity (12.1.1). That is unorthodox, but enlightened — the conventional approach to special relativity starts with the jarring concepts of non-simultaneity, time dilation, and length contraction. These notions are completely absent from non-relativistic physics.

  On the other hand, the relativistic velocity addition formula is a modification of something that we are quite used to, namely Eq. 50.

  GITE4 does not actually show how the relativistic velocity addition law arises until after his introduction to the Lorentz transformations. But there is a nice way to derive this result from the basic ideas of SR, first introduced by Mermin [4]. I will go over Mermin's derivation now.

- imagine taking a VIA train from Kitchener to Toronto.[8] You are standing at the back of a coach. You kick a soccer ball down the aisle towards the front of the train. Simultaneously you also send a short pulse of light in the same direction, whereupon reaching the front of the coach, it partially reflects on a window, and some of the light heads back towards you. The soccer ball and the returning light pulse will then meet somewhere along the aisle in front of you.



- you *could* characterize the meeting place of the soccer ball and light pulse by how far ahead of you that they meet; i.e., a distance. But any prior knowledge of SR will suggest to you that a distance such as this might be dependent on your reference frame. Instead, imagine characterizing the meeting place by the seat row number along the coach. You could rely on a person sitting in the seat adjacent to the meeting point to give you that information. That seat row number cannot not change between reference frames — it is an **invariant**.

  The strategy for derivation of the velocity addition law will be to determine this invariant in two **inertial frames**:

  (1) the train frame, in terms of $v_{BT,x}$, and

[8] Yes, yes, I know. That will involve highly non-relativistic speeds. I will also assume that you can *really* kick a soccer ball.

(2) the earth frame, in terms of $v_{BE,x}$ and $v_{TE,x}$,

where B refers to the ball, T to the train, and E to the earth.

By equating the invariant seat row number computed in these two frames, we (hope to be able to) determine $v_{BE,x}$ in terms of $v_{BT,x}$ and $v_{TE,x}$, without any reference to the specifics of the coach length, times, etc...

- suppose that there are N rows of seats in the coach (also an invariant), evenly spaced along the train, and that the ball and light pulse meet at row i. I will use $g := i/N$ as the frame invariant.

  We will find that g can be computed independently of L in each frame, which is good, as we will avoid assuming that the coach's length L, is the same in both the earth and coach's frame (its **rest frame**).

- we start with the earth frame case. The time duration from the kick until the meeting of the ball and light pulse will be divided into two parts:

$$t = \Delta t_1 + \Delta t_2, \tag{54}$$

where Δt_1 is the time it takes the light pulse to make it to the reflection at the front of the coach, and Δt_2 is the time that it takes from the reflection until the meeting with the soccer ball.

While the light pulse is travelling towards the front of the coach, the coach moves forward by $v_{TE,x}\Delta t_1$. Thus the amount of distance that the light pulse has to travel is $L + v_{TE,x}\Delta t_1$. Thus, $c\Delta t_1 = L + v_{TE,x}\Delta t_1$ which may be rearranged to give:

$$\Delta t_1 = \frac{L}{c - v_{TE,x}}. \tag{55}$$

To make it to the meeting point, the light pulse will have to travel a distance of $gL - v_{TE,x}\Delta t_2$. Note that now the required distance is *diminished* due to the train's motion. Thus $c\Delta t_2 = gL - v_{TE,x}\Delta t_2$, which may be rearranged to give:

$$\Delta t_2 = \frac{gL}{c + v_{TE,x}}. \tag{56}$$

Recall that it is g that we want; let us divide Eq. 56 by 55, to obtain:

$$\frac{\Delta t_2}{\Delta t_1} = g\,\frac{(c - v_{TE,x})}{(c + v_{TE,x})}, \tag{57}$$

which we may rearrange to give:

$$g = \frac{(c + v_{TE,x})}{(c - v_{TE,x})}\,\frac{\Delta t_2}{\Delta t_1}. \tag{58}$$

The ratio $\Delta t_2/\Delta t_1$ may be determined by considering the conditions for the collision to occur. From the kick to the collision, the ball travels: $v_{BE,x}(\Delta t_1 + \Delta t_2)$. Likewise, at the collision time, the light pulse will be a distance $c(\Delta t_1 - \Delta t_2)$ from where it started together with the ball. (Remember that the light pulse direction reverses, travelling in the $+v$ direction for Δt_1 and the negative direction for Δt_2.) For a collision, we equate these two distances, giving:

$$v_{BE,x}(\Delta t_1 + \Delta t_2) = c(\Delta t_1 - \Delta t_2) \tag{59}$$

which may be rearranged to obtain:

$$\frac{\Delta t_2}{\Delta t_1} = \frac{c - v_{BE,x}}{c + v_{BE,x}} \tag{60}$$

then substituted into Eq. 58 to give:[9]

$$g = \frac{(c + v_{TE,x})}{(c - v_{TE,x})} \frac{(c - v_{BE,x})}{(c + v_{BE,x})}. \tag{61}$$

- now we want the analogous expression to Eq. 61 for g, but computed in the train frame in terms of $v_{BT,x}$. But do we really have to redo everything in this new frame? No! We *could* rewrite everything, using $\Delta t'_1$ instead of Δt_1, $\Delta t'_2$ instead of Δt_2, and L' instead of L, using the $'$s to denote the fact that these times and distances may be different in the train frame, but we would just get the result of Eq. 61, with the replacements $v_{TE,x} \to v_{TT,x} = 0$ and $v_{BE,x} \to v_{BT,x}$:

$$g = \frac{(c - v_{BT,x})}{(c + v_{BT,x})}. \tag{62}$$

Almost finished! We now equate the invariant g, as computed in both frames, Eq.'s 61 and 62, to give the desired relationship between $v_{BE,x}$, $v_{BT,x}$, and $v_{BE,x}$:

$$\frac{(c - v_{BT,x})}{(c + v_{BT,x})} = \frac{(c + v_{TE,x})}{(c - v_{TE,x})} \frac{(c - v_{BE,x})}{(c + v_{BE,x})} \tag{63}$$

which may be rearranged to obtain:

$$\frac{(c - v_{BE,x})}{(c + v_{BE,x})} = \frac{(c - v_{BT,x})}{(c + v_{BT,x})} \frac{(c - v_{TE,x})}{(c + v_{TE,x})}. \tag{64}$$

A relevant piece of mathematical minutiae: some simple algebra shows that the function

[9] Recall a physicist's mantra: *check limiting cases*. If you kick the ball so hard that it travels close to the speed of light, then $v_{BE,x} \approx c$, and Eq. 61 gives $g \approx 0$; i.e., the ball and returning light pulse meet at the front of the coach, as expected. And if you do not kick the ball at all, $v_{BE,x} = v_{TE,x}$, and then Eq. 61 gives $g = 1$; i.e, the light pulse and ball meet at your location, also as expected.

$y = (1-x)/(1+x)$ is its own inverse, so that $x = (1-y)/(1+y)$. Applying this here gives:

$$\frac{v_{BE,x}}{c} = \frac{1 - \dfrac{(c - v_{BT,x})}{(c + v_{BT,x})}\dfrac{(c - v_{TE,x})}{(c + v_{TE,x})}}{1 + \dfrac{(c - v_{BT,x})}{(c + v_{BT,x})}\dfrac{(c - v_{TE,x})}{(c + v_{TE,x})}} \tag{65}$$

$$= \frac{(c + v_{BT,x})(c + v_{TE,x}) - (c - v_{BT,x})(c - v_{TE,x})}{(c + v_{BT,x})(c + v_{TE,x}) + (c - v_{BT,x})(c + v_{TE,x})} \tag{66}$$

$$= \frac{2cv_{TE,x} + 2cv_{BT,x}}{2c^2 + 2v_{TE,x}v_{BT,x}} \tag{67}$$

$$\boxed{v_{BE,x} = \frac{v_{BT,x} + v_{TE,x}}{1 + \dfrac{v_{BT,x}v_{TE,x}}{c^2}}} \tag{68}$$

as required □.

- later we will rederive the relativistic velocity addition formula, starting from the Lorentz transformations, and also consider the non-collinear case. But I like the preceding derivation (due to Mermin [4]), as it offers another context for consideration of the tenets of special relativity. I will recap the important points:

  (1) we took the speed of light[10] to be the same in the train and earth frame! Note how this assumption slipped into the derivation when I adapted the formula for g in the earth frame to the formula for g in the train frame (Eq. 61 to 62).

  (2) we assumed that in the two inertial frames, the soccer ball travelled at a constant velocity (obviously neglecting friction, air resistance, and so-on). In fact, this property should be taken as the definition of an inertial frame.[11]

  (3) we identified a relevant invariant: something that remained constant, independent of the inertial frame. In this derivation, we took the existence of the "seat row number" invariant as manifest. Sometimes we will justify invariance; sometimes we will assume it. For example, later we will assume that electric charge is an invariant, or more precisely, a **Lorentz invariant**; e.g., an electron's charge is the same in all inertial frames. This assumption can be, and has been, physically tested.

- there is an interesting condensation of the relativistic velocity addition formula. Consider the following purely mathematical identity (DLMF 4.35.3):

$$\tanh(u \pm v) = \frac{\tanh u \pm \tanh v}{1 \pm \tanh u \tanh v}. \tag{69}$$

[10] To be more accurate, it is the speed of light *in vacuum* that remains the same in different inertial frames. But for visible light there is only a small fractional difference between the speed of light in vacuum and in air, and thus we just use c in this section. Likewise, unless otherwise instructed, in this course you may assume that "air" is close enough to vacuum that we need not consider that the speed of light in air differs from that in vacuum. We shall discuss this issue in much more detail in the next section.

[11] "... a table will remain a table unless it gets broken or something." https://www.newyorker.com/humor/daily-shouts/an-oral-history-of-isaac-newton-discovering-gravity-as-told-by-his-contemporaries.

Notice the similarity in form to the relativistic velocity addition formula, Eq. 68. If we consider a change of variables of each velocity projection, v_x, to θ_x, so that $v_x/c = \tanh\theta_x$, then Eq. 69 allows the relativistic velocity addition formula, Eq. 68, to be written as:

$$\boxed{\theta_{AC,x} = \theta_{AB,x} + \theta_{BC,x}} \tag{70}$$

The θ representations of velocity components are known as **rapidities** and I will use them a bit more than GITE4 does.

10 An experimental test of relativistic velocity addition

- experimentally testing the relativistic velocity addition formula, Eq. 53, seems difficult when both $|v_{AB,x}|/c \ll 1$ and $|v_{BC,x}|/c \ll 1$, as it simply reduces to the non-relativistic formula.
- ideally for a test we want at least one of the speeds to be comparable to that of light. A historically important experiment due to Fizeau makes one of the speeds that of light in a moving material (comparable in magnitude to that of the speed in vacuum), and the other that of the moving material itself.

  In this section, I will discuss this so-called **Fizeau aether drag** experiment. In particular, I will show how the data from an informative recreation of this experiment due to Lahaye *et al.* [5] supports the relativistic velocity addition formula. (I mainly follow the notation and presentation of Lahaye *et al.'s* paper, which also contains a comprehensive set of references to the history of this influential experiment.)
- recall from optics the **refractive index**:

speed of light
in vacuum

$$n := \frac{c}{v_p} \tag{71}$$

phase velocity

At visible wavelengths, $n \approx 1.33$ for water, and $n \approx 1.003$ for air.

Let me be precise about what the phase velocity is: suppose that we have an electromagnetic wave, with wavelength λ and frequency f, travelling in the $+$ve x direction:

$$\mathbf{E}(x,t) = \mathbf{E}_0 \cos \underbrace{\left(\frac{2\pi}{\lambda}x - 2\pi f t + \phi\right)}_{=:\varphi} \tag{72}$$

and that we observe this wave while travelling at speed v in the positive x direction. i.e, our location is

$$x = vt. \tag{73}$$

Substituting this position into the "phase" of Eq. 72 we obtain

$$\varphi = \frac{2\pi vt}{\lambda} - 2\pi ft + \phi \tag{74}$$

$$= 2\pi t\left(\frac{v}{\lambda} - f\right) + \phi, \tag{75}$$

showing that if our speed v satisfies

$$v/\lambda = f \tag{76}$$

then we will "ride the wave" with constant phase.[12]

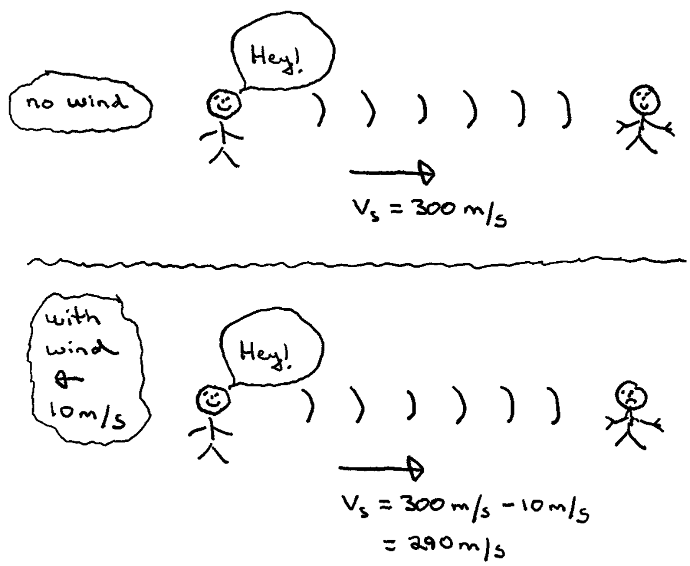
As such, the definition of **phase velocity** is

$$v_p := f\lambda. \tag{77}$$

The refractive index is just a handy way to compare v_p to c, the phase velocity in vacuum.

- what if the medium is moving?

  Think of a somewhat analogous situation with sound waves in air ($\approx 300\,\mathrm{m/s}$):



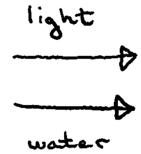
  The sound takes longer to get to you when it is travelling into an opposing wind.

  Clearly velocity addition is important when considering wave propagation in a moving medium.

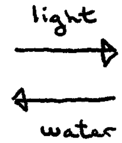
  It makes sense that n and thus v_p are for the medium when it is at rest, and that v_p will be different in a frame in which the medium is moving.

  The "idea" behind Fizeau's experiment is a comparison between the phase velocities of light, measured in the lab frame, between light travelling in the *same* direction as moving water:

[12] Einstein wondered about what you would observe when "riding along" with an electromagnetic wave (in vacuum) when he was 16. The paradoxes that result without SR — namely that Maxwell's equations are not satisfied in this moving frame — were precursors to his later thinking. In his later days, he reflected upon his early thinking: *"Invention is not the product of logical thought, even though the final product is tied to a logical structure."* For more details see Section 6d, pg 131, of Pais' biography of Einstein [6].

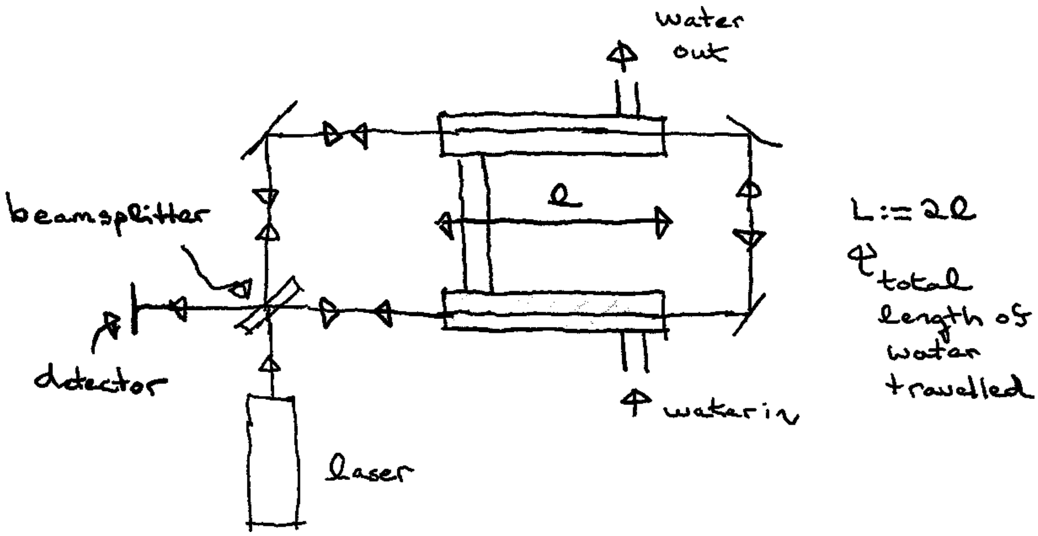


and in the *opposite* direction as water:



The results of the comparison depend on which velocity addition law we use, and provide confirmation that the relativistic law is the correct one.

Here is a simplified drawing of the modern apparatus of Ref. [5]:



There are two paths that the light can take from the laser to the detector:

1. counter-clockwise (CCW): reflect off the beam splitter, travel counterclockwise around loop (against the flow of the water), reflect off the beam splitter and then go into detector.
2. clockwise (CW): transmit through the beam splitter, travel clockwise around loop (with the flow of water) and go towards the detector.

The light from these two paths can *interfere* at the detector depending on the relative phase. e.g., if π phase difference is accumulated, the two paths destructively interfere, and no light is seen at the detector. For zero phase difference, or integer multiples of 2π phase difference, the interference is completely constructive — the maximum amount of light will be detected. The interferometer measures the relative phase of the two paths.

The only variation in the relative phase between the two paths will be due to their passage through the water, for which we can control the flow rate. Thus we will concentrate on the phase accumulated by passing through the water, and its dependence on the water velocity.

From Eq. 72, the phase accumulated by passage through the water will be:

$$\Delta\varphi = \frac{2\pi}{\lambda} L. \tag{78}$$

By Eq. 76, we substitute $\lambda = v_p/f$ to obtain:

$$\Delta\varphi = 2\pi \frac{Lf}{v_p}. \tag{79}$$

For water *stationary* in the laboratory frame, we may use $v_p = c/n$, and there is no difference in the phase accumulated by passage through the water in the CCW and CW paths.

The situation changes if water is flowing, because for the water flow direction shown in the diagram, the CCW beam travels against the water flow direction, and the CW beam travels with the flow direction.

I will first consider the phase difference between CW and CCW, assuming the non-relativistic velocity addition law is valid (which it is not). For the CCW beam:

$$v_{p,CCW} \stackrel{?}{=} \frac{c}{n} - v_{w,CW}, \tag{80}$$

analogous to the example that we gave for sound waves, with $v_{w,CW}$ positive for water flow in the clock-wise direction, as shown in the apparatus diagram.

Likewise, for the CW path:

$$v_{p,CW} \stackrel{?}{=} \frac{c}{n} + v_{w,CW} \tag{81}$$

As I have discussed, the interferometer measures the phase difference of the CCW and CW beams:[13]

$$\begin{aligned}
\Delta\varphi_{\text{non-relativistic}} &= \Delta\varphi_{CCW} - \Delta\varphi_{CW} && (82)\\
&= 2\pi L f \left(\frac{1}{v_{p,ccw}} - \frac{1}{v_{p,cw}} \right) && (83)\\
&= 2\pi L f \left(\frac{1}{c/n - v_{w,CW}} - \frac{1}{c/n + v_{w,CW}} \right). && (84)
\end{aligned}$$

Rewriting in a form suitable for an expansion:

$$\Delta\varphi_{\text{non-relativistic}} = \frac{2\pi L f}{c/n} \left(\frac{1}{1 - n v_{w,CW}/c} - \frac{1}{1 + n v_{w,CW}/c} \right). \tag{85}$$

Since $n_w \approx 1.33$ and $v_{w,CW} \ll c$, the quantity $|n v_{w,CW}/c| \ll 1$ and thus we may use $1/(1+x) = 1 - x + x^2 + \ldots$ on each of the bracketed terms to obtain:

$$\Delta\varphi_{\text{non-relativistic}} = \frac{2\pi L f}{c/n} \left[1 + \frac{n v_{w,CW}}{c} + \left(\frac{n v_{w,CW}}{c} \right)^2 + \ldots - \left(1 - \frac{n v_{w,CW}}{c} + \left(\frac{n v_{w,CW}}{c} \right)^2 + \ldots \right) \right]. \tag{86}$$

[13] Each path shares a common time dependent phase, which drops out of the difference. I am dropping any phase difference that exists with zero water flow, as it will just add a constant offset independent of $v_{w,CW}$.

To lowest non-vanishing order:

$$\Delta\varphi_{\text{non-relativistic}} \approx \frac{2\pi L f}{c/n}\frac{2nv_{w,CW}}{c} \tag{87}$$

$$\boxed{\Delta\varphi_{\text{non-relativistic}} \approx 4\pi\frac{Lf}{c}\frac{v_{w,CW}}{c}n^2} \tag{88}$$

Thus the phase difference scales linearly with $v_{w,CW}$: increasing $v_{w,CW}$ will cycle through constructive and destructive interference.

Now let us repeat this calculation of $\Delta\varphi$, but this time using the relativistic velocity addition formula.

The relativistic velocity addition formula gives $v_{p,CCW} = (c/n - v_{w,CW})/(1 - v_w(c/n)/c^2)$ and similarly for $v_{p,CW}$, which we substitute into Eq. 83 to obtain:

$$\Delta\varphi_{\text{relativistic}} = \Delta\varphi_{CCW} - \Delta\varphi_{CW} \tag{89}$$

$$= 2\pi L f\left[\frac{(1 - v_{w,CW}/(nc))}{c/n - v_{w,CW}} - \frac{(1 + v_{w,CW}/(nc))}{c/n + v_{w,CW}}\right]. \tag{90}$$

Again, we rearrange into a form suitable for expansion:

$$\Delta\varphi_{\text{relativistic}} = \frac{2\pi L f}{c/n}\left[\frac{(1 - v_{w,CW}/(nc))}{1 - nv_{w,CW}/c} - \frac{(1 + v_{w,CW}/(nc))}{1 + nv_{w,CW}/c}\right]. \tag{91}$$

Again, $|nv_{w,CW}/c| \ll 1$, so we expand the denominators to obtain:

$$\begin{aligned}\Delta\varphi_{\text{relativistic}} \approx \frac{2\pi L f}{c/n}&[(1 - v_{w,CW}/(nc))(1 + nv_{w,CW}/c - (nv_{w,CW}/c)^2 + \ldots) \\ &- (1 + v_{w,CW}/(nc))(1 - nv_{w,CW}/c + (nv_{w,CW}/c)^2 + \ldots)].\end{aligned} \tag{92}$$

which to lowest non-vanishing order gives:

$$\Delta\varphi_{\text{relativistic}} \approx \frac{2\pi L f}{c/n}2\left[\frac{nv_{w,CW}}{c} - \frac{v_{w,CW}}{nc}\right] \tag{93}$$

$$\boxed{\Delta\varphi_{\text{relativistic}} \approx 4\pi\frac{Lf}{c}\frac{v_{w,CW}}{c}(n^2 - 1)} \tag{94}$$

This result for $\Delta\varphi$ should be contrasted to the non-relativistic result, Eq. 88. In both cases $\Delta\varphi$ is proportional to $v_{w,CW}$. However, the relativistic result differs by a factor of $(n^2 - 1)/n^2$ from the non-relativistic one. For $n = 1.33$ (water at 532 nm), $(n^2 - 1)/n^2 \approx 0.43$.

The data of Lahaye *et al.* [5] conclusively support the relativistic velocity addition formula:

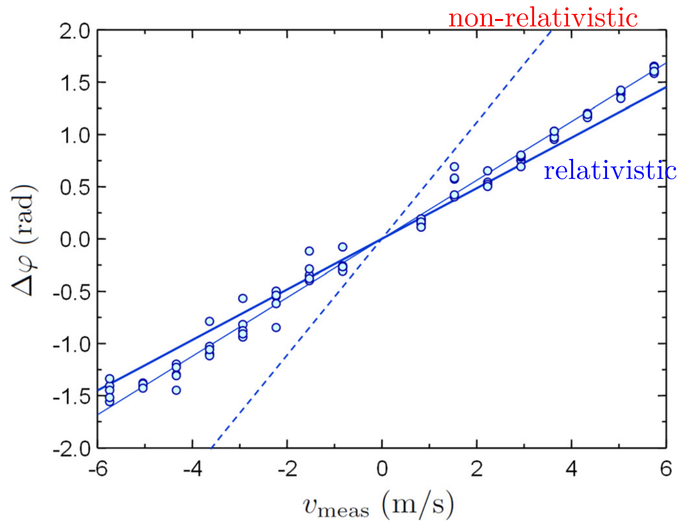


Figure 1: The phase difference between the CCW and CW paths as a function of the speed of the water. The non-relativistic prediction (Eq. 88) and relativistic predictions (Eq. 94) are shown with lines, together with a fit to the experimentally measured data points. (Figure 6 of Lahaye *et al.* [5]).

11 The relativity of simultaneity (GITE4 12.1.2(i))

- suppose that you release two pulses of light from the center of a train coach, one travelling towards the front, and one travelling towards the back. No matter whether or not the train is moving with respect to the earth, both pulses will reach the ends of the coach simultaneously according to people on the train.

  But if the train is moving forwards, someone at rest on the earth will think that the light pulse going backwards will arrive first, as it has less distance to travel (in the earth's frame) because of the train's forward motion.

  Thus for two events — such as the light pulses hitting the two ends of the coach — whether or not they occur at the same time, their so-called **simultaneity**, is dependent on the reference frame. In the train example, an observer in the reference frame of the train will consider the events simultaneous, whereas earth observers will not.

- in many situations there is a "trivial" *apparent* lack of simultaneity related to the propagation of signals. For example, when lighting strikes, we hear it after we see it. For a 1 s delay, we may estimate how far away lightning struck: $d \approx 340\,\text{m/s} \times 1\,\text{s} \approx 0.34\,\text{km}$, where we have taken the speed of sound to be $340\,\text{m/s}$ and assumed the light flash arrived instantaneously due to its much higher speed.

  The "simultaneous" events of the production of sound and light at the location of the lightning strike do not "appear" simultaneous to us. However, they would be if we corrected for the signal propagation times.

  Similarly, if after releasing the two light pulses from the center of the train, we walked a bit forward, we would *see* the light pulse hitting the front of the train first, because of the change in the time required for this light to make its way back to us.

All of these signal propagation delays may be corrected for — at least in principle. In this course we adopt GITE4's terminology and use the terms **observe** and **observer** in a very precise way to denote that our observations have accounted for the propagation effects. For example, with this terminology, we say that a person *observes* that thunder and lightning occur at the same time, no matter how far away they are from the lightning strike. All the passengers on the moving train *observe* that the light pulses hit the two ends at the same time, no matter where they are sitting. On the other hand, an *observer* in the earth's frame does not consider the pulses hitting the ends of the coach to be simultaneous. At the risk of belabouring a point: the relativity of simultaneity is not due to signal propagation delays! It is much more fundamental than that.

- one way that events may be *observed* is to fill an inertial reference frame up with a "lattice" of observers, evenly spaced, with synchronized clocks, and then rely on them to report back their observations to you. Note that "synchronization" implies simultaneity, which is only guaranteed if all clocks are stationary in a single common inertial reference frame.

12 Relativistic time dilation (GITE4 12.1.2(ii))

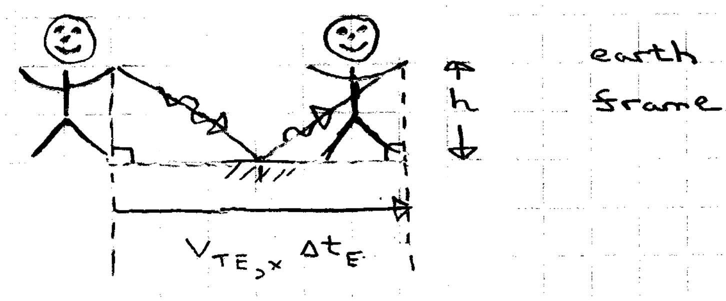
- my treatment here (and for relativistic length contraction and the Lorentz transformations) follows that found in GITE4, Section 12.1.2. It is concrete, but clunky: moving trains with a generous topping of anthropomorphization. To those already familiar with these arguments, who desire a cleaner, more abstract approach, see Landau and Lifshitz [2]. They base their derivation of the Lorentz transformations on the invariance of the interval. But if that means nothing to you (yet), read on.
- suppose that you let a light pulse travel from $h = 1\,\text{m}$ above ground to the ground directly beneath you. It reflects on a mirror and heads back to you.

You consider the time between emitting the pulse and receiving the reflected pulse back to be a single "tick" of a clock, of duration:

$$\Delta t_T = \frac{2h}{c} \tag{95}$$
$$\approx \frac{2\,\text{m}}{3\times 10^8\,\text{m/s}} \tag{96}$$
$$\approx 6\,\text{ns}. \tag{97}$$

Now suppose that you use this clock on a moving train coach. The train moves to the right, in a direction orthogonal to the direction that you send the light pulse.

To observers in the earth's frame the light pulse moves to the right, so as to meet you, since you are moving:



Our goal is now to determine how long a observer in the earth's frame thinks a tick of your clock is, Δt_E.

Critical to the following analysis is that h is the *same* in both the earth's and train's frame. This is a general result; specifically, that distances perpendicular to the relative velocity between two inertial frames remain unchanged. Imagine trying to run through a doorway holding barbells in the normal way. Do the barbells make it through or not? Everyone must agree, no matter what their reference frame.[14]

Using Pythagoras' theorem, the distance travelled in the earth's frame is

$$d = 2\sqrt{h^2 + \left(\frac{v_{TE,x}\Delta t_E}{2}\right)^2} \tag{98}$$

$$d = \sqrt{(2h)^2 + (v_{TE,x}\Delta t_E)^2}. \tag{99}$$

Using the *same* speed of light in earth's frame as in the train's frame:

$$c\Delta t_E = d \tag{100}$$

$$= \sqrt{(2h)^2 + (v_{TE,x}\Delta t_E)^2} \tag{101}$$

which may be rearranged to solve for Δt_E:

$$(c\Delta t_E)^2 - (v_{TE,x}\Delta t_E)^2 = (2h)^2 \tag{102}$$

$$(\Delta t_E)^2 = \frac{(2h)^2}{c^2 - v_{TE,x}^2} \tag{103}$$

[14] A key point is that both ends of the barbells passing through the door way are simultaneous events in both your and the door's rest frame.

We can write h in terms of the duration of the clock ticks in the train frame, $\Delta t_T = 2h/c$, so that:

$$(\Delta t_E)^2 = \frac{(\Delta t_T)^2}{1 - (v_{TE,x}/c)^2} \tag{104}$$

$$\boxed{\Delta t_E = \frac{\Delta t_T}{\sqrt{1 - (v_{TE,x}/c)^2}}} \tag{105}$$

It is common to introduce $\beta \coloneqq v_{TE,x}/c$ and $\gamma \coloneqq 1/\sqrt{1-\beta^2}$, so that $\Delta t_E = \gamma \Delta t_T$. Or equivalently, in terms of rapidity θ corresponding to the relative velocity of the two frames, $\Delta t_E = \cosh(\theta)\Delta t_T$. [15]

Since $\gamma \geq 1$, we see that the clock ticks observed in the earth's frame will always be of longer duration as compared to as observed in the train's frame.

This so-called time dilation ("expansion") gives rise to the slogan: *moving clocks are observed to run slow*, meaning that for whatever time we think has elapsed in our frame, the time will elapsed will be less in the rest frame of the clock. Or going in the other direction: the clock rest frame time durations are *dilated* in frames in which the clock is moving.

Time dilation is so abstract that I will immediately give some sort experimental evidence for it in the next section.

- food for thought: the "bouncing light" clock relied on the existence of two spatial dimensions. Is that assumption fundamental? i.e., is SR a theory of $N+1$ space-time, where $N \geq 2$ (but $N = 3$ for us), or is this a limitation of this particular derivation?

13 Experimental observation of time dilation

- the decay of unstable particles and atoms can be used to construct clocks. You may be familiar with the example of radioactive dating from geology. The decay process

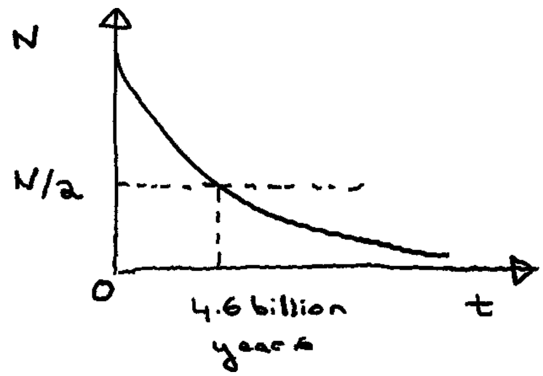
$$\underbrace{{}^{238}\mathrm{U}}_{\text{uranium}} -> \underbrace{{}^{206}\mathrm{Pb}}_{\text{lead}} + \dots \tag{106}$$

has a **half-life** of $t_{1/2} = 4.6$ billion years:

[15] From the mathematical identity, $\cosh^2\theta - \sinh^2\theta = 1$, we may determine that

$$\cosh\theta = \frac{1}{\sqrt{1 - (\tanh\theta)^2}}$$

and thus recognize that $\gamma = \cosh\theta$, since $v_{TE,x}/c = \tanh\theta$.



When the mineral zircon ($Zr\,Si\,O_4$) solidifies it contains small amounts of uranium, but no lead.

As time progresses, some of the uranium decays into lead.

Thus by measuring the amount of uranium and lead in a sample, we can establish how long ago the zircon crystallized. For example, suppose that we found a zircon sample with equal amounts ^{238}U and ^{206}Pb. We may infer that the zircon crystallized 4.5 billion years ago.

By this procedure, the *oldest* zircon that we can find on earth is determined to have crystallized about 4.4 billion years ago, thereby providing a *lower* bound on the age of the earth.

- now let us consider an analogous "clock" on a completely different time scale, based on the decay of **muons** (μ^-). A muon is roughly the 200 times the mass of an electron, and spontaneously decays with a **lifetime**[16] of $t_{1/e} = 2.2 \times 10^{-6}$ s:

$$\mu^- \to e^- + \bar{\nu}_e + \nu_\mu, \tag{107}$$

where e^- is an electron, $\bar{\nu}_e$ is an electron antineutrino, and ν_μ is a muon neutrino. Decaying muons are a "clock" that may be used to experimentally observe time dilation.

- more specifically, muons are made in the upper atmosphere of the earth and travel down towards the earth's surface with speeds close to c. The flux of muons F decreases as we go to lower altitudes in the earth's atmosphere because some of the muons decay on the way. Naively, we might think that we could take the muon speed v and divide it into the distance travelled d to get time elapsed $\Delta t = d/v$. Then the reduction of flux would be described by

$$F_2 \stackrel{?}{=} F_1 \exp\left(-\frac{\Delta t}{t_{1/e}}\right), \tag{108}$$

where F_1 is the flux at the higher elevation, and F_2 is the flux at a lower elevation.

But Eq. 108 with $\Delta t = d/v$ does not account for relativistic time dilation. Let us look at some specific numbers from an illustrative version of this so-called **Rossi-Hall** experiment, as performed by Frisch and Smith [7].[17]

[16] Instead of the "half-life", we will use "lifetime" $t_{1/e}$ so that the exponential decay factor may be conveniently written as $\exp(-\Delta t/t_{1/e})$, instead of $2^{-\Delta t/t_{1/2}}$.

[17] There is also a short movie to accompany Frisch and Smith's paper. It is helpful to know that in earlier literature, including the article by Frisch and Smith [7], muons are referred to as "μ-mesons". This terminology is no longer used.

At the top of Mount Washington, New Hampshire, 1920 m above sea-level, Frisch and Smith put a muon detector sensitive to muons travelling at a speed of $v \approx 0.9952c$ towards the earth. Over one hour, this detector counted $N_1 \approx 560$ muons.

The same detector was then brought to sea-level, where over one hour, $N_2 \approx 410$ muons were counted.

As we have noted, less muons were observed because some decay as they make their way down towards sea level. Let us examine this decay quantitatively. In the earth's frame E it takes muons of $v \approx 0.9952c$ at $h = 1920\,\mathrm{m}$,

$$\Delta t_E = \frac{\Delta h}{v} \tag{109}$$

to reach sea-level. Substituting numbers gives:

$$\begin{align} \Delta t_E &\approx \frac{1920\,\mathrm{m}}{0.9952 \times 2.998 \times 10^8\,\mathrm{m/s}} \tag{110} \\ &\approx 6.44 \times 10^{-6}\,\mathrm{s} \tag{111} \end{align}$$

Thus, we might naively expect the number of muons that we observe in one hour of counting at sea level to be

$$\begin{align} N_2 &\stackrel{?}{=} N_1 e^{-\Delta t_E / t_{1/e}} \tag{112} \\ &\stackrel{?}{\approx} 560\, e^{-6.44 \times 10^{-6}\,\mathrm{s} / 2.2 \times 10^{-6}\,\mathrm{s}} \tag{113} \\ &\stackrel{?}{\approx} 30. \tag{114} \end{align}$$

But as we have noted, $N_2 \approx 410$ muons were counted at sea level!

Remember: *moving clocks are observed to run slow.*

It is not the time elapsed in the earth's frame that is relevant to the muon decay, but instead the time elapsed in the *muon's frame*:

$$\begin{align} \Delta t_\mu &= \frac{\Delta t_E}{\gamma} \tag{115} \\ &\approx 6.44 \times 10^{-6}\,\mathrm{s} \times \sqrt{1 - 0.9952^2} \tag{116} \\ &\approx 6.3 \times 10^{-7}\,\mathrm{s}. \tag{117} \end{align}$$

So roughly a factor of 10 less time elapsed in the muon frame, as compared to the earth frame, and thus there is much less decay (than we naively expected):

$$\begin{align} N_2 &= N_1 e^{-\Delta t_\mu / t_{1/e}} \tag{118} \\ &\approx 560 e^{-6.3 \times 10^{-7}\,\mathrm{s} / 2.2 \times 10^{-6}\,\mathrm{s}} \tag{119} \\ &\approx 420, \tag{120} \end{align}$$

which is much closer to the number of muons that Frisch and Smith observed at sea-level $N_2 \approx 410$ [7].

Thus we see that when particles travel close to the speed of light, relativistic effects can be quite "non-subtle".

14 Relativistic length contraction (GITE4 12.1.2(iii))

- now let us consider the release of a pulse of light from the back of a moving train coach to the front, where a mirror then reflects it back.

  How long does this trip take in the train's frame? That is straightforward:

$$\Delta t_T = 2\frac{L_T}{c} \tag{121}$$

  where L_T is the length of the train coach in its rest frame.

  We will now compute the duration in the earth's frame in terms of the length of the train coach in the earth's frame. If we combine this relationship with Eq. 121 and the formula for relativistic time dilation, we can thereby obtain the desired relationship between L_T and L_E.

  How long does it take the pulse to travel from the back of the train, reflect on the front and go back, in the earth's frame Δt_E? Let us call the time required to get to the mirror $\Delta t_{E,1}$, and the time required to get back $\Delta t_{E,2}$, so that $\Delta t_E = \Delta t_{E,1} + \Delta t_{E,2}$.

  Considering $\Delta t_{E,1}$ first:

$$\Delta t_{E,1} = \frac{L_E + v_{TE,x}\Delta t_{E,1}}{c} \tag{122}$$

  We can rearrange to solve for $\Delta t_{E,1}$:

$$c\Delta t_{E,1} = L_E + v_{TE,x}\Delta t_{E,1} \tag{123}$$

$$\Delta t_{E,1} = \frac{L_E}{c - v_{TE,x}} \tag{124}$$

  Now we consider the time to return back:

$$\Delta t_{E,2} = \frac{L_E - v_{TE,x}\Delta t_{E,2}}{c} \tag{125}$$

  which we may rearrange to obtain:

$$\Delta t_{E,2} = \frac{L_E}{c + v_{TE,x}}. \tag{126}$$

  Thus the total time for the trip, as observed in the earth frame is:

$$\Delta t_E = \Delta t_{E,1} + \Delta t_{E,2} \tag{127}$$

$$= L_E\left(\frac{1}{c - v_{TE,x}} + \frac{1}{c + v_{TE,x}}\right) \tag{128}$$

$$= L_E\,\frac{(c + v_{TE,x} + c - v_{TE,x})}{c^2 - (v_{TE,x})^2} \tag{129}$$

$$= \frac{2cL_E}{c^2 - (v_{TE,x})^2} \tag{130}$$

Recall that we also have

$$\Delta t_T = 2\frac{L_T}{c}. \tag{131}$$

However $\Delta t_T \neq \Delta t_E$ due to time dilation. The Δt_T may be considered to be the tick of a moving clock in its rest frame; i.e., we have just reoriented the clock from Section 12. Rearranging Eq. 105 slightly gives:

$$\frac{\Delta t_E}{\Delta t_T} = \frac{1}{\sqrt{1-(v_{TE,x}/c)^2}} \tag{132}$$

We may obtain alternate expression for $\Delta t_E/\Delta t_T$ by combining Eq. 131 and Eq. 130.

$$\frac{\Delta t_E}{\Delta t_T} = \frac{c}{2L_T}\frac{2cL_E}{(c^2-(v_{TE,x})^2)} \tag{133}$$

$$= \frac{L_E}{L_T}\frac{1}{1-(v_{TE,x}/c)^2} \tag{134}$$

and equate this to Eq. 132 to obtain:

$$\frac{1}{\sqrt{1-(v_{TE,x}/c)^2}} = \frac{L_E}{L_T}\frac{1}{1-(v_{TE,x}/c)^2} \tag{135}$$

which may be rearranged to obtain L_E in terms of L_T:

$$\boxed{L_E = L_T\sqrt{1-(v_{TE,x}/c)^2}} \tag{136}$$

or $L_E = L_T/\gamma$, or in terms of rapidities $L_E = L_T/\cosh\theta$ (see footnote 15). Recall that $\gamma \geq 1$, so that the length observed in the earth's frame will always be *contracted* compared to its length in the train frame.

This is the phenomena of **length contraction**, giving rise to the slogan: *moving objects are observed to be shorter than in their rest frames.*.

- it would be nice if I could give you as dramatic illustration of length contraction as I gave in Section 13 for time dilation. But alas length contraction is not as readily manifest, at least from our normal rest frames. Objects that we observe to move at relativistic speeds are things such as electrons and protons, for which a length scale is either meaningless (electrons), or very small and not directly observable (protons).[18]

However, it is not completely hopeless. Consider the final expression that we used to compute the muon counts at sea-level:

$$N_2 = N_1 \exp\left(-\frac{1}{\tau}\frac{\Delta h_E}{v}\frac{1}{\gamma}\right) \tag{137}$$

Note that I have now made explicit that $\Delta h_E \approx 1920\,\mathrm{m}$ is a length in the *earth's* frame.

The expression for the counts, Eq. 137 can be interpreted in two different ways:

[18] Zangwill [8] (pg 830) discusses a nice example of length contraction in the "head-on" collisions of gold nuclei. It is clear that significant length contraction is important to understand the phenomena, but the connection to actual experimental observations is still a bit abstract.

(i) the time elapsed in the muon's frame is *less* than that elapsed in the earth frame, so that $\Delta t_\mu = \Delta h_E v/\gamma$. This is the viewpoint that we took in Section 13. Alternately,

(ii) the length travelled from the top of the mount to sea-level is seen as contracted in the in the muon frame, so that $\Delta h_\mu = \Delta h_E/\gamma$.

Both of these viewpoints give the same exponent in Eq. 137, but the second allows us to interpret the Rossi-Hall experiment as demonstrating length contraction — if only we could imagine ourselves travelling along with the muons at 0.9952c ☺.

15 Length contraction in electromagnetism (GITE4 12.3.1)

- from a relativistic point of view, the Lorentz force

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right) \tag{138}$$

  initially looks problematic. More specifically, we can always move to a frame where $\mathbf{v} = \mathbf{0}$, and thus there is no contribution of the $\mathbf{v} \times \mathbf{B}$ term to the force. Of course, we might expect the force to change between different inertial frames, but we cannot have acceleration in one inertial frame and none in another. As we shall see shortly — in full gory detail — the resolution is that both $\mathbf{E}$ and $\mathbf{B}$ may change in different inertial frames. The loss of $\mathbf{v} \times \mathbf{B}$ is compensated for by a change in $\mathbf{E}$.

- in the meantime, let us look at an important specific case that makes direct use of what we have learned about relativistic velocity addition and length contraction.

- two general principles will be employed in the analysis:

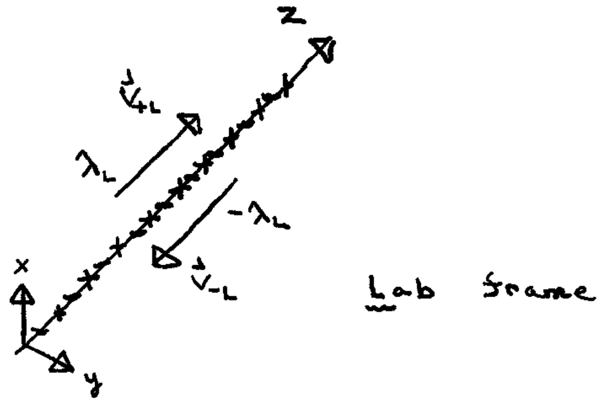
  (1) **charge invariance**: the charge on a particle (q in Eq. 138) remains unchanged between inertial reference frames. Although the invariance of charge fits "neatly" into the theory, it should be considered a postulate, supported by experimental observation.[19]

  Although we consider charge to be invariant, charge densities are certainly not: although a "clump" of charge has the same total amount of charge in different inertial frames, relativistic length contraction means that it may be more or less concentrated in volume. Thus charge densities differ between inertial frames.

  (2) **covariance**: the laws of physics "look the same in all inertial frames". In the case of electromagnetism, covariance means that the rules for computing the force on a particle (the Lorentz force law), and the rules for calculating the fields (Maxwell's equations) are precisely the same in all inertial reference frames. All the forces, charge densities, currents and fields may be different, but within any inertial reference frame, the relationships between them are the same.

[19] Jackson [9] (Section 11.9) discusses the strong experimental evidence for charge invariance.

- consider two superimposed uniform line charge densities of opposite signs, travelling in opposite directions:[20]



Since these two line charges are on top of one another, their opposite charges cancel, and thus there is no net charge density, and no $\mathbf{E}$ field.

But there is a non-zero $\mathbf{B}$ field.[21] Ampere's law gives us:

$$\mathbf{B}_L = \hat{\boldsymbol{\phi}}\, \frac{\mu_0}{2\pi} \frac{I}{s}, \tag{139}$$

where $I = 2v_{+L,z}\lambda_L$ and s is the distance to the "wire".

In what follows it is convenient to use the rest frame linear charge densities:

$$\lambda_0 = \frac{\lambda_L}{\cosh(\theta_{+L,z})}, \tag{140}$$

so that Eq. 139 may be written as:

$$\mathbf{B}_L = \hat{\boldsymbol{\phi}}\, \frac{\mu_0}{2\pi}\, \frac{1}{s}\, 2v_{+L,z} \cosh(\theta_{+L,z})\, \lambda_0. \tag{141}$$

Note that the linear charge densities are "larger" in the lab frame than in the rest frames. The charges are observed to be closer together in the lab frame — the distances between them *contract*.

- now consider the force on a point charge P located near the wire, moving parallel to the wire direction:

$$\mathbf{F}_L = q\, \mathbf{v}_{PL,z} \times \mathbf{B} \tag{142}$$

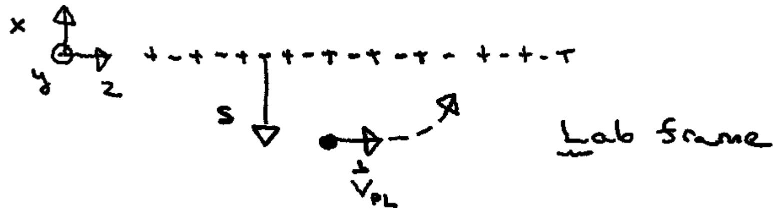
$$= -\hat{\mathbf{s}}\, q\, v_{PL,z}\, \frac{\mu_0}{2\pi}\, \frac{1}{s}\, 2v_{+L,z} \cosh(\theta_{+L,z})\, \lambda_0. \tag{143}$$

This force causes a +ve point charge to accelerate towards the wire if the charge is travelling in the same direction as the +ve line charge density.[22]

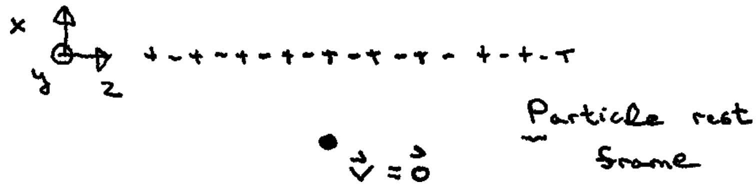
[20] I have followed the argument given in GITE4 Section 12.3.1. The setup is somewhat artificial but the same principles apply to real current-carrying wires. Purcell's original discussion [10] is more realistic.

[21] The subscript on B indicates that it is specific to the lab frame L.

[22] Two parallel wires with currents in the same direction attract; if the currents are in opposite directions, the wires repel. If you only want to remember one thing about magnetism, this should be it.



- now consider the inertial frame in which the point charge is stationary P:



At first it appears that there is no force on the charge in this frame ($\mathbf{v} = \mathbf{0} \Rightarrow \mathbf{v} \times \mathbf{B} = \mathbf{0}$), which is problematic because if the charge accelerated away from the wire in the original frame it must do so here as well — you cannot "transform away" this acceleration.

To resolve this discrepancy, consider that the two linear charge densities have different speeds in the point charge rest frame P. As a result they experience different length contractions, giving different line charge densities, and thus the densities no longer cancel. There is an electric field (!) in the rest frame of the particle, even though there was none in the lab frame.

Let us consider the details of this imperfect charge cancellation. The linear charge density of the positive charge in the frame of the point charge P is

$$\lambda_{+P} = \lambda_0 \cosh(\theta_{+P,z}), \tag{144}$$

with a similar expression for λ_{-P}.

Remember that for an infinitely long charged wire of line charge density λ, Gauss' law gives:

$$\mathbf{E} = \hat{\mathbf{s}} \frac{1}{2\pi\epsilon_0} \frac{\lambda}{s}. \tag{145}$$

Superposition of the fields due to both the positive and negative line charge densities gives:

$$\mathbf{E}_P = \hat{\mathbf{s}} \frac{1}{2\pi\epsilon_0} \frac{1}{s} \left[\lambda_0 \cosh(\theta_{+P,z}) - \lambda_0 \cosh(\theta_{-P,z})\right]. \tag{146}$$

Remember that velocity addition is easy[23] with rapidities (Eq. 70), so that:

$$\theta_{+P,z} = \theta_{+L,z} - \theta_{PL,z} \tag{147}$$
$$\theta_{-P,z} = \theta_{-L,z} - \theta_{PL,z}. \tag{148}$$

Substituting these rapidities into Eq. 146 gives:

$$\mathbf{E}_P = \hat{\mathbf{s}} \frac{1}{2\pi\epsilon_0} \frac{1}{s} \left[\lambda_0 \cosh(\theta_{+L,z} - \theta_{PL,z}) - \lambda_0 \cosh(\theta_{-L,z} - \theta_{PL,z}))\right]. \tag{149}$$

[23] It is almost unjust how much easier this derivation is with rapidities; cf. GITE4 Eq. 12.82.

Making use of the mathematical identity: $\cosh(u \pm v) = \cosh u \cosh v \pm \sinh u \sinh v$ (DLMF 4.35.2): gives:

$$\begin{aligned}\mathbf{E}_P = \hat{\mathbf{s}}\,\frac{1}{2\pi\epsilon_0}\frac{1}{s}\lambda_0\,[\,&\cosh(\theta_{+L,z})\cosh(\theta_{PL,z}) - \sinh(\theta_{+L,z})\sinh(\theta_{PL,z})\\ &- \cosh(\theta_{-L,z})\cosh(\theta_{PL,z}) + \sinh(\theta_{-L,z})\sinh(\theta_{PL,z})] \end{aligned} \tag{150}$$

Since $\theta_{-L,z} = -\theta_{+L,z}$ we may further simplify to:

$$\begin{aligned}\mathbf{E}_P = \hat{\mathbf{s}}\,\frac{1}{2\pi\epsilon_0}\frac{1}{s}\lambda_0\,[\,&\cosh(\theta_{+L,z})\cosh(\theta_{PL,z}) - \sinh(\theta_{+L,z})\sinh(\theta_{PL,z})\\ &+ \cosh(\theta_{+L,z})\cosh(\theta_{PL,z}) - \sinh(\theta_{+L,z})\sinh(\theta_{PL,z})] \end{aligned} \tag{151}$$

$$= -\hat{\mathbf{s}}\,\frac{1}{2\pi\epsilon_0}\frac{1}{s}\lambda_0\,2\sinh(\theta_{+L,z})\sinh(\theta_{PL,z})). \tag{152}$$

From $\tanh(\theta) = v/c$ we have $\sinh(\theta) = (v/c)\cosh(\theta)$, so that

$$\mathbf{E}_P = -\hat{\mathbf{s}}\,\frac{1}{2\pi\epsilon_0}\frac{1}{s}\lambda_0\,2\frac{v_{PL,z}v_{+L,z}}{c^2}\cosh(\theta_{+L,z})\cosh(\theta_{PL,z}). \tag{153}$$

Clearly there is no $\mathbf{v} \times \mathbf{B}$ force in the particle's rest frame. So the total force on the particle in its rest frame is given by $\mathbf{F}_P = q\,\mathbf{E}_P$, where $\mathbf{E}_P$ is from Eq. 153. With $c^2 = 1/(\mu_0\epsilon_0)$, the force is:

$$\mathbf{F}_P = -\hat{\mathbf{s}}\,q\,\frac{\mu_0}{2\pi}\frac{1}{s}\lambda_0\,2v_{PL,z}v_{+L,z}\cosh(\theta_{+L,z})\cosh(\theta_{PL,z}). \tag{154}$$

Apart from a factor of $\cosh(\theta_{PL,z})$, this force in the charge's frame — due to the imbalance of the line charge densities — is the same as the force in the lab frame (Eq. 143), computed using the line charge currents! Both forces cause the charge to accelerate towards the wire (for $q\lambda_0 v_{PL,z}v_{+L,z} > 0$).

In the non-relativistic limit we have $\cosh(\theta_{PL,z}) \to 1$ and thus the forces are identical $\mathbf{F}_P \to \mathbf{F}_L$. Later (GITE4, Section 12.2.4) we will see that the rules for transforming forces between inertial reference frames account for the $\cosh(\theta_{PL,z})$ factor.

Summarizing the differences between the fields in the two inertial reference frames:

in the lab frame (L)	in point charge rest frame (P)
$\mathbf{v} \times \mathbf{B} \neq \mathbf{0}$ $\mathbf{E} = \mathbf{0}$	$\mathbf{v} \times \mathbf{B} = \mathbf{0}$ $\mathbf{E} \neq \mathbf{0}$

but yet in the non-relativistic limit we have $\mathbf{F}_L = \mathbf{F}_P$.

Note that we did not need to consider $\mathbf{B}$ in the point charge rest frame, even though it is certainly non-zero. Soon we will derive the general rules for the transformation of both the $\mathbf{E}$ and $\mathbf{B}$ fields between two arbitrary inertial reference frames.

16 Lorentz transformations (GITE4 12.1.3, 12.1.4)

- the formulae that we have developed for time dilation and length contraction are appropriate when a clock or body are at rest in one inertial frame (their rest frame) and we are concerned with observations in another frame.

  We will now generalize these results to consider how arbitrary points in space-time, so-called **events**, are observed in two different inertial frames.

  For specificity, I take the relative velocity of the inertial frame B wrt inertial frame A to be $\mathbf{v}_{BA} = v_{BA,x}\hat{\mathbf{x}}$. We may always choose the orientation of the spatial axes to obtain this so-called **standard configuration**.

  We will take a single point in space-time as a "common origin" for both coordinate systems, so that

$$x_A = 0, \quad y_A = 0 \quad z_A = 0 \quad t_A = 0 \tag{155}$$

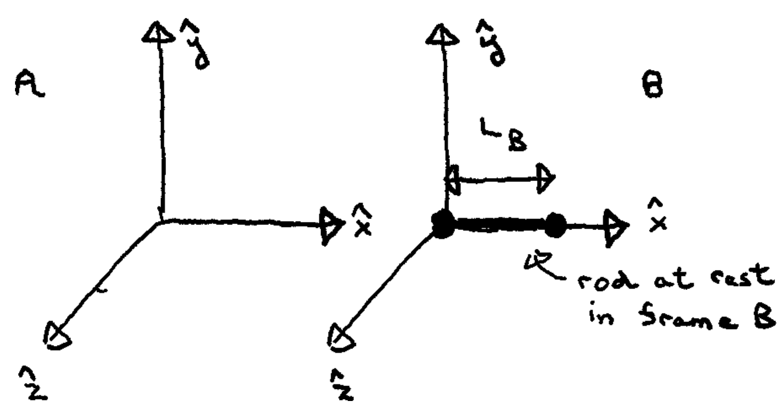
  and

$$x_B = 0, \quad y_B = 0 \quad z_B = 0 \quad t_B = 0 \tag{156}$$

  The origin of either of the systems may be shifted so that this is true.

  As was discussed in Section 9, lengths perpendicular to the relative motion of the two frames are unchanged, so that $y_A = y_B$ and $z_A = z_B$, and thus we only need to concentrate on the relationship between the coordinates in A: x_A, t_A and the coordinates in B: x_B, t_B.

  Consider a rod at rest in frame B, which is moving wrt to frame A:



  By length contraction: $L_A = L_B/\gamma = L_B/\cosh\theta_{BA,x}$ and thus:

$$x_A = v_{BA,x}t_A + \frac{x_B}{\cosh\theta_{BA,x}} \tag{157}$$

  I prefer rapidities, and thus write $v_{BA,x} = c\tanh\theta_{BA,x}$, so that after rearranging:

$$\begin{aligned} x_B &= \cosh\theta_{BA,x}\,(x_A - \tanh(\theta_{BA,x})\,ct_A) && (158)\\ &= \cosh(\theta_{BA,x})\,x_A - \sinh(\theta_{BA,x})\,ct_A. && (159)\end{aligned}$$

This relationship is what we want: a coordinate in B in terms of the coordinates in A. We still need to find the other "half" of the transformation: t_B in terms of x_A and t_A.

For that purpose, note that neither A or B is "preferred" in any way, so that we may swap them to obtain:

$$x_A = \cosh(\theta_{AB,x})\, x_B - \sinh(\theta_{AB,x})\, ct_A. \tag{160}$$

Both frames must agree on the magnitude of their relative velocities, but the direction is reversed. Specifically, $v_{BA,x} = -v_{AB,x}$ and thus $\theta_{AB,x} = -\theta_{BA,x}$, so that:

$$x_A = \cosh(\theta_{BA,x})\, x_B + \sinh(\theta_{BA,x})\, ct_B. \tag{161}$$

This equation may be combined with 159, by rearranging it to solve for x_B. Then we will have the "second-half" of the transformation: a relationship between x_A, t_A and t_B. Following this plan, Eq. 161 may be rearranged to obtain:

$$x_B = \frac{1}{\cosh\theta_{BA,x}}\left(x_A - \sinh(\theta_{BA,x})\, ct_B\right) \tag{162}$$

and equated to Eq. 159 to give:

$$\frac{1}{\cosh\theta_{BA,x}}\left(x_A - \sinh(\theta_{BA,x})\, ct_B\right) = \cosh(\theta_{BA,x})\, x_A - \sinh(\theta_{BA,x})\, ct_A. \tag{163}$$

Rearranging to solve for t_B gives:

$$ct_B = \frac{1}{-\sinh(\theta_{BA,x})}\left(\cosh^2(\theta_{BA,x})\, x_A - \sinh(\theta_{BA,x})\cosh(\theta_{BA,x})\, ct_A - x_A\right). \tag{164}$$

Recall the identity: $\cosh^2\theta - \sinh^2\theta = 1$, allowing $\cosh^2(\theta_{BA,x})$ to be rewritten so that:

$$ct_B = \frac{1}{-\sinh(\theta_{BA,x})}\left[\left(1 + \sinh^2\theta_{BA,x}\right)\ x_A - \sinh(\theta_{BA,x})\cosh(\theta_{BA,x})\, ct_A - x_A\right] \tag{165}$$

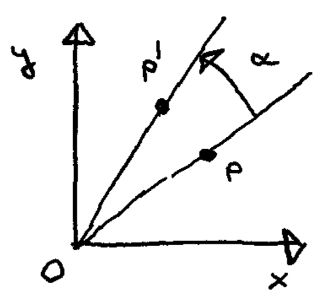
and then simplified to give:

$$ct_B = -\sinh(\theta_{BA,x})\, x_A + \cosh(\theta_{BA,x})\, ct_A. \tag{166}$$

Equations 159 and 166 may be summarized in matrix form:

$$\begin{bmatrix} ct_B \\ x_B \end{bmatrix} = \begin{bmatrix} \cosh\theta_{BA,x} & -\sinh\theta_{BA,x} \\ -\sinh\theta_{BA,x} & \cosh\theta_{BA,x} \end{bmatrix} \begin{bmatrix} ct_A \\ x_A \end{bmatrix}. \tag{167}$$

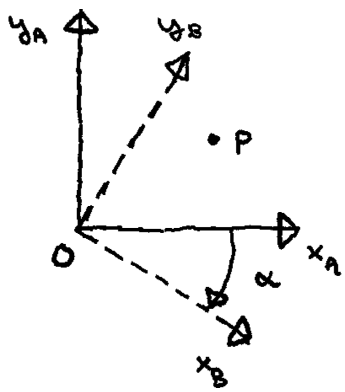
The form of this **Lorentz transformation** may remind you of a similar formula for rotations about the origin in two spatial dimensions:



$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \tag{168}$$

where α corresponds to a clockwise rotation of the point about the origin.

We may strengthen the analogy between these rotations in the plane and the Lorentz transformations by making a distinction between two types of transformations: active and passive. In Eq. 168 we rotated a point, moving it from P to P', but kept the same coordinate system — a so-called **active transformation**. On the other hand, before and after the Lorentz transforomation of Eq. 167 we are referring to the *same* point in space-time. It is the *coordinate* system (the inertial frame) that is changing. That is known as a **passive transformation**. The two-dimensional analog is a rotation of the coordinate system, while keeping the location of the point P constant:



$$\begin{bmatrix} x_B \\ y_B \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x_A \\ y_A \end{bmatrix}, \tag{169}$$

where we write the coordinates in the new system B in terms of the old system A using the same form as in the active transformation, but now $\alpha > 0$ corresponds to anti-clockwise rotations of the B *coordinate system* wrt the A *coordinate system*.

Our Lorentz transformations are *passive* transformations: we consider the space-time point to be fixed, but we want to know its coordinates in a different inertial reference frame.[24]

Let us explore the analogy between rotations in two spatial dimensions and Lorentz transformations a bit further. More specifically, rotations preserve distances from the origin; or equivalently:

$$x_A^2 + y_A^2 = x_B^2 + y_B^2, \tag{170}$$

which may be verified using the form of the transformation given by Eq. 169.

An analogous relationship holds for the Lorentz transformations. A small amount of algebra, based on Eq. 167, shows that:

$$-(ct_A)^2 + x_A^2 = -(ct_B)^2 + x_B^2. \tag{171}$$

Thus $-(ct)^2 + x^2$ is a Lorentz invariant.[25]

[24] The distinction between active and passive transformations is particularly important in the quantum theory of angular momentum.

[25] By consideration of rotations into and out of the “standard configuration” so that $y_A = y_B$ and $z_A = z_B$, we may conclude, more generally, that $-(ct)^2 + x^2 + y^2 + z^2$ is a Lorentz invariant. I will omit the two orthogonal coordinates when they are not apropos, their presence being understood as implicit.

Actually, we can say a bit more than this. Let us return to rotations in two dimensions and name the 2×2 matrix which appears on the LHS of Eq. 169 as $R(\alpha)$. Then computing $x_B^2+y_B^2$ in terms of x_A and x_B can be written as:

$$x_B^2 + y_B^2 = \begin{bmatrix} x_B & y_B \end{bmatrix} \begin{bmatrix} x_B \\ y_B \end{bmatrix} \tag{172}$$

$$= \begin{bmatrix} x_B \\ y_B \end{bmatrix}^T \begin{bmatrix} x_B \\ y_B \end{bmatrix} \tag{173}$$

$$= \left(R(\alpha) \begin{bmatrix} x_A \\ y_A \end{bmatrix} \right)^T \left(R(\alpha) \begin{bmatrix} x_A \\ y_A \end{bmatrix} \right). \tag{174}$$

With some elementary results of matrix algebra (the transpose of matrix products and associativity of matrix multiplication):

$$x_B^2 + y_B^2 = \begin{bmatrix} x_A & y_A \end{bmatrix} R(\alpha)^T R(\alpha) \begin{bmatrix} x_A \\ y_A \end{bmatrix}. \tag{175}$$

From the explicit form of $R(\alpha)$ given in Eq. 169 you may verify that $R(\alpha)^T R(\alpha) = \mathbb{1}$, the 2×2 identity matrix, thereby confirming that $x^2 + y^2$ is invariant under rotations about the origin. But the particular set of steps that we have taken makes a generalization obvious: if we consider two *different* points in space, labelled by 1 and 2, then their "dot product", $\mathbf{r}_1 \cdot \mathbf{r}_2$, is also invariant under rotations:

$$x_{1,A}x_{2,A} + y_{1,A}y_{2,A} = x_{1,B}x_{2,B} + y_{1,B}y_{2,B}. \tag{176}$$

Back to SR. The **Lorentz invariant scalar product**[26] of two space-time events, labelled as 1 and 2, is given by:

$$-ct_{1,A}\, ct_{2,A} + x_{1,A}x_{2,A} = \begin{bmatrix} ct_{1,A} & x_{1,A} \end{bmatrix} \begin{bmatrix} -1 & \\ & 1 \end{bmatrix} \begin{bmatrix} ct_{2,A} \\ x_{2,A} \end{bmatrix}. \tag{177}$$

Let us name the 2×2 transformation matrix in Eq. 167 as $L(\theta)$. By exactly the same matrix manipulations as for the rotation case, we may verify that

$$L(\theta)^T \begin{bmatrix} -1 & \\ & 1 \end{bmatrix} L(\theta) = \begin{bmatrix} -1 & \\ & 1 \end{bmatrix}, \tag{178}$$

independent of the value of θ, so that we conclude that for two space-time events, 1 and 2, the scalar $-ct_1\, ct_2 + x_1\, x_2$ is a Lorentz invariant. This result, and its generalizations, will be very important in our study of electricity and magnetism using SR.

- Suggested problem: show that $L(\theta)^{-1} = L(-\theta)$.

[26] GITE4 calls it the **four-dimensional scalar product**.

17 Four-vectors and tensors (GITE4 12.1.4)

- in this section I will provide a physical example of the Lorentz invariant scalar product, using it as a starting point for an introduction to four-vectors and tensors. A more expedient approach is taken by GITE4; you might prefer that treatment.

- it is straightforward to verify that the fields

$$\mathbf{E}(x,t) = E_0 \cos(kx - \omega t)\hat{\mathbf{y}} \quad \text{and} \quad \mathbf{B}(x,t) = \frac{1}{c} E_0 \cos(kx - \omega t)\hat{\mathbf{z}} \tag{179}$$

  satisfy Maxwell's equations in vacuum ($\rho = 0$ and $\mathbf{J} = \mathbf{0}$), where $\omega/k = c = 1/\sqrt{\mu_0 \epsilon_0}$.

  These solutions correspond what is known as a "linearly polarized plane electromagnetic wave", which we will discuss in much more detail later (the particular form of Eq.'s 179 correspond to Eq. 9.48 of GITE4, but with propagation in the x instead of z direction).

  For now let us concentrate on the "phase", $\phi = kx - \omega t$. We have already discussed this phase briefly in Section 10, in the context of phase velocities.

  The common phase in Eq. 179 is considered to be a Lorentz invariant. Why? Jackson says (on pg 519 of [9]): *The phase of a plane wave is an invariant quantity, the same in all coordinate frames. This is because the elapsed phase of a wave is proportional to the number of wave crests that have passed the observer. Since this is merely a counting operation, it must be independent of coordinate frame.*[27]

- let us look at the consequences of the Lorentz invariance of phase. Consider two inertial frames A and B, with a relative velocity $\mathbf{v}_{BA} = v_{BA,x}\hat{\mathbf{x}}$, which I will represent using a rapidity θ, such that $\tanh\theta = v_{BA,x}/c$. We Lorentz transform the x and t (in the standard configuration), writing the B coordinates in terms of the A coordinates, using Eq. 167, so that the phase may be written as

$$\phi = kx_B - \frac{\omega}{c}\, ct_B \tag{180}$$

$$= k\left(x_A \cosh\theta - ct_A \sinh\theta\right) - \frac{\omega}{c}\left(-x_A \sinh\theta + ct_A \cosh\theta\right). \tag{181}$$

  Slightly rearranging gives:

$$\phi = \left(k\cosh\theta + \frac{\omega}{c}\sinh\theta\right) x_A - \left(k\sinh\theta + \frac{\omega}{c}\cosh\theta\right) ct_A, \tag{182}$$

  suggesting that we consider both k and ω to be frame specific; i.e., write k as k_B and ω and ω_B so that the phase invariance condition can be written as

$$k_B x_B - \frac{\omega_B}{c} ct_B = k_A x_A - \frac{\omega_A}{c} ct_A, \tag{183}$$

[27] Despite its pedigree I must admit that I do not find this argument entirely compelling. For one thing, it does not rule out a constant phase difference. Perhaps a variant of an argument used in Section 15 removes my objection: during the zeros of the **E** and **B** field, a test particle will not accelerate. At all other phases there will be acceleration. The presence or absence of acceleration is a Lorentz invariant concept — acceleration can not be "transformed away" (as emphasized in Section 15). The invariance of these zeros of acceleration removes the possibility of a constant phase difference between inertial frames.

where by Eq. 182:

$$k_A = k_B \cosh\theta + \frac{\omega_B}{c}\sinh\theta, \text{ and} \tag{184}$$

$$\frac{\omega_A}{c} = k_B \sinh\theta + \frac{\omega_B}{c}\cosh\theta. \tag{185}$$

This an "inverse" Lorentz transformation that we may "invert" to obtain:

$$k_B = k_A \cosh\theta - \frac{\omega_A}{c}\sinh\theta, \text{ and} \tag{186}$$

$$\frac{\omega_B}{c} = -k_A \sinh\theta + \frac{\omega_A}{c}\cosh\theta. \tag{187}$$

Thus k_A and ω_A/c transform to k_B and ω_B/c in precisely the same way that x_A and ct_A transform to x_B and ct_B. This result generalizes to three spatial dimensions: the four components of ω/c and $\mathbf{k}$ transform in *precisely* the same way as the same four components of ct and $\mathbf{r}$. BTW: the transformations of Eq. 186 and 187 describe the Doppler effect.

As we proceed, we shall find many more instances of four entities that transform similarly, suggesting new terminology: we will call *anything* that transforms like $(ct, \mathbf{r})$ to be a **four-vector**. We have just shown that $(\omega/c, \mathbf{k})$ is a four-vector.

For obvious reasons, the first component of any four-vector is known as the **time-like** component, and the remaining components are known as the **space-like** components.

- the Lorentz invariant scalar product of Eq. 177 may be generalized, so as to apply to *any* two four-vectors. It is useful to define the symbol " $\diamondsuit$ " so that[28]

$$(ct_1, \mathbf{r}_1) \diamondsuit (ct_2, \mathbf{r}_2) \coloneqq -ct_1\, ct_2 + x_1x_2 + y_1y_2 + z_1z_2. \tag{188}$$

For example the phase of Eq. 179 is actually a Lorentz invariant scalar product between the $(\omega/c, \mathbf{k})$ and $(ct, \mathbf{r})$ four-vectors:

$$\phi = (\omega/c, \mathbf{k}) \diamondsuit (ct, \mathbf{r}) \tag{189}$$

$$= \mathbf{k}\cdot\mathbf{r} - \frac{\omega}{c}\, ct. \tag{190}$$

There is a slightly different way to view the Lorentz invariant scalar product:

$$(\omega/c, \mathbf{k}) \diamondsuit (ct, \mathbf{r}) = \underbrace{\begin{bmatrix}\omega/c & k_x & k_y & k_z\end{bmatrix} \overbrace{\begin{bmatrix}-1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1\end{bmatrix}}^{=:g}}_{1} \underbrace{\begin{bmatrix}ct \\ x \\ y \\ z\end{bmatrix}}_{2}. \tag{191}$$

[28] It is tempting to just reuse the "dot" of the dot product to represent the Lorentz invariant scalar product — like Taylor [11] and Jackson [9] do. But I have found — at least when teaching — that this reuse creates problems. Usage of the $\diamondsuit$ symbol is suggested in the Feynman lectures — the four-edges/corners remind us of its four-dimensional character.

The four components of (1) are known as the **covariant**[29] components of the four-vector $(\omega/c, \mathbf{k})$, and we write them as $k_0 = -\omega/c$, $r_1 = k_x$, $r_2 = k_y$, $r_3 = k_z$.

The four components of (2) are known as the **contravariant** components of the four-vector $(ct, \mathbf{r})$, and we write them as $r^0 = ct$, $r^1 = x$, $r^2 = y$, and $r^3 = z$.

With these conventions, the Lorentz invariant scalar product may be condensed to:

$$(\omega/c, \mathbf{k}_1) \diamondsuit (ct, \mathbf{r}) = \sum_{\alpha=0,1,2,3} k_\alpha r^\alpha. \tag{192}$$

It turns out to be incredibly convenient to adopt the convention that whenever a greek letter index is repeated, once as an upper index, and once as a lower index, the existence of a sum over that index is to be understood. For this example:

$$k_\alpha r^\alpha \qquad \text{means} \qquad \sum_{\alpha=0,1,2,3} k_\alpha r^\alpha. \tag{193}$$

This convention is known as the **(Einstein) summation convention**.

- note that any four-vector can be described by either its covariant or contravariant components. By (1) of Eq. 191 we may formally write the covariant components of any four-vector described by its contravariant components (F^0, F^1, F^2, F^3) as:

  $$F_\alpha = g_{\alpha\beta} F^\beta \tag{194}$$

  where $g_{00} = -1$, $g_{11} = g_{22} = g_{33} = 1$, and all other entries are zero (see Eq. 191), and I have employed the summation convention; i.e., β is being summed over on the RHS.

  Informally, we just "reverse the sign" of the time-like component to switch back and forth between the co- and contravariant components.

  From these rules we can verify that $k_\alpha r^\alpha = k^\alpha r_\alpha$.

- the components of four-vectors do not necessarily have to be numbers. For example, consider the Lorentz invariant phase "field": $\phi = \mathbf{k} \cdot \mathbf{r} - \omega t$. To get the contravariant components of the $(\omega/c, \mathbf{k})$ four-vector we may differentiate this phase field:

  $$k^0 = -\frac{1}{c}\frac{\partial \phi}{\partial t}, \quad k^1 = \frac{\partial \phi}{\partial x}, \quad k^2 = \frac{\partial \phi}{\partial y}, \quad k^3 = \frac{\partial \phi}{\partial z} \tag{195}$$

  which we write in the shorthand:

  $$k^\alpha = \partial^\alpha \phi. \tag{196}$$

  Since we may use this relation to obtain the $(\omega/c, \mathbf{k})$ four-vector in any inertial frame, the four ∂^α entities must transform like a four-vector, and thus we say that these are the four components of a contravariant "gradient" four-vector.

[29] The "covariant component" terminology should not be confused with the "principle of covariance" discussed earlier in Section 15. There is a loose connection, but the reuse of "covar-" is mostly just an unhelpful overloading.

By reversing the sign of the time-like component of the gradient four-vector we may obtain its covariant components:

$$\partial_0 = \frac{1}{c}\frac{\partial}{\partial t}, \quad \partial_1 = \frac{\partial}{\partial x}, \quad \partial_2 = \frac{\partial}{\partial y}, \quad \partial_3 = \frac{\partial}{\partial z}. \tag{197}$$

Note the absence of the negative sign in the time-like covariant component, which differs from both $(ct, \mathbf{r})$ and $(\omega/c, \mathbf{k})$.

- four-vectors and Lorentz scalars are just instances of the more general idea of a **(Lorentz) tensor**. While the components of a four-vector are specified by a single index, there may be an arbitrary number of indices involved in specifying the components of a tensor; e.g., $T^{\alpha\beta}$ refers to a single component of a tensor entity T with $4 \times 4 = 16$ components. In the same way that we could refer to the contravariant and covariant of a four-vector, we may also specify T using purely covariant components $T_{\alpha\beta}$, or mixed components $T^{\alpha}{}_{\beta}$ and $T_{\alpha}{}^{\beta}$. We may convert between these forms using g; for example $T_{\alpha}{}^{\beta} = g_{\alpha\nu}T^{\nu\beta}$. The number of upper and lower components may be written in a tuple-like form and is referred to as the **valence**; e.g., $T_{\alpha\beta}$ has a valence of $(0, 2)$, $T_{\alpha}{}^{\beta}$ has a valence of $(1, 1)$, scalars have a valence of $(0, 0)$. The total number of indices is a tensor's **rank**.[30]

  Recall that an arbitrary set of four numbers is not necessarily a four-vector. It has to follow the same Lorentz transformation rules as $(ct, \mathbf{r})$. Similarly for tensors — their components must follow certain transformation rules in order to earn the tensor "badge".

  Following GITE4, I use the 4×4 matrix Λ to represent a Lorentz transformation between two inertial reference frames, generalizing the $L(\theta)$ of Section 16, to any combination of rotations and so-called **boosts**: transformations between inertial frames with the same axes orientation, but with different relative velocities, such as discussed in Section 16. In analogy with Eq. 178, every Λ satisfies

$$\Lambda^T g \Lambda = g. \tag{198}$$

  By definition, a four-vector transforms like:

$$\tilde{H}^{\alpha} = \Lambda^{\alpha}_{\beta}\, H^{\beta}, \tag{199}$$

  where ˜ designates the components of the "new" system, and the absence of ˜'s indicates the "old" system.[31] The upper index corresponds to a row and the lower index to a column of Λ. For example, if B is the new system and A is the old system, then the pure boost $\mathbf{v}_{BA} = v_{BA,x}\hat{\mathbf{x}}$, with $\tanh\theta = v_{BA,x}/c$, has

$$\Lambda = \begin{bmatrix} \cosh\theta & -\sinh\theta & 0 & 0 \\ -\sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{200}$$

[30] "Rank" is yet another overloaded term; its usage here relates to, but differs from the "rank of a linear transformation".

[31] One slight disadvantage of the tensor notation is that we can no longer use subscript labels, but instead must use symbols like ˜ and ′ to distinguish between quantities in different inertial reference frames.

For the scalar product between two four-vectors to be invariant, covariant components must transform like:

$$\tilde{G}_\alpha = \Lambda^{-1}{}^{\gamma}_{\ \alpha}\, G_\gamma, \tag{201}$$

as this ensures

$$G_\alpha H^\alpha = \tilde{G}_\alpha \tilde{H}^\alpha, \tag{202}$$

since (by definition of the inverse)

$$\Lambda^{-1}{}^{\gamma}_{\ \alpha}\Lambda^\alpha_\beta = \delta^\gamma_\beta, \tag{203}$$

where δ^γ_β is the **Kronecker delta**, equal to one if $\gamma = \beta$, and zero otherwise.

Generalizing these transformation rules for four-vectors: tensor components that are purely contravariant $(\ldots, 0)$ must transform like:

$$\tilde{T}^{\alpha\beta} = \Lambda^\alpha_\nu \Lambda^\beta_\epsilon T^{\nu\epsilon}. \tag{204}$$

Likewise, tensor components that are purely covariant $(0, \ldots)$ must transform like:

$$\tilde{T}_{\alpha\beta} = \Lambda^{-1}{}^{\nu}_{\ \alpha}\Lambda^{-1}{}^{\epsilon}_{\ \beta} T_{\nu\epsilon}. \tag{205}$$

And mixed valence tensors should transform using a mixture of the elements of Λ and Λ^{-1}. For example

$$T^\alpha_\beta = \Lambda^\alpha_\nu \Lambda^{-1}{}^{\epsilon}_{\ \beta} T^\nu_\epsilon. \tag{206}$$

The generalization of all of these results to tensor components of arbitrary valence should be clear.

- it should be emphasized that the Λ's themselves are not tensors. It would be somewhat meaningless to discuss how they transform between frames — since a Λ describes the transformation between two frames.
- rarely do we have to test for tensorial character using the transformation rules directly. Instead there are rules by which new tensors may be constructed from objects that we already know to be tensors. These rules are readily established from the properties of g and the allowable Λ's.

For example, if G^α and H^β are the contravariant components of two four-vectors, then

$$T^{\alpha\beta} = G^\alpha H^\beta \tag{207}$$

are the components of a second-rank tensor, as are

$$T_\alpha{}^\beta = G_\alpha H^\beta. \tag{208}$$

And if $T_\alpha{}^\beta$ are the components of a tensor, then by **contraction**:

$$T_\alpha{}^\alpha \tag{209}$$

is a Lorentz scalar (a zero rank tensor).

Any mixed valence tensor can be contracted by summing over a common index in both a contravariant slot and covariant **slot** (index location). Given that we may convert contravariant slots to covariant slots and vice versa using g, tensors may often be contracted in multiple different ways to give tensors of lower rank, including of course, scalars.

- some final peripheral points regarding four-vectors and tensors:

  (1) A more "geometric" approach introduces special symbols for the four-vectors and their tensor generalizations, avoiding explicit component representations as much as possible. Thorne and Blandford [12] describe this philosophy and Steane [13] shows just how long the use of tensor components can be deferred in SR.

  (2) In general treatments of tensor fields [14], Eq. 201 — or more precisely its generalization (*vide infra*) — is sometimes used to *define* what is known as a "covariant vectors", and Eq. 199 helps define "contravariant vectors". We have been referring to four-vectors, and the simultaneous existence of their covariant and contravariant components. It is the **metric** g that permits us to refer to (unqualified) four-vectors, and freely convert between their covariant and contravariant component representations. When there is no metric, covariant and contravariant vectors are distinct entities.

  (3) I like to remember the components of Λ and Λ^{-1} in their more general forms:

$$\Lambda^{\alpha}_{\beta} = \frac{\partial \tilde{r}^{\alpha}}{\partial r^{\beta}} \quad \text{and} \quad {\Lambda^{-1}}^{\alpha}_{\beta} = \frac{\partial r^{\alpha}}{\partial \tilde{r}^{\beta}}. \tag{210}$$

  For arbitrary coordinate transformations — not just Lorentz transformations — these derivatives may depend on *location*; the matrix Λ is no longer a constant. That dependence complicates things. For example, the gradient of Eq. 197 is no longer a rank one tensor. These complications are a necessary and intrinsic part of the general theory of relativity, which we do not consider in this course.

18 Proper time and the terrible twins (GITE4 12.2.1)

- by itself, the Lorentz invariant $r^{\alpha} r_{\alpha}$ is not especially useful, as it depends on the origin of the coordinate system; i.e., the location in space-time of $ct = x = y = z = 0$.

  A more useful invariant, the **interval**, is defined by the *differences* between two space-time events:[32]

$$I := -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2, \tag{211}$$

  where Δ's are the differences in the respective coordinates for the two space-time points.

  Let us take a small detour for notation. In adopting tensor notation and the summation convention we have given up — or at least compromised — our ability to use subscript labels.

[32] Here is a good place to mention a convention. Since I is invariant, so is $-I$. The choice of $-I$ as the interval is consistent with an alternate convention (from these notes and GITE4) in which the metric is $g_{00} = 1$, $g_{11} = g_{22} = g_{33} = -1$; i.e., to go between co- and contravariant components with this metric the space-like components should be reversed instead of the time-like component. This alternate convention is used by Landau and Lifshitz [2] and others. But the difference between these metrics is not important — in the same way that it does not really matter what side of the road *we* drive on. But as Zare notes [15] — in a similar context — woe betide those who flout local customs!

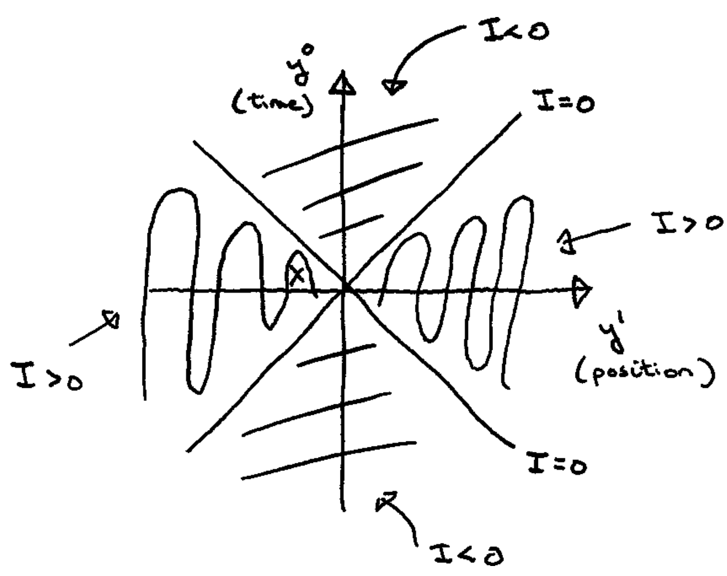
Consequently, it is beneficial to switch from using r^α for the components of a "position" four-vector, and instead use x^α for those same components and when necessary use y^α and z^α in a similar way, to label *different* points in space-time.

In this notation, the interval between two space-time events x and y can be written as[33]

$$I = (x_\alpha - y_\alpha)(x^\alpha - y^\alpha), \tag{213}$$

from which it is straightforward to see that the interval is Lorentz invariant. Alternately — and more intuitively — we can simply place the origin at one of the events, say y. Then invariance of the interval follows from the invariance of $x_\alpha x^\alpha$.

- the sign of the interval between two space-time points has a special significance: it establishes constraints on the signalling between two events. If Δd is the spatial distance between two events, writing the interval as $I = (\Delta d)^2 - (c\Delta t)^2$ allows us to see that for $I > 0$, a signal travelling between the two events would need to have a speed greater than the speed of light c. We call $I > 0$ a **space-like** separation. With some mild extrapolation: events with space-like separations cannot have any direct causal relation; i.e., event x cannot *cause* event y or vice-versa. On the other hand, if $I < 0$, so-called **time-like** separation, then the events *may* be causally linked, since a signal with a speed less than c *may* have travelled from one event to the other. With $I < 0$ it is impossible for a Lorentz transformation to change the time ordering of the two events.[34] Finally, two events with $I = 0$ are said to have a **light-like** separation.
- for motion in one spatial dimension, so-called **space-time diagrams** are quite useful. Placing one event at the origin (x below), we may label four distinct quadrants by the sign of I, their boundaries being light-like:

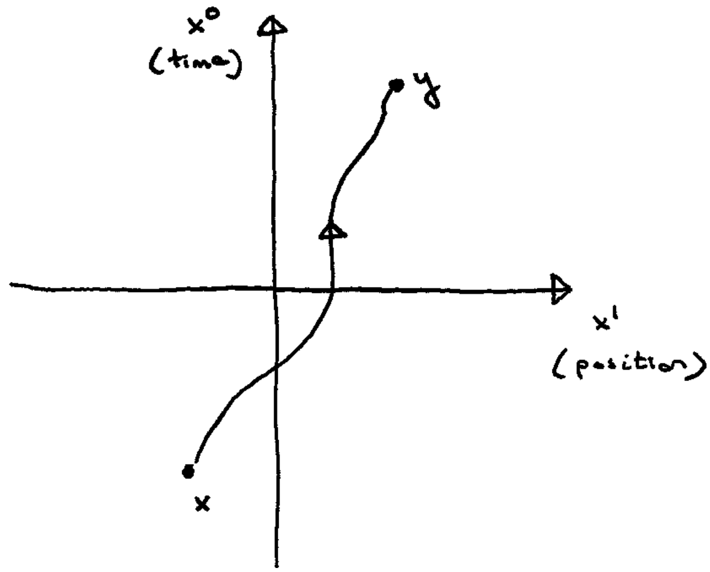


[33] Sometimes when I do not need to label the two points explicitly, I will just write

$$I = (\Delta x_\alpha)(\Delta x^\alpha). \tag{212}$$

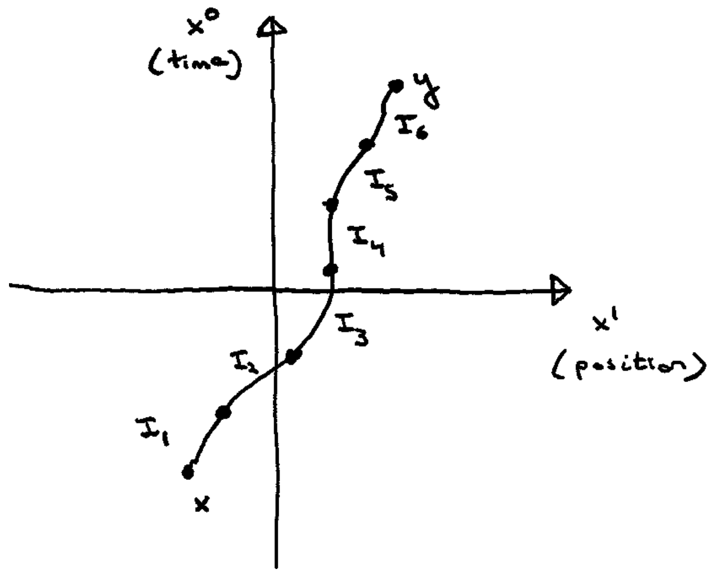
[34] Take that, Ted Chiang.

- now consider a particle undergoing a motion starting at one space-time point and ending at another:



As we have just discussed, the interval between the start and end events is an invariant. This interval does not depend on the specifics of the trajectory, the so-called **world-line** of the particle.

There are other Lorentz invariants that do depend on the world-line and not just its end-points. Consider the world-line to be broken up into a set of connected events, with Lorentz invariant intervals separating them:



Specifically, $\sum_i I_i$ is a Lorentz invariant (now i labels each segment; it is not a tensor index). However its value depends on where we choose to place the intermediate events along the world-line. But with some small tweaks we may construct an invariant characterizing the world-line that removes this arbitrariness.

Instead of summing I_i for each segment connecting events along the world-line, consider instead summing $\sqrt{-I_i}$:

$$J := \sum_i \sqrt{-I_i} \tag{214}$$

$$= \sum_i \sqrt{(c\Delta t_i)^2 - (\Delta d_i)^2} \tag{215}$$

$$= \sum_i c\Delta t_i \sqrt{1 - \left(\frac{\Delta d_i}{c\Delta t_i}\right)^2}, \tag{216}$$

suggesting conversion to a Riemann integral by letting $\Delta t_i \to 0$ with insertion of more intermediate events along the world-line:

$$J = \lim_{\Delta t_i \to 0} \sum_i c\Delta t_i \sqrt{1 - \left(\frac{\Delta d_i}{c\Delta t_i}\right)^2}. \tag{217}$$

Recognizing that in the same limit, $\Delta d_i/\Delta t_i \to v$, we have:

$$J = c\int_{x^0/c}^{y^0/c} dt\, \sqrt{1 - (v/c)^2}. \tag{218}$$

where in general v will change as we move along the world line. We now see why the negative sign was used in $\sqrt{-I_i}$ — it avoids a pesky $\sqrt{-1}$, assuming $v < c$.

By construction J is a Lorentz invariant; i.e., if we Lorentz transform the coordinates of the world-line to another frame, denoted by $\tilde{}$, then

$$\int_{\tilde{x}^0/c}^{\tilde{y}^0/c} d\tilde{t}\, \sqrt{1 - (\tilde{v}/c)^2} = \int_{x^0/c}^{y^0/c} dt\, \sqrt{1 - (v/c)^2}. \tag{219}$$

There is a nice physical interpretation of this invariant. Remember that "moving clocks run slow" (Section 12), so that in their rest frame R less time elapses than in the frame moving at v wrt the rest frame:

$$\Delta t_R = \Delta t\sqrt{1 - (v/c)^2}. \tag{220}$$

Recognizing this expression in Eq. 218, the integration now looks like a summation to obtain how much the time will elapse for a clock following this world-line. i.e., we are accumulating the instantaneous rest frame time elapsed. This time is known as the **proper time**, and it is customary to write it as

$$\boxed{\tau = \int_{x^0/c}^{y^0/c} dt\, \sqrt{1 - (v/c)^2}} \tag{221}$$

- the proper time for a world-line connecting two space-time points is maximal for a constant velocity trajectory and less for any other trajectory that (necessarily) involves acceleration. The proof is straightforward: since we are computing two proper times (that are necessarily

Lorentz invariant) it does not matter which inertial reference frame we compute the proper times in. Thus we may choose a reference frame in which the "straight line" between the start and end events is transformed to "no motion" at all; i.e., $v = 0$. In this reference frame we are comparing the proper time of the constant velocity world-line:

$$\tau_{\text{uniform}} = \int_{x^0/c}^{y^0/c} dt \tag{222}$$

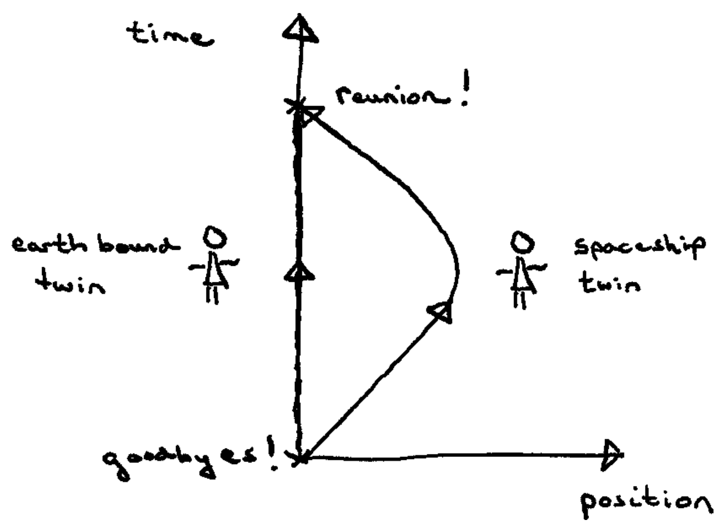
to the proper time of any other world-line which involves acceleration

$$\tau_{\text{non-uniform}} = \int_{x^0/c}^{y^0/c} dt\,\sqrt{1-(v/c)^2}. \tag{223}$$

Since $\sqrt{1-(v/c)^2} < 1$ for at least part of the non-uniform motion world-line (since it involves acceleration), the integral for second world-line is always less than the first. Thus $\tau_{\text{non-uniform}} < \tau_{\text{uniform}}$; the proper time is maximal for a constant velocity trajectory, and less for any other trajectory $\square$.

The fact that — in the absence of forces — a particle travels along world-lines that connect space-time points in such a way that proper time is maximal is sometimes known as the "principle of most proper time". This principle may remind you of Lagrangian mechanics. Hold that thought — it is going to be important.

- with proper time, the antics of the notorious twins are less opaque:



If we compare the twin's proper times at their reunion, the twin who stays at home will always be older than the spaceship-travelling twin.

No Lorentz transformation can "swap" the twins; i.e., the Lorentz transformation maps straight lines on space-time diagrams to straight lines. You cannot take the "head back home" bend out of the spaceship world-line.

- again the twin paradox seems so unusual that some sort of experimental evidence is at least mildly comforting. One famous test [16] compared the decay of muons travelling around a circular accelerator ring and found their lifetime was extended significantly compared to muons that "stayed at home" (64 μs compared to 2.2 μs), in excellent agreement with SR.

19 Momenergy from an invariant action (GITE4 12.2.2)

- introductions to special relativity typically motivate the definition of relativistic momentum by examination of a collision between two particles in different inertial reference frames. By the principle of covariance (Section 15), the equations describing momentum conservation should have the same form in all inertial reference frames. With varying degrees of rigour, this covariance constraint is used to determine an expression for relativistic momentum in terms of previously defined quantities.

  The naturally occurring "time-like" component accompanying the three spatial components of relativistic momentum is then identified with energy, as its non-relativistic limit coincides with the normal non-relativistic kinetic energy (modulo the infamous mc^2). As with the relativistic momentum this identification has to be backed up by experimental evidence.

  GITE4 essentially follows this strategy; i.e., Problem 12.29 is used to justify Eq. (12.46). Here I will follow Landau and Lifshitz's variational approach [2], partly for variety, partly because it gets us to the rules for transforming the electric and magnetic fields between inertial frames in a different way. But I also recommend reviewing GITE4's section 12.2.1 and 12.2.2, especially the problems.

- we will start by considering a free particle; electric and magnetic fields will be considered later, building upon the free particle case. Based on the principle of most proper time, we hypothesize that the action[35] for a world-line connecting a start and end event is of the form:

$$S = -\alpha \int d\tau \tag{224}$$

  where $\alpha > 0$, but is otherwise to be determined. In what follows, the two space-time events that define the limits of integrals such as these will be omitted but should be understood to be present. As written, the action is an Lorentz invariant, so we expect that it will lead to the same world-line, independent of which specific inertial frame that we apply the Lagrangian procedure in.[36]

  Choosing a specific inertial reference frame:

$$S = -\alpha \int dt \sqrt{1-(v/c)^2}. \tag{225}$$

[35] The **action** is the time integral of the Lagrangian over a trajectory; see Appendix B. Since the action may be computed for different world-lines connecting the same start and end events, it may be considered to be a device which takes a function(s) and spits out a number — a so-called **functional**.

[36] If a trajectory minimizes an *invariant* action then it is clear that this trajectory corresponds to a world-line which minimizes the action in *all* inertial frames. Unfortunately, the simplicity (but not the outcome) of this argument is spoiled by the fact that the "principle of least action" is a misnomer — it really should be called the "principle of stationary action". See this nice example.

Since by definition $S = \int dt\mathcal{L}$, we have

$$\mathcal{L} = -\alpha\sqrt{1-(v/c)^2}. \tag{226}$$

We may determine α by comparing the non-relativitic limit of this Lagrangian with the ordinary non-relativistic Lagrangian for a free particle $\mathcal{L} = \frac{1}{2}mv^2$. Expanding Eq. 226 in the parameter v/c:

$$\mathcal{L} = -\alpha\left(1-\frac{1}{2}\left(\frac{v}{c}\right)^2+\ldots\right) \tag{227}$$

$$= -\alpha + \alpha\frac{1}{2}\frac{v^2}{c^2}+\ldots \tag{228}$$

The first term in the expansion $(-\alpha)$ is a constant and thus does not influence motion. For consistency of the second term with $\mathcal{L} = \frac{1}{2}mv^2$, we must have $\alpha = mc^2$, and thus we take the relativistic Lagrangian of a free particle to be:

$$\boxed{\mathcal{L} = -mc^2\sqrt{1-(v/c)^2}} \tag{229}$$

What are corresponding canonical momenta? Since all of the components behave similarly, I will just write out x-component, using subscripts in the normal way — I will drop tensor notation for a while, reinstating it at the end of the argument. By definition:

$$p_x = \frac{\partial\mathcal{L}}{\partial v_x}. \tag{230}$$

Substituting the Lagrangian (Eq. 229):

$$p_x = \frac{\partial}{\partial v_x}(-mc^2)\sqrt{1-(v/c)^2} \tag{231}$$

$$= \frac{-mc^2}{2\sqrt{1-(v/c)^2}}\frac{-1}{c^2}\frac{\partial}{\partial v_x}(v_x^2+v_y^2+v_z^2) \tag{232}$$

$$= \frac{mv_x}{\sqrt{1-(v/c)^2}} \tag{233}$$

and similarly for the y and z components, so that we may write

$$\boxed{\mathbf{p} = m\frac{\mathbf{v}}{\sqrt{1-(v/c)^2}}} \tag{234}$$

which reduces to $p = m\mathbf{v}$ in the non-relativistic limit.

- recall that one of the advantages of the Lagrangian approach is that symmetries of a Lagrangian give constants of motion. Here we have a time-independent Lagrangian (a type of symmetry), which leads to a time-independent energy.

First I will review the general case (following Ref. [17]). Consider a system characterized by a set of generalized coordinates q_i and corresponding generalized velocities $\dot{q}_i$, so that:

$$\frac{d\mathcal{L}}{dt} = \sum_i \frac{\partial \mathcal{L}}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\ddot{q}_i. \tag{235}$$

Recall that in Lagrangian mechanics, the Euler-Lagrange equations

$$\frac{d}{dt}\underbrace{\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right)}_{=:p_i} = \frac{\partial \mathcal{L}}{\partial q_i} \tag{236}$$

are the equations of motion, so that we can write Eq. 235 as:

$$\frac{d\mathcal{L}}{dt} = \sum_i \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right)\dot{q}_i + \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\ddot{q}_i. \tag{237}$$

The product rule for derivatives allows the simplification:

$$\frac{d\mathcal{L}}{dt} = \sum_i \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\dot{q}_i\right). \tag{238}$$

Rearranging gives:

$$\frac{d}{dt}\left(\sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}}\dot{q}_i - \mathcal{L}\right) = 0. \tag{239}$$

Thus

$$\sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}}\dot{q}_i - \mathcal{L} \tag{240}$$

is a constant of motion. We call it the energy E. Recalling that $p_i = \partial\mathcal{L}/\partial\dot{q}_i$, we may write: $E \coloneqq -\mathcal{L} + \sum p_i \dot{q}_i$. For the relativistic free particle:

$$E = mc^2\sqrt{1-(v/c)^2} + \frac{m}{\sqrt{1-(v/c)^2}}(v_x^2 + v_y^2 + v_z^2) \tag{241}$$

$$= \frac{mc^2(1-(v/c)^2) + mv^2}{\sqrt{1-(v/c)^2}} \tag{242}$$

$$= \frac{mc^2 - mv^2 + mv^2}{\sqrt{1-(v/c)^2}} \tag{243}$$

$$\boxed{E = \frac{mc^2}{\sqrt{1-(v/c)^2}}} \tag{244}$$

Note the common $1/\sqrt{1-(v/c)^2}$ factor in both this expression and the one for the relativistic momentum (Eq. 234). Recall from Section 18 that this factor relates changes in time to changes in proper time:

$$\frac{dt}{d\tau} = \frac{1}{\sqrt{1-(v/c)^2}}, \tag{245}$$

which suggests a rewriting of the relativistic momentum expression (Eq. 234):

$$\mathbf{p} = m\frac{1}{\sqrt{1-(v/c)^2}}\frac{d\mathbf{r}}{dt} \tag{246}$$

$$= m \underbrace{\frac{1}{\sqrt{1-(v/c)^2}}\frac{d\tau}{dt}}_{=1} \frac{d\mathbf{r}}{d\tau} \tag{247}$$

$$= m\frac{d\mathbf{r}}{d\tau}. \tag{248}$$

Consider that the relativistic energy (Eq. 244), may also be written as a derivative wrt the proper time:

$$E = mc^2\frac{dt}{d\tau}. \tag{249}$$

Since changes in proper time along a world line are Lorentz invariants, if we start with a position four-vector x^α $(= (ct, x, y, z)$; back to tensor notation), and differentiate wrt proper time, we now also have a four-vector $dx^\alpha/d\tau$ by construction. By Eq. 249 our relativistic energy is $E = (mc)\, dx^0/d\tau$ and by Eq. 248 our relativistic momentum components are $m\, dx^\alpha/d\tau$, where $\alpha = 1, 2, 3$. And thus if we consider m to be a Lorentz invariant, then $(E/c, \mathbf{p})$ corresponds to the contravariant components of a four vector. In tensor notation:

$$\boxed{p^\alpha = m\frac{dx^\alpha}{d\tau}} \tag{250}$$

the so-called **momenergy**.[37]

The base $dx^\alpha/d\tau$ (p^α without the mass factor) is known as the **proper velocity**,[38] which GITE4 writes as

$$\boxed{\eta^\alpha := \frac{dx^\alpha}{d\tau}} \tag{251}$$

but is more commonly called u^α; e.g., Ref. [2].

- note that the relativistic velocity addition law (Eq. 68) indicates that the *ordinary* velocity $d\mathbf{r}/dt$ is *not* the space part of a four-vector. The transformation rules for ordinary velocities are quite different than those for four-vectors.

[37] Taylor and Wheeler [18] coined the "momenergy" terminology. It would be great if all of the four-vectors had their own special names distinguishing them from their three-dimensional cousins.

[38] Depending on the context, "proper velocity" may also refer to just the space part of the η^α four-vector.

- since η^α is a four-vector, the quantity $\eta^\alpha \eta_\alpha$ is a Lorentz scalar. Let us compute its value:

$$\begin{aligned} \eta_\alpha \eta^\alpha &= -c\frac{dt}{d\tau}c\frac{dt}{d\tau} + \left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2 + \left(\frac{dz}{d\tau}\right)^2 && (252) \\ &= -c^2\left(\frac{dt}{d\tau}\right)^2 + \left(\frac{dt}{d\tau}\right)^2 v^2 && (253) \\ &= \frac{v^2 - c^2}{1-(v/c)^2} && (254) \\ &= -c^2. && (255) \end{aligned}$$

from which it follows that

$$p_\alpha p^\alpha = -m^2c^2. \quad (256)$$

Writing this equality in terms of the components of p^α, dropping tensor notation:

$$-\left(\frac{E}{c}\right)^2 + p^2 = -m^2c^2. \quad (257)$$

Rearranging gives:

$$E^2 = (pc)^2 + (mc^2)^2. \quad (258)$$

From Eq. 244, we know that $E > 0$, so we select the positive root, to give:

$$\boxed{E = \sqrt{(pc)^2 + (mc^2)^2}} \quad (259)$$

which is one of the most important equations in physics!

20 The Lorentz force law from an invariant action

- we determined a Lagrangian for the relativistic free particle from a simple (simplest?) form for an Lorentz invariant action. Now we will do the same in the presence of a field which influences the particle's motion.
- suppose that the field is described by some sort of four-vector A^α, which varies throughout spacetime; i.e., has time and space dependence. Given a particular world line of a particle, $\int A_\alpha dx^\alpha$ is a Lorentz invariant by construction. Thus a possible form for the invariant action is:[39]

$$S = -mc^2 \int d\tau + q\int A_\alpha dx^\alpha, \quad (260)$$

where q determines the amount of "coupling" of the particle to the field. When $q = 0$, the action just reduces to that of the free particle, which was dealt with in the previous section.

[39] As with the previous section, we omit the starting and ending events that define the limits of the action integral, but they should be understood to be implicit.

- now we follow the same steps as for the free particle, starting with determination of the Lagrangian.

  Recall that with no fields, writing

$$\frac{d\tau}{dt} = \sqrt{1-(v/c)^2} \tag{261}$$

  established that

$$\mathcal{L} = -mc^2\sqrt{1-(v/c)^2}. \tag{262}$$

  To consider the contribution of the second term of Eq. 260 to the Lagrangian, we break A_α up into its time and space parts:

$$A_\alpha = (-A^0, \mathbf{A}) \tag{263}$$

  and drop tensor notation for the rest of this section, so that

$$\int A_\alpha dx^\alpha = \int -A^0 c\, dt + \int \mathbf{A}\cdot\mathbf{dr} \tag{264}$$

  and thus

$$S = -mc^2\int dt\sqrt{1-(v/c)^2} - qc\int A^0 dt + q\int \mathbf{A}\cdot\mathbf{dr}. \tag{265}$$

  Action is the time integration of a Lagrangian, so we rewrite the last term as:

$$q\int \mathbf{A}\cdot\mathbf{dr} = q\int \mathbf{A}\cdot\mathbf{v}dt, \tag{266}$$

  so that from Eq. 265 we may conclude that

$$\mathcal{L} = -mc^2\sqrt{1-(v/c)^2} - qcA^0 + q\mathbf{A}\cdot\mathbf{v}. \tag{267}$$

  From this Lagrangian we may proceed to find the generalized momenta and equations of motion.

  The generalized momenta are

$$P_i = \frac{\partial\mathcal{L}}{\partial\dot{q}_i}. \tag{268}$$

  We know the contribution of the first term of Eq. 267 from the previous section. Only the last term makes an additional contribution, so that:

$$P_x = \frac{mv_x}{\sqrt{1-(v/c)^2}} + qA_x \tag{269}$$

  and similarly for the y and z components, so that

$$\mathbf{P} = \frac{m\mathbf{v}}{\sqrt{1-(v/c)^2}} + q\mathbf{A}. \tag{270}$$

  It is a bit confusing, but in the presence of fields it is conventional to refer to the first term as the **momentum** $=:$ $\mathbf{p}$ and

$$\mathbf{P} = \mathbf{p} + q\mathbf{A} \tag{271}$$

as the **canonical momentum**.

For Euler-Lagrange equations of motion, we need to know $\partial\mathcal{L}/\partial q_i$. For the x coordinate:

$$\frac{\partial \mathcal{L}}{\partial x} = -qc\frac{\partial A^0}{\partial x} + q\frac{\partial}{\partial x}(\mathbf{A}\cdot\mathbf{v}), \tag{272}$$

and similarly for the y and z components. From the Euler-Lagrange equations we now have the equations of motion:

$$\frac{d}{dt}\left(\frac{m\mathbf{v}}{\sqrt{1-(v/c)^2}} + q\mathbf{A}\right) = -qc\nabla A^0 + q\nabla(\mathbf{A}\cdot\mathbf{v}). \tag{273}$$

Rearranging slightly gives:

$$\frac{d}{dt}\left(\frac{m\mathbf{v}}{\sqrt{1-(v/c)^2}}\right) = -q\frac{d\mathbf{A}}{dt} - qc\nabla A^0 + q\nabla(\mathbf{A}\cdot\mathbf{v}). \tag{274}$$

For the $\nabla(\mathbf{A}\cdot\mathbf{v})$ term I will use the general identity

$$\nabla(\mathbf{a}\cdot\mathbf{b}) = \mathbf{a}\times(\nabla\times\mathbf{b}) + \mathbf{b}\times(\nabla\times\mathbf{a}) + (\mathbf{a}\cdot\nabla)\mathbf{b} + (\mathbf{b}\cdot\nabla)\mathbf{a} \tag{275}$$

for arbitrary $\mathbf{a}$ and $\mathbf{b}$. The ∇ in $\nabla(\mathbf{A}\cdot\mathbf{v})$ is a *partial* differentiation with respect to coordinates, with the velocities held constant, so the $\mathbf{A}\times(\nabla\times\mathbf{v})$ and $(\mathbf{A}\cdot\nabla)\mathbf{v}$ terms vanish. Thus

$$\nabla(\mathbf{A}\cdot\mathbf{v}) = \mathbf{v}\times(\nabla\times\mathbf{A}) + (\mathbf{v}\cdot\nabla)\mathbf{A} \tag{276}$$

which we may substitute into Eq. 274 to obtain

$$\frac{d}{dt}\left(\frac{m\mathbf{v}}{\sqrt{1-(v/c)^2}}\right) = -q\frac{d\mathbf{A}}{dt} - qc\nabla A^0 + q\mathbf{v}\times(\nabla\times\mathbf{A}) + q(\mathbf{v}\cdot\nabla)\mathbf{A}. \tag{277}$$

On the other hand, the derivative of $\mathbf{A}$ is a *total* derivative;[40] i.e., how does $\mathbf{A}$ change in time along the particle's trajectory? The chain rule gives:

$$\frac{d\mathbf{A}}{dt} = \frac{\partial\mathbf{A}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{A}. \tag{278}$$

Substituting this derivative into Eq. 277 leads to some cancellation of terms giving:

$$\frac{d}{dt}\left(\frac{m\mathbf{v}}{\sqrt{1-(v/c)^2}}\right) = q\left[-c\nabla A^0 - \frac{\partial\mathbf{A}}{\partial t} + \mathbf{v}\times(\nabla\times\mathbf{A})\right]. \tag{279}$$

The RHS looks familiar if we make the identifications:

$$\mathbf{E} = -c\nabla A^0 - \frac{\partial\mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} = \nabla\times\mathbf{A} \tag{280}$$

[40] This derivative is sometimes known as the **convective derivative**, but also has many other names. A rose something, something ...

since that gives:

$$\frac{d}{dt}\left(\frac{m\mathbf{v}}{\sqrt{1-(v/c)^2}}\right) = q\left[\mathbf{E} + \mathbf{v}\times\mathbf{B}\right]. \tag{281}$$

Equivalently, making use of the relativistic momentum, we have the **relativistic Lorentz force law**:

$$\boxed{\frac{d\mathbf{p}}{dt} = q\left[\mathbf{E} + \mathbf{v}\times\mathbf{B}\right]} \tag{282}$$

which reduces to the non-relativistic version as $v/c \to 0$.

- we commonly call $\mathbf{A}$ the **vector potential** and $\Phi := cA^0$ the **scalar potential**, so that $A^\alpha = (\Phi/c, \mathbf{A})$ and it is a bit cleaner to rewrite Eq. 280 as:

$$\boxed{\mathbf{E} = -\nabla\Phi - \frac{\partial \mathbf{A}}{\partial t}} \quad \text{and} \quad \boxed{\mathbf{B} = \nabla\times\mathbf{A}} \tag{283}$$

- while highly plausible, we have not *proven* the relationships (Eq. 283) between the A^α that appears in the invariant action and the vector fields $\mathbf{E}$ and $\mathbf{B}$, as given by Maxwell's equations. The correctness of these relationships must be experimentally tested. See the example problem: "Bucherer's test of the Lorentz force law".

21 The Minkowski force (GITE4 12.3.4)

- your special relativity spidey-sense should suggest to you that there is something missing in the relativistic Lorentz force law, at least as presented in Eq. 282. Where is the "missing" time-like component? And does the force law obey the principle of covariance? In this section these questions are answered by expressing the Lorentz force law in tensor notation.

  When deriving the equations of motion from the action in the last section, the time coordinate was given a different treatment from the position coordinates; i.e., the action is obtained by integrating the Lagrangian over time. Now we will take a fully-covariant approach and avoid the "special treatment" of time.[41] We will not use the Euler-Lagrange equations — as they give special preference to the time coordinate — but rather consider variations of the action directly — in much the same way that the Euler-Lagrange equations are derived.

- it is helpful to first take the fully-covariant approach with a free particle in the absence of the A^α field, as it gets us warmed up to the relevant manipulations, and yields an intermediate result that will be applied in the general case.

  The action of the free particle S_F is as before:

$$S_F = -mc^2 \int d\tau, \tag{284}$$

[41] Here I am following the treatment found in section 23 of Landau and Lifshitz [2], with some slight differences due to the unit systems. L&L use the more condensed "δ" notation of the calculus of variations.

where integration is along a world-line with fixed starting and ending events. Parameterize the world line using θ, running say from 0 to 1 so that:

$$S_F = -mc \int_0^1 d\theta \sqrt{-\frac{dx^\alpha}{d\theta}\frac{dx_\alpha}{d\theta}}. \tag{285}$$

For a world-line with a stationary action, we should perturb the world-line:

$$x^\alpha \to x^\alpha + \lambda\, \epsilon^\alpha \tag{286}$$

and then differentiate the action wrt λ and then set $\lambda = 0$. The criteria that this derivative be zero — *independent* of the specifics of the ϵ^α perturbation — gives the equations of motion that the world-line must obey. This procedure is how the Euler-Lagrange equations are derived, but with the trajectory parameterized by t.

Following this plan:

$$\frac{1}{-mc}\frac{dS_F}{d\lambda} = \frac{d}{d\lambda}\int_0^1 d\theta \sqrt{-\frac{d(x^\alpha + \lambda\epsilon^\alpha)}{d\theta}\frac{d(x_\alpha + \lambda\epsilon_\alpha)}{d\theta}} \tag{287}$$

$$= \int_0^1 d\theta\, \frac{-1}{2\sqrt{\cdots}}\left[\frac{d\epsilon^\alpha}{d\theta}\frac{d(x_\alpha + \lambda\epsilon_\alpha)}{d\theta} + \frac{d(x^\alpha + \lambda\epsilon^\alpha)}{d\theta}\frac{d\epsilon_\alpha}{d\theta}\right]. \tag{288}$$

Set $\lambda = 0$, to obtain:

$$\frac{1}{-mc}\frac{dS_F}{d\lambda}\bigg|_{\lambda=0} = \int_0^1 d\theta\, \frac{-1}{2\sqrt{\cdots}}\left[\frac{d\epsilon^\alpha}{d\theta}\frac{dx_\alpha}{d\theta} + \frac{dx^\alpha}{d\theta}\frac{d\epsilon_\alpha}{d\theta}\right] \tag{289}$$

$$= \int_0^1 d\theta \frac{-1}{\sqrt{-\frac{dx^\alpha}{d\theta}\frac{dx_\alpha}{d\theta}}}\ \frac{d\epsilon^\alpha}{d\theta}\frac{dx_\alpha}{d\theta}. \tag{290}$$

But

$$\sqrt{-\frac{dx^\alpha}{d\theta}\frac{dx_\alpha}{d\theta}} = c\frac{d\tau}{d\theta} \tag{291}$$

so that:

$$\frac{1}{-m}\frac{dS_F}{d\lambda}\bigg|_{\lambda=0} = -\int_0^1 d\theta\, \frac{dx_\alpha}{d\tau}\frac{d\epsilon^\alpha}{d\theta}. \tag{292}$$

Recognizing the proper velocity η_α,

$$\frac{1}{-m}\frac{dS_F}{d\lambda}\bigg|_{\lambda=0} = -\int_0^1 d\theta\, \eta_\alpha \frac{d\epsilon^\alpha}{d\theta}. \tag{293}$$

Apply integration by parts:

$$\frac{1}{-m}\frac{dS_F}{d\lambda}\bigg|_{\lambda=0} = -\eta_\alpha\epsilon^\alpha\bigg|_0^1 + \int_0^1 d\theta\ \epsilon^\alpha\frac{d\eta_\alpha}{d\theta}. \tag{294}$$

The first term is zero, because all possible perturbations ϵ^α should be zero at the starting and ending events, so that:

$$\frac{1}{-m}\frac{dS_F}{d\lambda}\bigg|_{\lambda=0} = \int_0^1 d\theta\ \epsilon^\alpha \frac{d\eta_\alpha}{d\theta}. \tag{295}$$

This equation shows us that to obtain a stationary action, $dS_F/d\lambda\,|_{\lambda=0} = 0$, we must have $d\eta_\alpha/d\theta = 0$ for all θ along the world-line, since ϵ^α is arbitrary (apart from being zero at the start and end points). Thus the proper velocity should be constant for a free particle, which is certainly no surprise.

Now we consider the presence of the A_α field. We will write the action as $S = S_F + S_A$, where S_F is the free particle part, which we have just studied in isolation, and

$$S_A \coloneqq q\int A_\alpha dx^\alpha. \tag{296}$$

Again we will need to perturb the world line and set $dS/d\lambda\,|_{\lambda=0} = 0$ to obtain the equations of motion.

For this purpose, we compute:

$$\frac{1}{q}\frac{dS_A}{d\lambda} = \frac{d}{d\lambda}\int_0^1 d\theta\ A_\alpha(x^0+\lambda\epsilon^0, x^1+\lambda\epsilon^1, x^2+\lambda\epsilon^2, x^3+\lambda\epsilon^3)\ \frac{d(x^\alpha+\lambda\epsilon^\alpha)}{d\theta}, \tag{297}$$

where the explicit dependence of A_α on position in space-time has been indicated — to emphasize that as the world-line is varied, A_α is sampled along a different path in space-time.

By the chain rule:[42]

$$\frac{dA_\alpha}{d\lambda}\bigg|_{\lambda=0} = \epsilon^\beta\partial_\beta A_\alpha \tag{298}$$

and thus

$$\frac{1}{q}\frac{dS_A}{d\lambda}\bigg|_{\lambda=0} = \int_0^1 d\theta\ \left[A_\alpha\frac{d\epsilon^\alpha}{d\theta} + (\epsilon^\beta\partial_\beta A_\alpha)\frac{dx^\alpha}{d\theta}\right]. \tag{299}$$

For the first term we may use integration by parts to obtain:

$$\frac{1}{q}\frac{dS_A}{d\lambda}\bigg|_{\lambda=0} = A_\alpha\epsilon^\alpha\bigg|_0^1 + \int_0^1 d\theta\ \left[-\frac{dA_\alpha}{d\theta}\epsilon^\alpha + (\epsilon^\beta\partial_\beta A_\alpha)\frac{dx^\alpha}{d\theta}\right]. \tag{300}$$

The first term vanishes because the perturbing ϵ is constrained to be zero at the end points. For $dA_\alpha/d\theta$ we apply the chain rule, so that

$$\frac{1}{q}\frac{dS_A}{d\lambda}\bigg|_{\lambda=0} = \int_0^1 d\theta\ \left[-(\partial_\beta A_\alpha)\frac{dx^\beta}{d\theta}\epsilon^\alpha + (\epsilon^\beta\partial_\beta A_\alpha)\frac{dx^\alpha}{d\theta}\right]. \tag{301}$$

[42] This application of the chain rule is analogous to the second term in the convective derivative of Eq. 278.

To obtain a tidier integrand, the α and β indices in the first term can be interchanged, allowing a factorization, giving:

$$\frac{1}{q}\frac{dS_A}{d\lambda}\bigg|_{\lambda=0} = \int_0^1 d\theta\,\frac{dx^\alpha}{d\theta}\,\epsilon^\beta\,\left[-(\partial_\alpha A_\beta)+(\partial_\beta A_\alpha)\right]. \tag{302}$$

Now we can combine this result with Eq. 295 to obtain the first order variation in the action:

$$\frac{dS}{d\lambda}\bigg|_{\lambda=0} = \frac{dS_F}{d\lambda}\bigg|_{\lambda=0} + \frac{dS_A}{d\lambda}\bigg|_{\lambda=0} \tag{303}$$

$$= \int_0^1 d\theta\,\epsilon^\beta\left[-m\frac{d\eta_\beta}{d\theta} + q\frac{dx^\alpha}{d\theta}(\partial_\beta A_\alpha - \partial_\alpha A_\beta)\right]. \tag{304}$$

Again, as for the free particle, the perturbation ϵ^β is arbitrary (except at its end points), so the bracketed factor must be zero for stationary action, $dS/d\lambda\,|_{\lambda=0} = 0$. Setting this factor equal to zero and rearranging gives:

$$m\frac{d\eta_\beta}{d\theta} = q\frac{dx^\alpha}{d\theta}(\partial_\beta A_\alpha - \partial_\alpha A_\beta) \tag{305}$$

$$m\frac{d\tau}{d\theta}\frac{d\eta_\beta}{d\tau} = q\frac{d\tau}{d\theta}\frac{dx^\alpha}{d\tau}(\partial_\beta A_\alpha - \partial_\alpha A_\beta) \tag{306}$$

$$\frac{dp_\beta}{d\tau} = q\eta^\alpha(\partial_\beta A_\alpha - \partial_\alpha A_\beta). \tag{307}$$

On the LHS, the covariant components $dp_\beta/d\tau$ correspond to a four-vector known as the **Minkowski force** ($=: K_\beta$). The **Faraday tensor**[43] is defined as $F_{\beta\alpha} := (\partial_\beta A_\alpha - \partial_\alpha A_\beta)$, so that the relativistic Lorentz force law can be written in a wonderfully concise and manifestly covariant[44] form:

$$\boxed{\frac{dp_\beta}{d\tau} = q\eta^\alpha F_{\beta\alpha}} \tag{308}$$

- it is straightforward to verify that the space part of Eq. 308 is equivalent to the "three-vector" relativistic Lorentz force law of the previous section, Eq. 282. However, we now have an additional "time-like" component to consider:

$$\frac{dp_0}{d\tau} = q\eta^0 F_{00} + q\eta^1 F_{01} + q\eta^2 F_{02} + q\eta^3 F_{03}. \tag{309}$$

These components of the Faraday tensor may be worked out from its definition:

$$F_{\beta\alpha} := (\partial_\beta A_\alpha - \partial_\alpha A_\beta) \tag{310}$$

[43] GITE4 calls it the "field tensor", but adding the "Faraday" qualifier makes it a bit more clear that it is related to the *electric and magnetic* fields.

[44] Just a reminder about the double meaning of "covariant" (as previously discussed in footnote 29): here I mean that the Lorentz force law has the same form in all inertial reference frames — the principle of covariance introduced in Section 15. It is also true that the LHS of Eq. 308 gives the covariant *components* of the Minkowski force — as opposed to the contravariant components — but we could just as easily write an equivalent equation for its contravariant components following the normal rules described in Section 17.

which makes it clear that it is an **anti-symmetric tensor**, $F_{\beta\alpha} = -F_{\alpha\beta}$, and for the purpose of simplifying Eq. 309, $F_{00} = 0$. For the other components required in Eq. 309, we have:

$$F_{01} = \partial_0 A_1 - \partial_1 A_0 \tag{311}$$
$$= \frac{1}{c}\frac{\partial A^1}{\partial t} + \frac{\partial \Phi/c}{\partial x} \tag{312}$$
$$= -E_x/c. \tag{313}$$

Similarly, $F_{02} = -E_y/c$ and $F_{03} = -E_z/c$. Thus, Eq. 309 may be written as:

$$\frac{dp_0}{d\tau} = -q\frac{1}{\sqrt{1-(v/c)^2}}\mathbf{v}\cdot\mathbf{E}/c \tag{314}$$
$$\frac{d}{d\tau}\left(\frac{-mc^2}{\sqrt{1-(v/c)^2}}\frac{1}{c}\right) = -q\frac{1}{\sqrt{1-(v/c)^2}}\mathbf{v}\cdot\mathbf{E}/c. \tag{315}$$

Remember that $d\tau/dt = \sqrt{1-(v/c)^2}$, so we may convert the derivative on the LHS to a derivative wrt ordinary time to obtain:

$$\boxed{\frac{d}{dt}\left(\frac{mc^2}{\sqrt{1-(v/c)^2}}\right) = q\mathbf{v}\cdot\mathbf{E}} \tag{316}$$

which is a relativistically correct work-energy type theorem: the rate of change of energy appearing on the LHS and a "power" provided by the electric field on the RHS; cf. Section 12.2.4 of of GITE4, esp. the discussion preceding Eq. 12.64. Note that since $\mathbf{v}\cdot(\mathbf{v}\times\mathbf{B}) = 0$, the magnetic field makes no contribution to the work.

22 Transformations of electric and magnetic fields between inertial frames (GITE4 12.3.2, 12.3.3)

- in the previous section we have seen that the components of the **E** field appear as components of the Faraday tensor. Since the Faraday tensor is anti-symmetric ($F_{\beta\alpha} = -F_{\alpha\beta}$), it only has $(4\times 4 - 4)/2 = 6$ independent components. We know from the previous section that $F_{01} = -E_x/c$, $F_{02} = -E_y/c$, and $F_{03} = -E_z/c$. The other three independent components of the Faraday tensor correspond to the three components of the magnetic field. For example:

$$F_{12} = \partial_1 A_2 - \partial_2 A_1 \tag{317}$$
$$= \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}. \tag{318}$$

Recall that we have previously noted (Eq. 283) that $\mathbf{B} = \nabla\times\mathbf{A}$, so

$$F_{12} = B_z \tag{319}$$

and similarly $F_{23} = B_x$ and $F_{31} = B_y$. For convenience, I have listed the components of $F_{\mu\nu}$, $F^{\mu}{}_{\nu}$, $F_{\mu}{}^{\nu}$, and $F^{\mu\nu}$ in the Formulae summary.[45]

[45] GITE4 notes variations in the definition of the Faraday tensor. The definition here (Eq. 310) is equivalent to the one in GITE4.

- the Faraday tensor is a Lorentz tensor by construction, which means that we know precisely how it transforms between different inertial references frames (see Section 17). If we denote quantities in the new inertial frame by ˜'s, then:

$$\tilde{F}^{\mu\beta} = \Lambda^{\mu}_{\alpha}\,\Lambda^{\beta}_{\kappa}\,F^{\alpha\kappa}. \tag{320}$$

  Since the Faraday tensor consists of the components of **E** and **B**, this equation tells us how to transform **E** and **B** between inertial reference frames. (We discussed a special case in Section 15.)

- it is useful to write plain-old "three-vector" versions of the transformation rules of Eq. 320, making explicit reference to **E** and **B** in both frames.

  For this purpose, let us consider two inertial frames in the standard configuration, with relative motion along the x-axis, quantified by a rapidity θ, so that Λ is of the form given in Eq. 200.

  First consider $\tilde{E}_x$, the component of the electric field parallel to the relative velocity between the two frames:

$$\begin{aligned}
\tilde{E}_x &= c\tilde{F}^{01} && (321)\\
&= c\,\Lambda^{0}_{\alpha}\,\Lambda^{1}_{\kappa}\,F^{\alpha\kappa} && (322)\\
&= c\,\Lambda^{0}_{\alpha}\,\left(\Lambda^{1}_{0}F^{\alpha 0} + \Lambda^{1}_{1}F^{\alpha 1} + \Lambda^{1}_{2}F^{\alpha 2} + \Lambda^{1}_{3}F^{\alpha 3}\right) && (323)\\
&= c\,\Lambda^{0}_{\alpha}\,\left(\Lambda^{1}_{0}F^{\alpha 0} + \Lambda^{1}_{1}F^{\alpha 1}\right) && (324)\\
&= c\,\Lambda^{0}_{\alpha}\,\left(-\sinh\theta F^{\alpha 0} + \cosh\theta F^{\alpha 1}\right) && (325)\\
&= c\,\Lambda^{0}_{0}\,\left(-\sinh\theta F^{00} + \cosh\theta F^{01}\right) + c\,\Lambda^{0}_{1}\,\left(-\sinh\theta F^{10} + \cosh\theta F^{11}\right)\\
&\quad + c\,\Lambda^{0}_{2}\,\left(-\sinh\theta F^{20} + \cosh\theta F^{21}\right) + c\,\Lambda^{0}_{3}\,\left(-\sinh\theta F^{30} + \cosh\theta F^{31}\right) && (326)\\
&= c\,\Lambda^{0}_{0}\,\left(\cosh\theta F^{01}\right) + c\,\Lambda^{0}_{1}\,\left(-\sinh\theta F^{10}\right) && (327)\\
&= c\,\cosh\theta\,\left(\cosh\theta F^{01}\right) + c\,(-\sinh(\theta))\,\left((-\sinh\theta)F^{10}\right). && (328)
\end{aligned}$$

  By the anti-symmetry of the Faraday tensor, $F^{10} = -F^{01}$, and thus a considerable simplification results:

$$\begin{aligned}
\tilde{E}_x &= c(\cosh^2\theta - \sinh^2\theta)F^{01} && (329)\\
&= cF^{01} && (330)\\
&= E_x. && (331)
\end{aligned}$$

  In other words, the component of the electric field along the direction of relative frame motion remains unchanged. By another set of equally straightforward — but tedious(!) — manipulations, the analogous result is found for the magnetic field: $\tilde{B}_z = B_z$.

  The components of the **E** and **B** fields which are orthogonal to the relative frame motion are more interesting. Consider, for example:

$$\begin{aligned}
\tilde{E}_y &= c\tilde{F}^{02} && (332)\\
&= c\,\Lambda^0_{\ \alpha}\,\Lambda^2_{\ \kappa}\,F^{\alpha\kappa} && (333)\\
&= c\,\Lambda^0_{\ \alpha}\,\left(\Lambda^2_{\ 0}F^{\alpha 0} + \Lambda^2_{\ 1}F^{\alpha 1} + \Lambda^2_{\ 2}F^{\alpha 2} + \Lambda^2_{\ 3}F^{\alpha 3}\right) && (334)\\
&= c\,\Lambda^0_{\ \alpha}\,\Lambda^2_{\ 2}\,F^{\alpha 2} && (335)\\
&= c\,\Lambda^2_{\ 2}\left(\Lambda^0_{\ 0}\,F^{02} + \Lambda^0_{\ 1}\,F^{12} + \Lambda^0_{\ 2}\,F^{22} + \Lambda^0_{\ 3}\,F^{32}\right) && (336)\\
&= c\left(\Lambda^0_{\ 0}\,F^{02} + \Lambda^0_{\ 1}\,F^{12}\right) && (337)\\
&= c\left(\Lambda^0_{\ 0}\,E_y/c + \Lambda^0_{\ 1}\,B_z\right) && (338)\\
&= \cosh\theta\,E_y - \sinh\theta\,cB_z. && (339)
\end{aligned}$$

By similar manipulations:

$$\begin{aligned}
\tilde{E}_z &= \cosh\theta\,E_z + \sinh\theta\,cB_y && (340)\\
\tilde{B}_y &= \cosh\theta\,B_y + \sinh\theta\,E_z/c && (341)\\
\tilde{B}_z &= \cosh\theta\,B_z - \sinh\theta\,E_y/z. && (342)
\end{aligned}$$

These results — specific to a relative frame velocity in the x direction — are inelegant. But by using rotational symmetry they may be put into a more general form. For this purpose, I will denote the velocity of the ˜ frame wrt to the original frame by $\mathbf{v}$ (= $\hat{\mathbf{x}}c\tanh\theta$ so far). The $\mathbf{E}$ field may be decomposed:

$$\mathbf{E} = \underbrace{(\hat{\mathbf{v}}\cdot\mathbf{E})\hat{\mathbf{v}}}_{:=\mathbf{E}_\parallel} + \underbrace{\mathbf{E} - (\hat{\mathbf{v}}\cdot\mathbf{E})\hat{\mathbf{v}}}_{:=\mathbf{E}_\perp} \tag{343}$$

and similarly: $\mathbf{B} = \mathbf{B}_\parallel + \mathbf{B}_\perp$. As we have shown, $\tilde{\mathbf{B}}_\parallel = \mathbf{B}_\parallel$ and $\tilde{\mathbf{E}}_\parallel = \mathbf{E}_\parallel$. For the perpendicular component of $\mathbf{E}$:

$$\begin{aligned}
\tilde{\mathbf{E}}_\perp &= \hat{\mathbf{y}}(\cosh\theta\,E_y - \sinh\theta\,cB_z) + \hat{\mathbf{z}}(\cosh\theta\,E_z + \sinh\theta\,cB_y) && (344)\\
&= \cosh\theta\,(\hat{\mathbf{z}}E_z + \hat{\mathbf{y}}E_y) + \sinh\theta\,(\hat{\mathbf{z}}cB_y - \hat{\mathbf{y}}cB_z) && (345)\\
&= \cosh\theta\,\mathbf{E}_\perp + \sinh\theta\,c\left[\hat{\mathbf{z}}(\hat{\mathbf{y}}\cdot\mathbf{B}) - \hat{\mathbf{y}}(\hat{\mathbf{z}}\cdot\mathbf{B})\right]. && (346)
\end{aligned}$$

Recall the "bac-cab" identity: $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b})$, which can be used here to write $\hat{\mathbf{z}}(\hat{\mathbf{y}}\cdot\mathbf{B}) - \hat{\mathbf{y}}(\hat{\mathbf{z}}\cdot\mathbf{B}) = \hat{\mathbf{x}}\times\mathbf{B}$, so that:

$$\begin{aligned}
\tilde{\mathbf{E}}_\perp &= \cosh\theta\,\mathbf{E}_\perp + \sinh\theta\,\hat{\mathbf{x}}\times c\mathbf{B} && (347)\\
&= \cosh\theta\,\left[\mathbf{E}_\perp + \tanh\theta\,\hat{\mathbf{x}}\times c\mathbf{B}\right] && (348)\\
&= \cosh\theta\,\left[\mathbf{E}_\perp + \mathbf{v}\times\mathbf{B}\right]. && (349)
\end{aligned}$$

By a similar set of manipulations:

$$\tilde{\mathbf{B}}_\perp = \cosh\theta\,\left[\mathbf{B}_\perp - \mathbf{v}/c\times\mathbf{E}/c\right]. \tag{350}$$

Summarizing:

$$\tilde{\mathbf{E}}_\parallel = \mathbf{E}_\parallel \tag{351}$$

$$\tilde{\mathbf{E}}_\perp = \cosh\theta\ \left[\mathbf{E}_\perp + \mathbf{v} \times \mathbf{B}\right] \tag{352}$$

$$\tilde{\mathbf{B}}_\parallel = \mathbf{B}_\parallel \tag{353}$$

$$\tilde{\mathbf{B}}_\perp = \cosh\theta\ \left[\mathbf{B}_\perp - \mathbf{v}/c \times \mathbf{E}/c\right] \tag{354}$$

where $\mathbf{v}$ is the velocity of the $\tilde{}$ frame wrt to the initial frame and $\tanh\theta = v/c$.

- GITE4 derives these transformations differently — starting from the assumption that Maxwell's equations have the same form in all inertial reference frames (the principle of covariance). By transforming some well-chosen special cases of the "source fields", $\mathbf{J}$ and ρ, between two inertial reference frames, the same rules as given by Eq.'s 351 to 354 may be inferred.

  The approach that we took here is quite different — I made no reference to Maxwell's equations!

  We can view the agreement between these two approaches as reinforcing the hypothesized correspondence (Eq. 283) between

  1. the A^α four-vector arising from our variational arguments for a relativistic correct Lorentz force law, and
  2. the $\mathbf{E}$ and $\mathbf{B}$ fields as computed from Maxwell's equations.

  Aside from the annoyance associated with rewriting the clean(!) transformation rules for the Faraday tensor in terms of the "plain-old" $\mathbf{E}$ and $\mathbf{B}$ vectors, the variational approach is quite appealing in its economy. L&L go on to motivate Maxwell's equations using variational methods, but here we will (unfortunately) break with them. However, in the next section we will rewrite Maxwell's equations in a manifestly covariant form, similar in spirit to the way we rewrote the Lorentz Force law (Eq. 282) in a manifestly covariant form (Eq. 308).

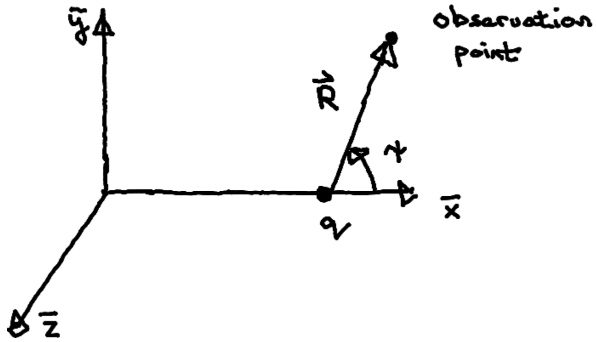
23 Fields due to a uniformly moving point charge (GITE4 12.3.4)

Summary:

- determining the fields due to uniformly moving point charge is a straightforward but important application of the rules for transforming fields between inertial reference frames.
- take two inertial reference frames with relative motion along the x axis — the "standard configuration". In the first frame, the point charge is stationary and located at the origin. In this inertial reference frame the electric field is just given by Coulomb's law and the magnetic field is zero.

  We then transform these fields into the second inertial reference frame, travelling at a velocity $c\boldsymbol{\beta} = V_{21,x}\hat{\mathbf{x}}$ with respect to the first. Once we have the fields in this second reference frame, we have what we want: the fields for a point charge moving with a velocity of $-c\boldsymbol{\beta}$.

- define, in the second frame, in which the particle is moving:



where

$$\mathbf{R} := (\bar{x} - v\bar{t})\,\hat{\mathbf{x}} + \bar{y}\,\hat{\mathbf{y}} + \bar{z}\,\hat{\mathbf{z}} \tag{355}$$

and the angle ψ is such that

$$\hat{\mathbf{x}} \cdot \mathbf{R} = R\cos\psi \tag{356}$$

and

$$\frac{y^2 + z^2}{R^2} = \sin^2\psi. \tag{357}$$

In the frame in which the particle is moving, the fields are:

$$\boxed{\bar{\mathbf{E}} = \frac{q}{4\pi\epsilon_0}\frac{\hat{\mathbf{R}}}{R^2}\frac{(1-\beta^2)}{(1-\beta^2\sin^2\psi)^{3/2}}} \tag{358}$$

and

$$\boxed{\bar{\mathbf{B}} = \frac{\mathbf{v}}{c} \times \frac{\bar{\mathbf{E}}}{c}} \tag{359}$$

where $\beta = v/c$ and $\mathbf{v}$ is the velocity of the particle in this frame.

24 Maxwell's equations in a manifestly covariant form (GITE4 12.3.3)

Summary:

- justified treating $(\rho c, \mathbf{J})$ as a four-vector.
- described how using A^α ($\equiv (\varphi/c, \mathbf{A})$) to describe $\mathbf{E}$ and $\mathbf{B}$ satisfied two of Maxwell's equations.
- wrote the other two of Maxwell's equations in terms of φ and $\mathbf{A}$.
- explained gauge invariance and that we have the freedom to set $\nabla \cdot \mathbf{A}$ to our own specifications.

- mentioned why the **Coulomb gauge** $\nabla \cdot \mathbf{A} = 0$ is so useful in magnetostatics.
- introduced the **Lorenz gauge** $\nabla \cdot \mathbf{A} = -(1/c^2)\partial\varphi/\partial t$ and showed how it simplifies Maxwell's equations (when written in terms of $\mathbf{A}$ and φ) to a pair of "wave equations".
- showed how these two wave equations could be written together in the highly condensed and manifestly covariant form:

$$\boxed{-\partial_\alpha \partial^\alpha A^\beta = \mu_0 J^\beta} \tag{360}$$

25 Poynting's theorem (GITE4 8.1.2)

- Poynting's theorem concerns energy conservation.
- final result:

$$\underbrace{-\oint \mathbf{S} \cdot d\mathbf{a}}_{\substack{\text{power flow}\\ \text{into volume}}} = \underbrace{\frac{dE}{dt}}_{\substack{\text{power being put}\\ \text{into particles}\\ \text{in volume}}} + \underbrace{\frac{d}{dt}\int d\tau \left[\frac{\epsilon_0}{2}E^2 + \frac{B^2}{2\mu_0}\right]}_{\substack{\text{power being put}\\ \text{into fields}\\ \text{in volume}}} \tag{361}$$

where the Poynting vector $\mathbf{S}$ is defined as:

$$\mathbf{S} \coloneqq \frac{1}{\mu_0}\mathbf{E} \times \mathbf{B}. \tag{362}$$

26 Momentum conservation and Maxwell's stress tensor (GITE4 8.2.1, 8.2.2, 8.2.3)

- final result:

$$\underbrace{\oint \overleftrightarrow{\mathbf{T}} \cdot d\mathbf{a}}_{\substack{\text{rate of momentum}\\ \text{flow into volume}}} = \underbrace{\frac{d\mathbf{p}}{dt}}_{\substack{\text{rate of}\\ \text{particle momentum}\\ \text{change in volume}}} + \underbrace{\int d\tau\, \epsilon_0 \frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B})}_{\substack{\text{rate of}\\ \text{electromagnetic momentum}\\ \text{change in volume}}} \tag{363}$$

where the components of the symmetric Maxwell stress tensor $\overleftrightarrow{\mathbf{T}}$ are defined as:

$$T_{ij} \coloneqq \epsilon_0(E_i E_j - \frac{1}{2}E^2\delta_{ij}) + \frac{1}{\mu_0}(B_i B_j - \frac{1}{2}B^2\delta_{ij}). \tag{364}$$

27 Wave equations (GITE4 9.1.1, 9.1.2)

- XX

28 Electromagnetic plane waves (GITE4 9.2.1, 9.2.2)

- XX

29 Energy and momentum transport in electromagnetic waves (GITE4 9.2.3)

- XX

30 Solving Maxwell's equations in the Lorenz gauge (GITE4 10.2.1)

- XX

31 Introduction to radiation

We will study two topics:

1. the Hertzian dipole. This is the simplest radiating system and of great historical importance, as it was used by Hertz to demonstrate that electromagnetic waves could be generated electrically.
2. Purcell's treatment of radiation of a point charge. It is the *acceleration* of point charges that give rise to radiation. There is a simple but intuitive picture that allows us to see this.

All of the videos for radiation were made in Fall 2020. Some things that I say are slightly outdated (reference to problem sets etc...).

32 Hertzian dipole (GITE4 11.1.2)

- *Where are we going?*
- *Determining the vector potential*

  Assuming that the wavelength is small compared to the size of the dipole, we obtain:

$$\mathbf{A}(r,\theta,t) = -\frac{\mu_0 p_0 \omega}{4\pi r}\sin\left[\omega(t-r/c)\right]\hat{\mathbf{z}} \qquad \text{(GITE4 Eq. 11.17)}$$

- *Determining the scalar potential*

  The Lorenz gauge condition may then be integrated to determine:

$$V(r,\theta,t) = \frac{p_0 \cos\theta}{4\pi\epsilon_0 r}\left\{-\frac{\omega}{c}\sin[\omega(t-r/c)] + \frac{1}{r}\cos[\omega(t-r/c)]\right\} \qquad \text{(GITE4 Eq. 11.12)}$$

- *Determining the* **E** *and* **B** *fields*

  When determining the **E** and **B** fields we only retain terms that give oscillating amplitudes that scale like $1/r$, ignoring $1/r^2$ and quicker fall-offs, as these do not carry energy off to infinity. To indicate this we write $\mathbf{E}_{\rm rad}$ and $\mathbf{B}_{\rm rad}$ where "rad" is for radiation:

$$\mathbf{E}_{\rm rad} = -\frac{\mu_0 p_0 \omega^2}{4\pi}\left(\frac{\sin\theta}{r}\right)\cos[\omega(t-r/c)]\hat{\theta} \qquad \text{(GITE4 Eq. 11.18)}$$

$$\mathbf{B}_{\rm rad} = -\frac{\mu_0 p_0 \omega^2}{4\pi c}\left(\frac{\sin\theta}{r}\right)\cos[\omega(t-r/c)]\hat{\phi} \qquad \text{(GITE4 Eq. 11.19)}$$

- *Determining the radiated power*

  These fields may be used to compute the time-averaged power radiated off to infinity per unit solid angle:

$$\frac{dP}{d\Omega} = r^2 \langle \mathbf{S}\cdot\hat{r}\rangle_T \qquad (365)$$

  where $\langle\ldots\rangle$ indicates a time average (here over a single cycle $T = 2\pi/\omega$).

$$\frac{dP}{d\Omega} = \frac{1}{2\mu_0 c}\left(\frac{\mu_0 dq\omega^2}{4\pi}\right)^2 \sin^2\theta \qquad (366)$$

  Integrating over all solid angles gives:

$$P = \frac{\mu_0 p_0^2 \omega^4}{12\pi c}. \qquad \text{(GITE Eq. 11.22)}$$

33 Intuitive treatment of point charge radiation (GITE4 11.2.1)

- broken up into three videos:
  1. *Determination of E field*
  2. *Determination of B field*
  3. *Determination of radiated power*

- the argument presented is originally due to J. J. Thomson[46] and was popularized by Purcell's textbook.

  The only thing that I add (in the videos) to Purcell's treatment is an argument for the radiation magnetic field.

  Here is a helpful animation of the fields generated by the charge's deceleration.

- final result is:
$$\frac{dP}{d\Omega} = \frac{\mu_0 q^2 a^2}{6\pi c}\frac{3}{8\pi}\,\sin^2\theta \tag{367}$$
  where a is the acceleration of a particle of charge q. After integration over all solid angles we have:
$$P = \frac{\mu_0 q^2 a^2}{6\pi c} \tag{GITE4 Eq. 11.70}$$
  which is known as **Larmor's formula**.

.

Part I

Appendices

A Electromagnetic unit systems demystified

For "reasons" some people continue to weigh themselves in "pounds" and measure their height in "feet" and "inches".[47] Thus, it should be no surprise that not everyone measures charge in Coulombs, magnetic fields in Tesla, and so-on. Although I use the SI system in these notes (as do most undergraduate textbooks), you can expect to encounter the Gaussian and Heaviside-Lorentz systems "in the wild" — especially in more advanced works. This appendix covers most (but not all) of what you might need to know about these alternatives to the SI system.

Actually it is a bit more than the "units" that change between the different systems — the form of purely symbolic equations can differ significantly between the systems as well. For example: if you have ever thought that Coulomb's law should be rid of its awkward $1/(4\pi\epsilon_0)$ factor and charge redefined so that $F = q_1 q_2/r^2$ instead, then you will understand the appeal of Gaussian unit system, in which Coulomb's law takes this simplified form. In contrast, consider $\mathbf{F} = m\mathbf{a}$, where $\mathbf{F}$ is in N, m is in kg and $\mathbf{a}$ is in m/s^2. We also have $\mathbf{F} = m\mathbf{a}$ when $\mathbf{F}$ is in pounds-force, m is in "slugs" and $\mathbf{a}$ is in ft/s^2. The form of equation $\mathbf{F} = m\mathbf{a}$ remains unchanged despite using a different unit system. But as the example of Coulomb's law illustrates, equations *do* change their form between the different

[46] Thomson's presentation may be found here, but it is mainly of historical interest.

[47] From the Carter to Reagan transition of my childhood I recall the metric system being vaguely associated with communism. "Mr. Gorbachev tear down this wall."

electromagnetic unit systems. Despite its lack of precision the "unit system" terminology is what we are stuck with, and I will continue to use it here.

Virtually every book on electromagnetism contains an appendix like this one. Perhaps the most comprehensive is in Jackson's *Classical Electrodynamics* [9], and I refer you to there for a different outlook and more details on some of the more obscure systems — here I limit the scope to the relationship between the SI system and the two of the other commonly used systems: Gaussian and Heaviside-Lorentz. But I will write everything out and motivate the differences between the systems, in what may seem like excessive detail (to some).

I will start with fields, currents, and charge densities written in the SI system and then consider two "successive changes" of the field variables parameterized using "scale factors" — α and β. The different unit systems will correspond to specific choices for α and β.

The first change of variables allows $\mathbf{B}$ to be expressed in the same units as $\mathbf{E}$ — in harmony with special relativity. A new magnetic field $\mathbf{B'}$ is defined such that the SI magnetic field is given by $\mathbf{B} = \mathbf{B}'/\alpha$. The electric field remains unchanged, as do all the current and charge densities, so that $\mathbf{E} = \mathbf{E}'$, $\rho = \rho'$, $\mathbf{J} = \mathbf{J}'$, $q = q'$. The units of length, mass, and time remain unchanged and thus I will not use primes for non-electrical quantities.[48]

We now rewrite the Lorentz force law in SI form $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, in terms of the new variables:

$$\mathbf{F} = q'(\mathbf{E}' + \frac{\mathbf{v}}{\alpha} \times \mathbf{B}'). \tag{368}$$

Normally we would select $\alpha = c$ so that $\mathbf{B'}$ is now in the same units as $\mathbf{E'}$, but for emphasis on the generic nature of the transformations, it will be kept unspecified for now.

The motivation for the second transformation is the goal of "cleaning up" Coulomb's law:

$$F = \frac{q^2}{4\pi\epsilon_0}\frac{1}{r^2} \qquad \rightarrow \qquad F = \frac{q''^2}{r^2}. \tag{369}$$

Or in terms of Maxwell's equations:

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0 \qquad \rightarrow \qquad \nabla \cdot \mathbf{E}'' = 4\pi\rho''. \tag{370}$$

To accomplish these transformations let us define a "new" electric field $\mathbf{E}''$ such that

$$\mathbf{E} = \mathbf{E}' = \frac{\mathbf{E}''}{\beta} \tag{371}$$

where β is now an "adjustable" parameter for this second transformation in the same way that α was for the first.

If we want the Lorentz force law to retain the form of Eq. 368 after the replacement of $\mathbf{E'}$ with $\mathbf{E}''/\beta$ then we must also rescale q; i.e., $q' = \beta q''$, so that

$$\mathbf{F} = q'\mathbf{E}' \tag{372}$$

$$= \beta q''\,\mathbf{E}''/\beta \tag{373}$$

$$= q''\mathbf{E}''. \tag{374}$$

[48] It is conventional in the Gaussian and Heaviside-Lorentz systems to use centimetres, grams, and seconds instead of metres, kilograms, and seconds; but that does not change the symbolic form of equations; i.e., $\mathbf{F} = m\mathbf{a}$ in both.

For consistency, the transformation of q' to q'' must be accompanied by changes in the charge density: $\rho' = \beta\rho''$. And then if we want to preserve the form of the charge continuity equation $\nabla \cdot \mathbf{J}' = -\partial\rho'/\partial t$, we must change the current density: $\mathbf{J}' = \beta\mathbf{J}''$.

The scaling of q also means that we need to scale $\mathbf{B}$ to preserve the form of the

$$q\,\frac{\mathbf{v}}{\alpha} \times \mathbf{B} \tag{375}$$

term of the Lorentz force law. Specifically, $\mathbf{B}' = \mathbf{B}''/\beta$, so that

$$\begin{aligned} \mathbf{F} &= q'\frac{\mathbf{v}}{\alpha} \times \mathbf{B}' && (376)\\ &= q''\beta\frac{\mathbf{v}}{\alpha} \times \frac{\mathbf{B}''}{\beta} && (377)\\ &= q''\frac{\mathbf{v}}{\alpha} \times \mathbf{B}'' && (378) \end{aligned}$$

With the second change of variables: $\mathbf{E}' = \mathbf{E}''/\beta$, $\mathbf{B}' = \mathbf{B}''/\beta$, $q' = \beta q''$, $\rho' = \beta\rho''$, $\mathbf{J}' = \beta\mathbf{J}''$, Lorentz's force law and Maxwell's equations now have the forms:

$$\mathbf{F} = q''\left(\mathbf{E}'' + \frac{\mathbf{v}}{\alpha} \times \mathbf{B}''\right) \tag{379}$$

$$\begin{aligned} &\nabla \cdot \mathbf{E}'' = \beta^2\rho''/\epsilon_0 && \nabla \times \mathbf{E}'' = -\frac{1}{\alpha}\frac{\partial \mathbf{B}''}{\partial t} && (380)\\ &\nabla \cdot \mathbf{B}'' = 0 && \nabla \times \mathbf{B}'' = \mu_0\alpha\beta^2\mathbf{J}'' + \mu_0\epsilon_0\alpha\frac{\partial \mathbf{E}''}{\partial t} && (381) \end{aligned}$$

And for summary, we note that in terms of the original SI quantities (unprimed), these new variables (double-primed) are:

$$\mathbf{E} = \mathbf{E}''/\beta, \qquad \mathbf{B} = \mathbf{B}''/(\beta\alpha), \qquad q = \beta q'', \qquad \rho = \beta\rho'', \qquad \mathbf{J} = \beta\mathbf{J}'' \tag{382}$$

In principle α and β may be set arbitrarily. But there are only two choices of interest to us here:

(i) the Gaussian system, corresponding to $\alpha = c$, $\beta = \sqrt{4\pi\epsilon_0}$, and

(ii) the Heaviside-Lorentz system, corresponding to $\alpha = c$, $\beta = \sqrt{\epsilon_0}$.

In these systems:

Gaussian	Heaviside-Lorentz
$\nabla \cdot \vec{E}'' = 4\pi \rho''$	$\nabla \cdot \vec{E}'' = \rho''$
$\nabla \times \vec{B}'' = \frac{4\pi}{c} \vec{J}'' + \frac{1}{c} \frac{\partial \vec{E}''}{\partial t}$	$\nabla \times \vec{B}'' = \frac{\vec{J}''}{c} + \frac{1}{c} \frac{\partial \vec{E}''}{\partial t}$

$$\left.\begin{aligned} \nabla \cdot \vec{B}'' &= 0 \\ \nabla \times \vec{E}'' &= -\frac{1}{c} \frac{\partial \vec{B}''}{\partial t} \\ \vec{F} &= q'' \left(\vec{E}'' + \frac{\vec{v}}{c} \times \vec{B}'' \right) \\ \nabla \cdot \vec{J}'' &= -\frac{\partial \rho''}{\partial t} \end{aligned}\right\} \text{the same in both systems}$$

In both of these systems $\mathbf{E}''$ and $\mathbf{B}''$ have the same dimensions and ϵ_0 and μ_0 are dispensed with. The difference between the two systems is whether 4π appears in Gauss' law (Gaussian) or Coulomb's law (Heaviside-Lorentz) and likewise for Ampere's law and the the Biot-Savart law.

Now let us consider a straightforward example of converting an equation from one system to another. Zangwill's [8] Eq. 20.115 gives the power radiated by an accelerating non-relativistic charge as

$$P = \frac{1}{4\pi\epsilon_0} \frac{2q^2}{3c^3} a^2 \tag{383}$$

where a is the acceleration. Zangwill uses the SI system. What is the equivalent expression in the Gaussian system? Using our transformation approach: $q = \beta q''$ with $\beta = \sqrt{4\pi\epsilon_0}$, so that in the Gaussian system we have

$$P = \frac{2q''^2}{3c^3} a^2 \tag{384}$$

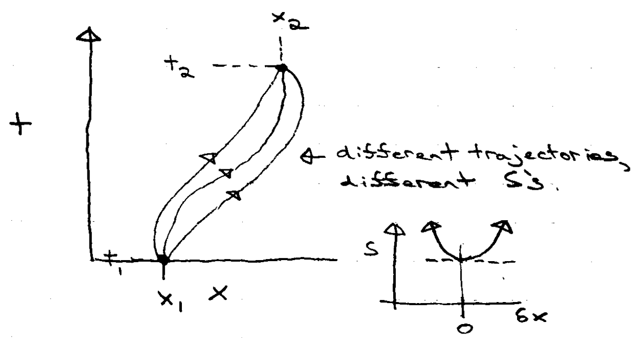
agreeing with the Gaussian system expression given by Jackson [9] (his Eq. 14.22 on page 665):

$$P = \frac{2}{3} \frac{e^2}{c^3} |\dot{v}|^2. \tag{385}$$

As an aside: usage of "e" is usually a hint (but not a guarantee) that the Gaussian or Heaviside-Lorentz systems are being used.

B Review of Lagrangian and Hamiltonian dynamics

Brief review of Lagrangian and Hamiltonian Dynamics (classical)

In general	1-d example (conservative forces)
Start with Lagrangian $\mathcal{L}(\underbrace{q, \dot{q}}_{\text{generalized coordinates and velocities}}, t)$	$\mathcal{L} = T - U = \frac{1}{2}m\dot{x}^2 - V(x)$
Action $S \equiv \int_{t_1}^{t_2} dt\, \mathcal{L}(q, \dot{q}, t)$	$S = \int_{t_1}^{t_2} dt \left(\frac{1}{2}m\dot{x}^2 - V(x)\right)$
must be "stationary" gives us Euler-Lagrange equations: $\underbrace{\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = \frac{\partial \mathcal{L}}{\partial q_i}}$ these should reproduce Newton's 2nd law.	
Define canonical momenta: $p_i \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$	$p_x = m\dot{x}$
Define Hamiltonian: $H \equiv -\mathcal{L} + \sum_i \dot{q}_i p_i$	$H = -\frac{1}{2}m\dot{x}^2 + V(x) + \dot{x}m\dot{x} = \frac{1}{2}m\dot{x}^2 + V(x)$

In general

The time evolution of any quantity $A(p,q)$ is given by:

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

where the Poisson bracket $\{\ldots,\ldots\}$ is defined as

$$\{A,B\} = \sum_i \left[\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i}\right]$$

1-d example

$$\frac{dx}{dt} = \{x, H\}$$

$$= \frac{\partial x}{\partial x}\frac{\partial H}{\partial p} - \underbrace{\cancel{\frac{\partial x}{\partial p}}}_{=0}\frac{\partial H}{\partial x}$$

$$= \frac{p_x}{m}$$

$$\frac{dp_x}{dt} = \{p_x, H\}$$

$$= \underbrace{\cancel{\frac{\partial p_x}{\partial x}}}_{=0}\frac{\partial H}{\partial p_x} - \frac{\partial p_x}{\partial p_x}\frac{\partial H}{\partial x}$$

$$= -\frac{\partial H}{\partial x}$$

$$\frac{dp_x}{dt} = -\frac{\partial V}{\partial x}$$

That was more of a summary than a review. I highly recommend Chapter 2 of Shankar's, *Principles of Quantum Mechanics*, 2nd ed. [19], which despite being a book on *quantum* mechanics, gives an excellent review of classical mechanics using Lagrangians.

C Cyclotron motion

Consider the motion of a charged particle in a homogeneous $\mathbf{B}$ field, with $\mathbf{E} = \mathbf{0}$. In the non-relativistic limit

$$m\dot{\mathbf{v}} = q(\mathbf{v} \times \mathbf{B}), \tag{386}$$

and for concreteness I will take $\mathbf{B}$ to be along the z-axis so that $\mathbf{B} = B_z\hat{\mathbf{z}}$ and thus

$$m\dot{v}_x = qv_yB_z \tag{387}$$
$$m\dot{v}_y = -qv_xB_z \tag{388}$$
$$m\dot{v}_z = 0. \tag{389}$$

There is not much to say about the uniform motion in the z direction (in the non-relativistic case) but the xy motion is a coupled linear system. To make use of some standard results of linear algebra I will use matrix notation:

$$\begin{bmatrix}\dot{v}_x \\ \dot{v}_y\end{bmatrix} = \frac{qB_0}{m}\underbrace{\begin{bmatrix}0 & 1 \\ -1 & 0\end{bmatrix}}_{=:C}\begin{bmatrix}v_x \\ v_y\end{bmatrix} \tag{390}$$

or equivalently:

$$\dot{\mathbf{v}} = \frac{qB}{m} C \mathbf{v}. \tag{391}$$

It is useful to observe that the matrix C is unitary:

$$\begin{aligned} C^\dagger C &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} && (392) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} && (393) \\ &= \mathbb{1}, && (394) \end{aligned}$$

because any unitary matrix may be diagonalized by a unitary matrix, meaning that we may write:

$$C = UDU^\dagger \tag{395}$$

where $UU^\dagger = \mathbb{1}$, and D is a diagonal matrix. The columns of U are a set of orthonormal eigenvectors of C. The corresponding eigenvalues are along the diagonal of D.

Finding two orthogonal eigenvectors to construct U may be done in various ways — I just tried some obvious choices and identified the eigenvectors

$$\begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ -i \end{bmatrix} \tag{396}$$

corresponding to the eigenvalues of $-i$ and i, respectively. Since they are non-degenerate, these eigenvectors are unique to within constant factors. Normalizing them and assembling them into a unitary matrix gives:

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \tag{397}$$

so that we may write:

$$\begin{aligned} C &= UDU^\dagger && (398) \\ &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} -i & \\ & i \end{bmatrix} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}. && (399) \end{aligned}$$

Substituting $C = UDU^\dagger$ into Eq. 391 gives:

$$\dot{\mathbf{v}} = \frac{qB}{m} UDU^\dagger \mathbf{v}. \tag{400}$$

Applying $U^\dagger$ from the left to both sides and making the change of variables $\mathbf{u} \coloneqq U^\dagger \mathbf{v}$ gives:

$$\dot{\mathbf{u}} = \frac{qB_z}{m} D \mathbf{u}. \tag{401}$$

Since D is diagonal, we now have two uncoupled 1st order linear differential equations, which are straightforward to solve:

$$\mathbf{u}(t) = \begin{bmatrix} e^{-i\omega t} & \\ & e^{i\omega t} \end{bmatrix} \mathbf{u}(0) \tag{402}$$

where $\mathbf{u}(0)$ are the "initial conditions" at $t = 0$, and $\omega = qB_z/m$. Of course, we are interested in $\mathbf{v}(t)$, so let us put this solution back in these original variables:

$$U^\dagger \mathbf{v}(t) = \begin{bmatrix} e^{-i\omega t} & \\ & e^{i\omega t} \end{bmatrix} U^\dagger \mathbf{v}(0). \tag{403}$$

Equivalently:

$$\mathbf{v}(t) = U \begin{bmatrix} e^{-i\omega t} & \\ & e^{i\omega t} \end{bmatrix} U^\dagger \mathbf{v}(0) \tag{404}$$

$$= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} e^{-i\omega t} & \\ & e^{i\omega t} \end{bmatrix} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \mathbf{v}(0) \tag{405}$$

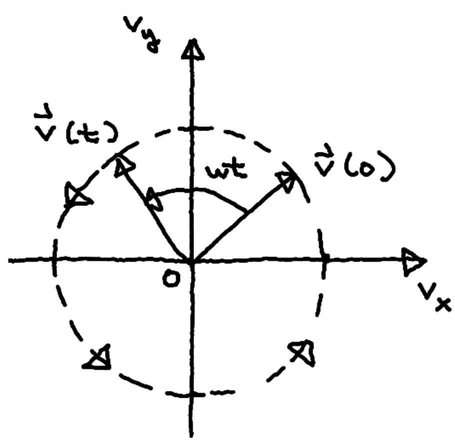
$$= \frac{1}{2} \begin{bmatrix} e^{-i\omega t} & e^{i\omega t} \\ ie^{-i\omega t} & -ie^{i\omega t} \end{bmatrix} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \mathbf{v}(0) \tag{406}$$

$$= \frac{1}{2} \begin{bmatrix} e^{-i\omega t} + e^{i\omega t} & -ie^{-i\omega t} + ie^{i\omega t} \\ ie^{-i\omega t} - ie^{i\omega t} & e^{-i\omega t} + e^{i\omega t} \end{bmatrix} \mathbf{v}(0). \tag{407}$$

Recognizing that $e^{i\omega t} + e^{-i\omega t} = 2\cos(\omega t)$ and $-i(e^{i\omega t} - e^{-i\omega t}) = 2\sin(\omega t)$, permits the simplification:

$$\mathbf{v}(t) = \begin{bmatrix} \cos(\omega t) & -\sin(\omega t) \\ \sin(\omega t) & \cos(\omega t) \end{bmatrix} \mathbf{v}(0). \tag{408}$$

The matrix that appears here is familiar: it appeared in Eq. 168, describing the active rotation of points in two spatial dimensions about the origin. It has an analogous role here: as time proceeds, the velocity vector traces out a circle in "velocity space", centered around zero velocity.



To obtain the corresponding position vector, we integrate Eq. 408, giving:

$$\mathbf{r}(t) = \frac{1}{\omega} \begin{bmatrix} \sin(\omega t) & \cos(\omega t) \\ -\cos(\omega t) & \sin(\omega t) \end{bmatrix} \mathbf{v}(0) + \mathbf{c} \tag{409}$$

where $\mathbf{c}$ is an arbitrary vector determined by the initial conditions; i.e.,

$$\mathbf{r}(0) = \frac{1}{\omega} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{v}(0) + \mathbf{c} \tag{410}$$

which we may rearrange to obtain:

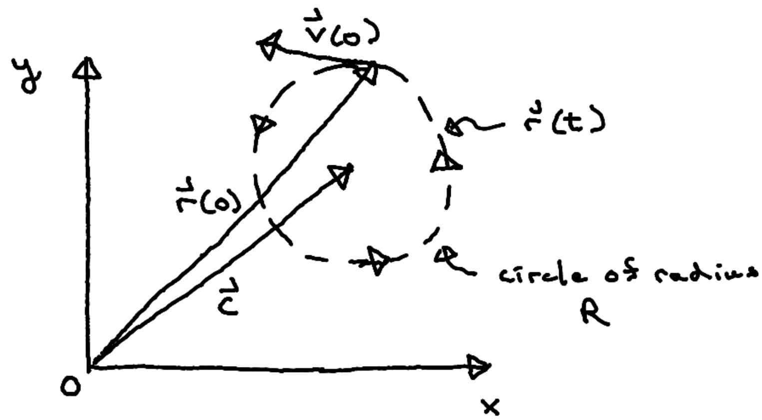
$$\mathbf{c} = \mathbf{r}(0) - \frac{1}{\omega} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{v}(0) \tag{411}$$

so that

$$\mathbf{r}(t) = \mathbf{r}(0) + \frac{1}{\omega} \begin{bmatrix} \sin(\omega t) & \cos(\omega t) - 1 \\ -\cos(\omega t) + 1 & \sin(\omega t) \end{bmatrix} \mathbf{v}(0) \tag{412}$$

describing circular orbits of radius:

$$\boxed{R = \frac{|\mathbf{v}(0)|}{\omega}} \tag{413}$$



This "circling" behaviour is referred to as **cyclotron motion** and the corresponding frequency $f = \omega/2\pi$, or equivalently

$$\boxed{f = \frac{1}{2\pi}\frac{qB}{m}} \tag{414}$$

is known as the **cyclotron frequency**.

It is quite convenient that the cyclotron frequency is independent of the speed/kinetic energy of the particle — this independence is the basis for so-called **cyclotrons**: early high energy particle accelerators.[49] However, as the energy of the particle increases and its motion becomes relativistic, the cyclotron frequency *decreases*. (Recall that we started this section with the non-relativistic form of Lorentz's force law.) This dependence limits the particle energies available from cyclotrons and different types of accelerators must be used to produce particles with higher energies.

[49] Some details and the early history of cyclotrons are given in M. S. Livingston, "Part i, History of the cyclotron", Phys. Today **12**, 18–23 (1959). The biography: L. W. Alvarez, *Alvarez: adventures of a physicist* (Basic Books, New York, 1987) contains some interesting personal recollections from the early days of cyclotrons.

D Formulae summary

Maxwell's equations

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0 \tag{D.1}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{D.2}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{D.3}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{D.4}$$

Potentials

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A} \tag{D.5}$$

$$\mathbf{E} = -\boldsymbol{\nabla} \phi - \frac{\partial \mathbf{A}}{\partial t} \tag{D.6}$$

Gauge conditions

$$\boldsymbol{\nabla} \cdot \mathbf{A} = 0 \qquad \text{Coulomb} \tag{D.7}$$

$$\boldsymbol{\nabla} \cdot \mathbf{A} = -\frac{1}{c^2}\frac{\partial \phi}{\partial t} \qquad \text{Lorenz} \tag{D.8}$$

Conservation laws

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} \tag{D.9}$$

$$\mathbf{J} \cdot \mathbf{E} + \frac{\partial}{\partial t}\left(\frac{\epsilon_0}{2}E^2 + \frac{1}{2\mu_0}B^2\right) = -\nabla \cdot \mathbf{S} \tag{D.10}$$

$$\mathbf{S} \coloneqq \frac{1}{\mu_0}\mathbf{E} \times \mathbf{B} \tag{D.11}$$

$$\mathbf{f} = \nabla \cdot \overleftrightarrow{T} - \epsilon_0 \mu_0 \frac{\partial \mathbf{S}}{\partial t} \tag{D.12}$$

$$\begin{aligned} T_{ij} \coloneqq \epsilon_0 &\left(E_i E_j - \frac{1}{2}\delta_{i,j}E^2\right) \\ + \frac{1}{\mu_0} &\left(B_i B_j - \frac{1}{2}\delta_{i,j}B^2\right) \end{aligned} \tag{D.13}$$

Coulomb's Law

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\int d\tau' \frac{\rho(\mathbf{r}')}{\mathscr{r}^2}\hat{\boldsymbol{\mathscr{r}}} \tag{D.14}$$

Biot-Savart

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} I \oint \frac{d\boldsymbol{\ell}' \times \hat{\boldsymbol{\mathscr{r}}}}{\mathscr{r}^2} \tag{D.15}$$

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi}\int \frac{\mathbf{J}(\mathbf{r}') \times \hat{\boldsymbol{\mathscr{r}}}}{\mathscr{r}^2}\, d\tau' \tag{D.16}$$

Auxiliary fields in materials

$$\mathbf{D} \coloneqq \epsilon_0 \mathbf{E} + \mathbf{P} \tag{D.17}$$

$$\mathbf{H} \coloneqq \mathbf{B}/\mu_0 - \mathbf{M} \tag{D.18}$$

$$\rho_b = -\nabla \cdot \mathbf{P} \tag{D.19}$$

$$\sigma_b = \mathbf{P} \cdot \mathbf{n} \tag{D.20}$$

$$\mathbf{J}_b = \nabla \times \mathbf{M} + \frac{\partial \mathbf{P}}{\partial t} \tag{D.21}$$

$$\mathbf{K}_b = \mathbf{M} \times \mathbf{n} \tag{D.22}$$

Materials with linear response

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E} \tag{D.23}$$

$$\epsilon = \epsilon_0(1+\chi) \tag{D.24}$$

$$u_{\mathrm{EM}}(t) \approx \frac{1}{2}\left\{\frac{\partial}{\partial\omega}\left[\omega\tilde{\epsilon}\,'(\omega)\right]|\mathbf{E}(t)|^2 + \frac{\partial}{\partial\omega}\left[\omega\tilde{\mu}\,'(\omega)\right]|\mathbf{H}(t)|^2\right\} \tag{D.25}$$

$$Q_E(t) \approx \omega\epsilon''(\omega)|\mathbf{E}(t)|^2 \tag{D.26}$$

$$\mathbf{P}(t) = \epsilon_0 \int_{-\infty}^{\infty} dt'\,\chi(t-t')\,\mathbf{E}(t') \tag{D.27}$$

$$\check{\chi}(\omega) = \int_{-\infty}^{\infty} dt\,\chi(t)\,e^{i\omega t} \tag{D.28}$$

$$\chi(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\,\check{\chi}(\omega)e^{-i\omega t} \tag{D.29}$$

$$\check{\mathbf{P}}(\omega) = \epsilon_0\check{\chi}(\omega)\check{\mathbf{E}}(\omega) \tag{D.30}$$

$$\check{\chi}'(\omega) = -\frac{1}{\pi}\,P\int_{-\infty}^{\infty} ds\,\frac{\check{\chi}''(s)}{\omega-s}$$

$$\check{\chi}''(\omega) = \frac{1}{\pi}\,P\int_{-\infty}^{\infty} ds\,\frac{\check{\chi}'(s)}{\omega-s}.$$

Retardation and Radiation

$$\mathbf{A}_{\mathrm{rad}}(\mathbf{r},t) = \frac{\mu_0}{4\pi r}\int d\tau'\mathbf{J}\left(\mathbf{r}', t-\frac{r}{c}+\frac{r'\cos\gamma}{c}\right) \tag{D.31}$$

$$c\,\mathbf{B}_{\mathrm{rad}}(\mathbf{r},t) = -\hat{\mathbf{r}}\times\frac{\partial\mathbf{A}_{\mathrm{rad}}(\mathbf{r},t)}{\partial t} \tag{D.32}$$

$$\mathbf{E}_{\mathrm{rad}}(\mathbf{r},t) = -\hat{\mathbf{r}}\times c\,\mathbf{B}_{\mathrm{rad}}(\mathbf{r},t) \tag{D.33}$$

$$\frac{dP_{\mathrm{Larmor}}}{d\Omega} = \frac{\mu_0}{c(4\pi)^2}\;q^2a_{\mathrm{ret}}^2\sin^2\theta \tag{D.34}$$

$$P_{\mathrm{Larmor}} = \frac{q^2a_{\mathrm{ret}}^2}{6\pi\epsilon_0c^3} \tag{D.35}$$

Special Relativity

$$\beta \coloneqq v/c \tag{D.36}$$

$$\gamma \coloneqq \frac{1}{\sqrt{1-(v/c)^2}} \tag{D.37}$$

$$\tanh\theta \coloneqq v/c \tag{D.38}$$

$$J^\alpha \equiv (\rho c, \mathbf{J}) \tag{D.39}$$

$$A^\alpha \equiv (\varphi/c, \mathbf{A}) \tag{D.40}$$

$$\partial_i \equiv (\frac{1}{c}\partial_t, \boldsymbol{\nabla}) \tag{D.41}$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} \cosh\theta & -\sinh\theta & 0 & 0 \\ -\sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{D.42}$$

In the following μ and ν correspond to the rows and columns of the corresponding arrays.

$$F_{\mu\nu} \equiv \left\{\begin{array}{cccc} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{array}\right\} \tag{D.43}$$

$$F^{\mu}{}_{\nu} \equiv \left\{ \begin{array}{cccc} 0 & E_x/c & E_y/c & E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{array} \right\} \quad \text{(D.44)}$$

$$F_{\mu}{}^{\nu} \equiv \left\{ \begin{array}{cccc} 0 & -E_x/c & -E_y/c & -E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{array} \right\} \quad \text{(D.45)}$$

$$F^{\mu\nu} \equiv \left\{ \begin{array}{cccc} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{array} \right\} \quad \text{(D.46)}$$

Transformation of fields

$$\mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel} \qquad \mathbf{B}'_{\parallel} = \mathbf{B}_{\parallel} \quad \text{(D.47)}$$

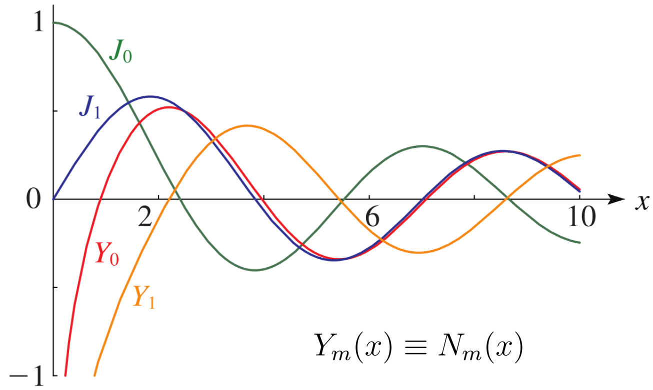
$$\mathbf{E}'_{\perp} = \gamma(\mathbf{E} + \boldsymbol{\beta} \times c\mathbf{B})_{\perp} \quad \text{(D.48)}$$

$$c\mathbf{B}'_{\perp} = \gamma(c\mathbf{B} - \boldsymbol{\beta} \times \mathbf{E})_{\perp} \quad \text{(D.49)}$$

Bessel functions

$$\frac{d^2R}{dx^2} + \frac{1}{x}\frac{dR}{dx} + \left[1 - \frac{m^2}{x^2}\right] R = 0 \quad \text{(D.50)}$$

From the DLMF:



$$H_m^{(1)}(x) = J_m(x) + iN_m(x) \quad \text{(D.51)}$$

$$H_m^{(2)}(x) = J_m(x) - iN_m(x) \quad \text{(D.52)}$$

$$H_m^{(1)}(x) \sim \sqrt{\frac{2}{\pi x}}\; e^{i(x - \frac{1}{2}m\pi - \frac{1}{4}\pi)} \quad \text{(D.53)}$$

$$H_m^{(2)}(x) \sim \sqrt{\frac{2}{\pi x}}\; e^{-i(x - \frac{1}{2}m\pi - \frac{1}{4}\pi)} \quad \text{(D.54)}$$

$$e^{iks\cos\phi} = \sum_{m=-\infty}^{m=\infty} i^m J_m(ks) e^{im\phi} \quad \text{(D.55)}$$

Helmholtz Decomposition

$$\mathbf{C}(\mathbf{r}) = \nabla \times \mathbf{F}(\mathbf{r}) - \nabla\Omega(\mathbf{r}) \quad \text{(D.56)}$$

$$\Omega(\mathbf{r}) = \frac{1}{4\pi}\int d\tau' \,\frac{\nabla' \cdot \mathbf{C}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad \text{(D.57)}$$

$$\mathbf{F}(\mathbf{r}) = \frac{1}{4\pi}\int d\tau' \,\frac{\nabla' \times \mathbf{C}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad \text{(D.58)}$$

Miscellaneous results

$$\delta\left(x - a\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(x-a)t}\mathrm{d}t. \quad \text{(D.59)}$$

$$\nabla \frac{1}{\mathscr{r}} = -\frac{\hat{\mathscr{r}}}{\mathscr{r}^2} \tag{D.60}$$

$$\nabla \cdot \frac{\hat{\mathscr{r}}}{\mathscr{r}^2} = 4\pi\delta^3(\mathscr{r}) \tag{D.61}$$

$$\nabla^2 \frac{1}{\mathscr{r}} = -4\pi\delta^3(\mathscr{r}) \tag{D.62}$$

$$\cosh^2\theta - \sinh^2\theta = 1 \tag{D.63}$$

$$\mu_0 = 4\pi \times 10^{-7}\,\mathrm{H/m} \tag{D.64}$$

$$\epsilon_0 \approx 8.854187 \times 10^{-12}\,\mathrm{F/m} \tag{D.65}$$

$$c = 2.99792458 \times 10^{8}\,\mathrm{m/s} \tag{D.66}$$

$$c^2 = \frac{1}{\mu_0\epsilon_0} \tag{D.67}$$

E Comments on additional resources

First let me warn you that there are a variety of differing conventions in electromagnetism that you should be conscious of when consulting other sources. As an example, in the so-called **Gaussian unit system**, the electric and magnetic fields are measured with the *same* units. And the first of Maxwell's equations is $\nabla \cdot \vec{E} = 4\pi\rho$, not the familiar $\nabla \cdot \vec{E} = \rho/\epsilon_0$ that we are used to. I stick to GITE4's choice throughout these notes (SI). Do not be surprised if you find formulae in other sources that look slightly different. Further information may be found in Electromagnetic unit systems demystified.

A textbook with much physical insight, but at a lower level than GITE4, is Purcell's Electricity and Magnetism [10]. I really like this book. Due to the conditions under which it was originally published, it is available here.

Apart from their use of the Gaussian unit system, Heald and Marion's *Classical electromagnetic radiation* [22] would be a suitable text for this course. It is available for loan here, and is relatively inexpensive to purchase. There is an accompanying solutions manual freely provided by the publisher.

The canonical graduate textbook in electromagnetism is Jackson's *Classical electrodynamics* [9]. Jackson's appendix gives a thorough discussion of the differing unit systems in electromagnetism.

An interesting, high-quality alternative to Jackson has appeared in recent years: Zangwill's *Modern electrodynamics* [8].

I mentioned in the introduction that Landau and Lifshitz's *Classical theory of fields* [2] is inspirational, but impractical as a textbook. There is an accompanying volume concentrating on electromagnetism in materials [23], which is equally interesting.

Two books on special relativity that I like and think are appropriate for the level of this course are Williams [24] and Steane [13].

Moving away from the realm of traditional textbooks: students have told me that they have found `http://physicspages.com/` useful. These are notes based on the author's study of various standard physics textbooks, including Griffiths. I have not looked at them in too much detail, but I want to call your attention to their existence.

Just as it is unwise to take medical advice based on the results of google searches,[50] you should not use mathematical formulae of questionable provenance. Use authoritative standard sources for any mathematical results that you need. For the type of mathematics needed for the physics of electromagnetism, the NIST Digital Library of Mathematical Functions is the gold standard, building on its predecessor "Abramowitz and Stegun" (A&S). Physical book versions of both the DLMF and A&S are available [25, 26] — in this format I like the older A&S a bit more than DLMF, due to some sort of combination of aesthetics, usability, and nostalgia.

For the fundamental physical constants, the web-site:

`https://physics.nist.gov/cuu/Constants/`

is authoritative, and is kept up to date with the widely accepted CODATA values [27].

A useful source of information regarding the elementary particles is the "Review of Particle Physics" [28], available on-line.

[50] Even the usually reliable wikipedia should never be fully trusted: I was recently surprised to find some rather dubious claims in an entry. No, it was not about something important or controversial, but instead regarding the rather mundane virial expansion: a systematic way of accounting for corrections to the ideal gas law.

F Video pages

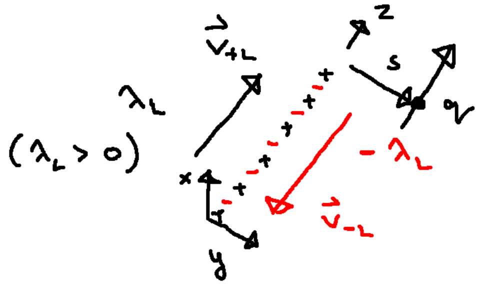
F.1 Video pages for length contraction in electromagnetism

1

Length contraction in electromagnetism

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$$

Covariance: laws "look the same" in different inertial reference frames.



No $\vec{E}$ field. $\vec{E}_L = \vec{0}$

$$\vec{B}_L = \frac{\mu_0}{2\pi}\frac{1}{s} I \hat{\phi}$$

$$I = 2\lambda_L v_{+L,z}$$

$$\frac{C}{m} \quad \frac{m}{s}$$

2

λ_0: line charge density in rest frame

$$\lambda_0 = \frac{\lambda_L}{\cosh\theta_{+L,z}}$$

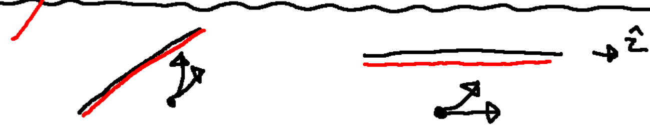
charges are observed to be "squeezed" together in lab frame.

⇒ force on charge q initially moving ∥ to wire

$$\vec{F}_L = q\,\vec{v} \times \vec{B}$$

$$= q\,\hat{z}\, v_{PL,z} \times \frac{\mu_0}{2\pi}\frac{1}{s}\lambda_0 \cosh\theta_{+L,z}\, 2 v_{+L,z}\,\hat{\phi}$$

$$\boxed{\vec{F}_L = \underbrace{\hat{z}\times\hat{\phi}}_{-\hat{s}}\; q\, v_{PL,z}\, \frac{\mu_0}{2\pi}\frac{1}{s}\, \lambda_0 \cosh\theta_{+L,z}\, 2 v_{+L,z}}$$



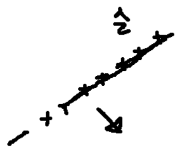
3

Now consider problem in the rest frame of particle P, where $\vec{v} \times \vec{B} = \vec{0}$.

Now the two charge densities experience different length contractions and no longer cancel
$\vec{E}_p \neq \vec{0}$

$$\lambda_{+p} = \cosh(\theta_{+P,z})\,\lambda_0$$

Gauss' Law $\qquad \vec{E} = \hat{s}\,\frac{\lambda}{2\pi\epsilon_0}\,\frac{1}{s}$



$$\vec{E}_p = \hat{s}\,\frac{1}{2\pi\epsilon_0}\,\frac{1}{s}\,\lambda_0\left(\overbrace{\cosh(\underbrace{\theta_{+P,z}}_{})}^{+ve} - \overbrace{\cosh(\underbrace{\theta_{-P,z}}_{})}^{-ve}\right)$$

4

$$\theta_{+P,z} = \theta_{+L,z} - \theta_{PL,z}$$
$$\theta_{-P,z} = \theta_{-L,z} - \theta_{PL,z}$$

$$\vec{E}_p = \hat{s}\,\frac{1}{2\pi\epsilon_0}\,\frac{\lambda_0}{s}\left[\underbrace{\cosh(\theta_{+L,z} - \theta_{PL,z})}_{} - \cosh(\underbrace{\theta_{-L,z}}_{-\theta_{+L,z}} - \theta_{PL,z})\right]$$

Mathematical identity $\quad \cosh(u+v) = \cosh u \cosh v + \sinh u \sinh v$

$$= \hat{s}\,\frac{1}{2\pi\epsilon_0}\,\frac{\lambda_0}{s}\Big[\underbrace{\cosh(\theta_{+L,z})\cosh(\theta_{PL,z}) - \sinh(\theta_{+L,z})\sinh(\theta_{PL,z})}_{''} - \underbrace{\qquad\qquad}_{-}\Big]$$

$$\vec{E}_p = -\hat{s}\,\frac{1}{2\pi\epsilon_0}\,\frac{\lambda_0}{s}\,2\,\underbrace{\sinh\theta_{+L,z}}_{\frac{v_{+L,z}}{c}\cosh\theta_{+L,z}}\;\underbrace{\sinh\theta_{PL,z}}_{\frac{v_{PL,z}}{c}\cosh\theta_{PL,z}}$$

$$c^2 = \frac{1}{\mu_0\epsilon_0} \Rightarrow \frac{1}{c^2\epsilon_0} = \mu_0$$

5

$$\vec F_p = q(\vec E_p + \underbrace{\overbrace{\vec v_p}^{=\vec 0} \times \vec B_p}_{=\vec 0})$$

$$\boxed{\vec F_p = -\hat s\, q\, \frac{\mu_0}{2\pi}\, \frac{1}{s}\, \lambda_0\, 2 v_{PL,z} V_{+L,z} \cosh(\Theta_{+L,z}) \underbrace{\cosh(\Theta_{PL,z})}_{\text{transformation of force.}}}$$

for non-relativistic point particle motion

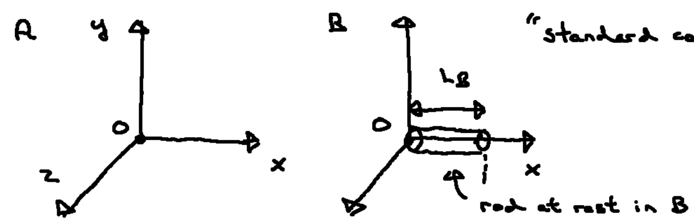
$\cosh(\Theta_{PL,z}) \to 1$

and $\vec F_p \to \vec F_L$

F.2 Video pages for Lorentz transformations

1

The Lorentz transformations



"standard configuration"

$$\vec v_{BA} = \underbrace{v_{BA,x}}\, \hat x$$

Take $x_A = y_A = z_A = t_A = 0$ to be the same point in spacetime as

$x_B = y_B = z_B = t_B = 0$

By length contraction: $L_A = \overset{\sim x_B}{L_B} / \underbrace{\cosh\Theta_{BA,x}}$

$$x_A = t_A v_{BA,x} + \frac{x_B}{\cosh\Theta_{BA,x}} \qquad (\to c\tanh\Theta_{BA,x})$$

②

$$\boxed{x_B = x_A \cosh \theta_{BA,x} - ct_A \sinh \theta_{BA,x}} \qquad (*)$$

Now we want t_B in terms of x_A and t_A

By symmetry:

$$x_A = x_B \cosh \theta_{AB,x} - ct_B \sinh \theta_{AB,x}$$

$$\vec{V}_{AB} = -\vec{V}_{BA} \quad \Rightarrow \quad \theta_{AB,x} = -\theta_{BA,x}$$

$$x_A = x_B \cosh \theta_{BA,x} + ct_B \sinh \theta_{BA,x} \qquad (\ddagger)$$

$$x_B = \frac{1}{\cosh \theta_{BA,x}} \left(x_A - ct_B \sinh \theta_{BA,x} \right)$$ & equate to $(*)$

③

$$\frac{1}{\cosh(\theta_{BA,x})} \left(x_A - ct_B \sinh \theta_{BA,x} \right) = x_A \cosh \theta_{BA,x} - ct_A \sinh \theta_{BA,x}$$

Rearrange to solve for t_B:

$$-ct_B \sinh \theta_{BA,x} = -x_A + \cosh \theta_{BA,x} \left(x_A \cosh \theta_{BA,x} - ct_A \sinh \theta_{BA,x} \right)$$

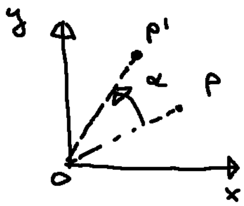
$$ct_B = \frac{1}{\sinh \theta_{BA,x}} \left[x_A - x_A \cosh^2 \theta_{BA,x} + ct_A \cosh \theta_{BA,x} \sinh \theta_{BA,x} \right]$$

analogous to $\cos^2\theta + \sin^2\theta = 1$: $\cosh^2\theta - \sinh^2\theta = 1$

$$\boxed{ct_B = -x_A \sinh \theta_{BA,x} + ct_A \cosh \theta_{BA,x}}$$

4

$$\begin{bmatrix} ct_B \\ x_B \end{bmatrix} = \underbrace{\begin{bmatrix} \cosh\Theta_{BA,x} & -\sinh\Theta_{BA,x} \\ -\sinh\Theta_{BA,x} & \cosh\Theta_{BA,x} \end{bmatrix}}_{L(\Theta)} \begin{bmatrix} ct_A \\ x_A \end{bmatrix}$$



$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \underbrace{\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}}_{=:\,R(\alpha)} \begin{bmatrix} x \\ y \end{bmatrix}$$

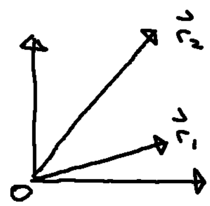
active transformation

$\rightsquigarrow$ $R^T(\alpha)R(\alpha) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $R(\alpha)$ orthogonal matrices

$$x'^2 + y'^2 = [x' \; y'] \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix}^T \begin{bmatrix} x' \\ y' \end{bmatrix} = \left(R(\alpha) \begin{bmatrix} x \\ y \end{bmatrix} \right)^T R(\alpha) \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \begin{bmatrix} x \\ y \end{bmatrix}^T \underbrace{R(\alpha)^T R(\alpha)}_{I} \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + y^2 \quad \checkmark$$

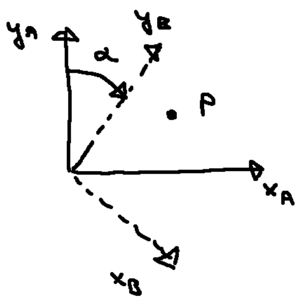
5



rotations preserve the dot product

$$\overbrace{x_1 x_2 + y_1 y_2} = \overbrace{\vec{r}_1 \cdot \vec{r}_2}$$

is invariant under rotations about O



$$\underbrace{\begin{bmatrix} x_B \\ y_B \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x_A \\ y_A \end{bmatrix}}_{\text{passive transformation}}$$

Lorentz transformations are passive.

6

$\underbrace{ct_1 ct_2 - x_1 x_2}_{\text{Lorentz scalar product}}$ is a Lorentz invariant.

$$
\begin{aligned}
& ct_{1B}ct_{2B} - x_{1B}x_{2B} \\
&= \begin{bmatrix} ct_{1B} & x_{1B} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} ct_{2B} \\ x_{2B} \end{bmatrix} \\
&= \left(L(\theta) \begin{bmatrix} ct_{1A} \\ x_{1A} \end{bmatrix} \right)^T \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \left(L(\theta) \begin{bmatrix} ct_{2A} \\ x_{2A} \end{bmatrix} \right) \\
&= \begin{bmatrix} ct_{1A} & x_{1A} \end{bmatrix} \underbrace{L(\theta)^T \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} L(\theta)}_{= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}} \begin{bmatrix} ct_{2A} \\ x_{2A} \end{bmatrix} \\
&= \underbrace{ct_{1A} ct_{2A} - x_{1A} x_{2A}}
\end{aligned}
$$

F.3 Video pages for four vectors and tensors

1

Four-vectors and tensors

$$
\left.
\begin{aligned}
\vec{E}(x,t) &= E_0 \cos(kx - \omega t)\,\hat{y} \\
\vec{B}(x,t) &= \frac{E_0}{c} \cos(\underbrace{kx - \omega t}_{\phi})\,\hat{z}
\end{aligned}
\right\}
$$

"linearly polarized plane EM wave" Eq. 9.48 GITEY

$\phi = k_B x_B - \frac{\omega_B}{c} ct_B$

frame B wrt A $\vec{v}_{BA} = v_{BA,x}\,\hat{x}$

$\tanh\theta = \frac{v_{AB,x}}{c}$

$$
\begin{aligned}
&= k_B (x_A \cosh\theta - ct_A \sinh\theta) \\
&\quad - \frac{\omega_B}{c} \left(- x_A \sinh\theta + ct_A \cosh\theta \right) \\
&= x_A \underbrace{\left(k_B \cosh\theta + \frac{\omega_B}{c} \sinh\theta \right)}_{=: k_A} - ct_A \underbrace{\left(k_B \sinh\theta + \frac{\omega_B}{c} \cosh\theta \right)}_{=: \frac{\omega_A}{c}}
\end{aligned}
$$

2

$$k_B = k_A \cosh\theta - \frac{\omega_A}{c}\sinh\theta$$

$$\frac{\omega_B}{c} = -k_A\sinh\theta + \frac{\omega_A}{c}\cosh\theta$$

same transformation as for (ct, x)

$(ct, \vec{r})$ and $(\frac{\omega}{c}, \vec{k})$ transform between inertial frames in the same way.

four-vectors

$(\rho c, \vec{J})$

$$\phi = \vec{k}\cdot\vec{r} - \omega t = k_x x + k_y y + k_z z - \frac{\omega}{c}ct$$

$$= \left(\frac{\omega}{c}, \vec{k}\right) \cdot_L \left(ct, \vec{r}\right)$$

Lorentz invariant scalar product.

3

$$\left(\frac{\omega}{c}, \vec{k}\right) \cdot_L \left(ct, \vec{r}\right)$$

$$= \underbrace{\left[\frac{\omega}{c}, k_x, k_y, k_z\right]\begin{bmatrix}-1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1\end{bmatrix}}_{k_i}\underbrace{\begin{bmatrix}ct\\ x\\ y\\ z\end{bmatrix}}_{r^i}$$

$k_0 = -\frac{\omega}{c}$, $k_1 = k_x$, $k_2 = k_y$, $k_3 = k_z$

$r^0 = ct$, $r^1 = x$, $r^2 = y$, $r^3 = z$

$$= \sum k_i r^i$$

$$= k_i r^i$$

summation over indices repeated in "top and bottom" are implicit

Einstein summation convention

(4)

Λ is a matrix that describes Lorentz transformations (combinations of rotations and boosts)

$$\Lambda = \begin{bmatrix} c & -s & 0 & 0 \\ -s & c & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \Lambda = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & s \\ 0 & 0 & -s & c \end{bmatrix} \qquad K_\alpha r^\alpha = K^\alpha r_\alpha$$

$$g = \begin{bmatrix} -1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{bmatrix} \qquad \Lambda^T g \Lambda = g$$

covariant components of 4-vector: $(ct, \vec{r})$

r^α want r_α $\qquad r_\beta = g_{\beta\alpha} r^\alpha$

contravariant components.

$\tilde{r}^\beta = \Lambda^\beta{}_\alpha r^\alpha$

new coordinate system

old coordinate system

(5)

How do covariant components transform between inertial reference frames?

$$\underbrace{(\Lambda^{-1})^\alpha{}_\gamma \Lambda^\beta{}_\alpha}_{} = \delta^\beta{}_\gamma$$

$= 0$ unless $\gamma = \beta$ in which case it's one.

Kronecker delta

$A_\alpha B^\alpha = \tilde{A}_\gamma \tilde{B}^\gamma$ ← should be true

$$\tilde{A}_\alpha \tilde{B}^\alpha = \tilde{A}_\alpha \Lambda^\alpha{}_\gamma B^\gamma = (\Lambda)^{-1\,\nu}{}_\alpha A_\nu \Lambda^\alpha{}_\gamma B^\gamma$$

if $\tilde{A}_\alpha = (\Lambda)^{-1\,\nu}{}_\alpha A_\nu$

$$= \underbrace{\delta^\nu{}_\gamma A_\nu}_{} B^\gamma$$

$$= A_\gamma B^\gamma$$

6

$$\underset{\text{4-vector}}{A_\alpha} B^\beta =: \underbrace{T_\alpha{}^\beta}_{\text{rank 2 tensor}} \Big\}$$

$$\tilde{T}_\nu{}^\epsilon = (\Lambda^{-1})^\beta_\nu (\Lambda)^\epsilon_\alpha \, T_\beta{}^\alpha$$

rank 2 tensors $\quad T_{\alpha\beta} = g_{\alpha\nu} T_\beta{}^\nu$

$$\underbrace{A_\alpha{}^{\beta\nu}}\xrightarrow[\text{define contraction}]{} B^\nu = A_\alpha{}^{\alpha\nu}$$

$$A_\alpha B^\beta =: T_\alpha{}^\beta \qquad T_\alpha{}^\alpha = \underbrace{A_\alpha B^\alpha}$$

7

$$\partial_0 := \frac{1}{c}\frac{\partial}{\partial t}, \; \partial_1 := \frac{\partial}{\partial x}, \; \partial_2 := \frac{\partial}{\partial y}, \; \partial_3 = \frac{\partial}{\partial z}$$

$$\partial_i J^i = \frac{1}{c}\frac{\partial}{\partial t}(\rho c) + \underbrace{\frac{\partial}{\partial x}J_x + \frac{\partial}{\partial y}J_y + \frac{\partial J_z}{\partial z}}_{\nabla\cdot\vec{J}}$$

$$\underbrace{\partial_i J^i = 0}_{\frac{\partial \rho}{\partial t} = -\nabla\cdot\vec{J}} \Rightarrow \tilde{\partial}_i \tilde{J}^i = 0$$

$*$

F.4 Video pages for proper time and the terrible twins

1

Proper time and the terrible twins

- $r_\alpha r^\alpha$ $\qquad r^0 = ct,\ r^1 = x, \ldots$

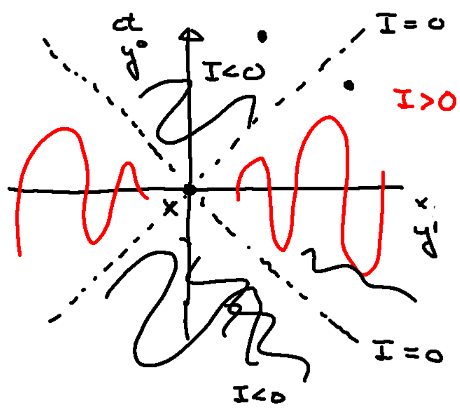
$$I := -(c\Delta t)^2 + \underbrace{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}_{(\Delta d)^2}$$ } invariant !!!

interval between two spacetime points

Between two spacetime events x and y

$$I = (x_\alpha - y_\alpha)(x^\alpha - y^\alpha)$$ } invariance obvious.

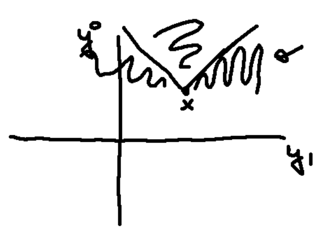
2



$I<0$: time-like

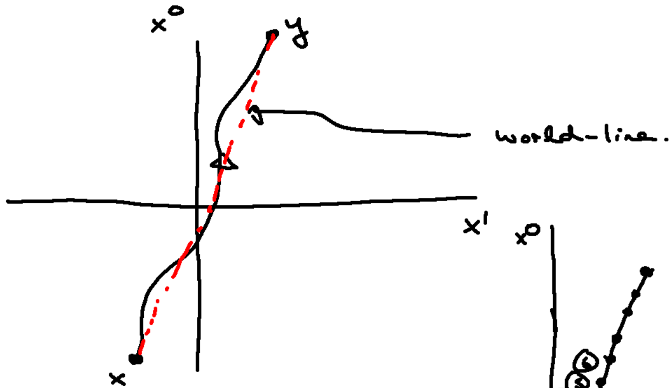
$I>0$: spacelike

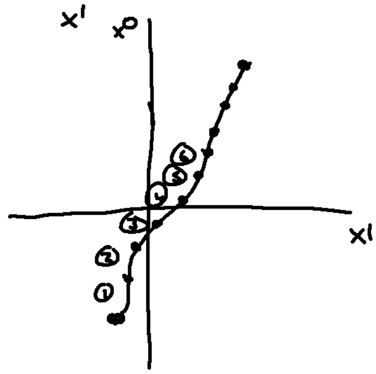
$I=0$: light like.



3

imagine a particle moving through spacetime





$\sum_i I_i$ is an invariant,

but depends on placement of points

4

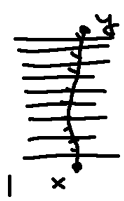
$$J = \underbrace{\sum_i \sqrt{\frac{-I_i}{?}}}$$

$$= \sum_i \sqrt{(c\Delta t_i)^2 - (\Delta d_i)^2}$$

$$= c \sum_i \Delta t_i \sqrt{1 - \left(\frac{\Delta d_i}{c \Delta t_i}\right)^2}$$

take limit as $\Delta t_i \to 0$

$$J = c \int_{x^0/c}^{y^0/c} \underbrace{dt \sqrt{1 - \left(\frac{v}{c}\right)^2}}_{\text{accumulation of time in rest-frame of clock.}} \quad \Big\} \text{ still invariant!!!}$$



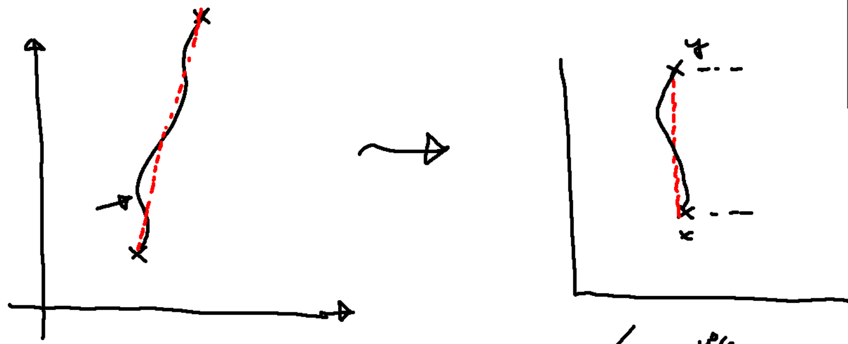
Remember "moving clocks run slow":

$$\underset{\substack{\uparrow \\ \text{rest frame of clock}}}{\Delta t_R} = \underbrace{\Delta t \sqrt{1 - \left(\frac{v}{c}\right)^2}}$$

"proper time"

$$\boxed{\underline{\underline{\tau}} = \int_{x^0/c}^{y^0/c} dt \sqrt{1 - \left(\frac{v}{c}\right)^2}}$$

5



$$\tau = \int_{x/c}^{y/c} dt = \Delta t$$

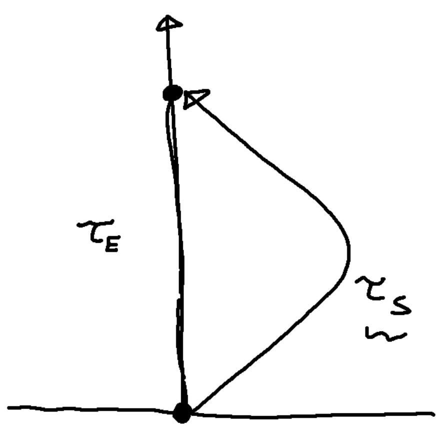
uniform motion

$$\tau = \int_{x/c}^{y/c} dt \underbrace{\sqrt{1-\left(\frac{v}{c}\right)^2}}_{<1}$$

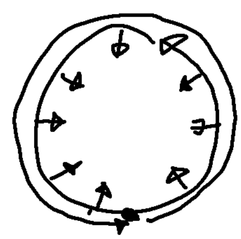
non-uniform motion

$$\underbrace{\tau_{\text{non-uniform}} < \tau_{\text{uniform}}}.$$

6



$$\tau_S < \tau_E$$



F.5 Video pages for momenergy from an invariant action

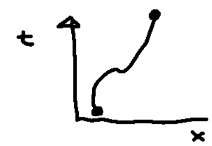
1

Momenergy from an invariant action

- based on principle of "most proper time"

$$S = -\alpha \int_{\text{world line}} d\tau$$

$\alpha > 0$



$$= -\alpha \left\{ \int dt \sqrt{1-(v/c)^2} \right\}$$

$$S = \int dt\, \mathcal{L}\,, \qquad \mathcal{L} = -\alpha\sqrt{1-(v/c)^2}$$

- for a non-relativistic particle: $\mathcal{L} = \frac{1}{2}mv^2$

Get α by requiring that $\mathcal{L}$ has this non-relativistic limit.

2

$$\mathcal{L} = -\alpha\left(1 - \frac{1}{2}\left(\frac{v}{c}\right)^2 + \cdots\right)$$

$$= -\alpha + \frac{\alpha}{2}\frac{v^2}{c^2}$$

doesn't affect motion since constant

$$\frac{\alpha}{\cancel{2}}\frac{1}{c^2} = \frac{m}{\cancel{2}} \Rightarrow \alpha = mc^2$$

$$\mathcal{L} = -mc^2\sqrt{1-\left(\frac{v}{c}\right)^2}$$

$$p_x := \frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{\partial \mathcal{L}}{\partial v_x} = \frac{\partial}{\partial v_x}(-mc^2)\sqrt{1-(v/c)^2}$$

$$= \frac{-mc^2}{2\sqrt{1-\left(\frac{v}{c}\right)^2}}\left(-\frac{1}{c^2}\right)\frac{\partial}{\partial v_x}\left(v_x^2 + v_y^2 + v_z^2\right)$$

3

$$\boxed{\vec{p} = \frac{m\vec{v}}{\sqrt{1-(v/c)^2}}}$$

$$\frac{d\mathcal{L}}{dt} = \sum_i \frac{\partial \mathcal{L}}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\ddot{q}_i$$

$$\frac{d}{dt}\left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right] = \frac{\partial \mathcal{L}}{\partial q_i}$$

$$= \sum_i \frac{d}{dt}\left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right]\dot{q}_i + \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\ddot{q}_i$$

$$= \sum_i \frac{d}{dt}\left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\dot{q}_i\right] \qquad p_i \qquad E := -L + \sum_i p_i \dot{q}_i$$

$$\frac{d}{dt}\Big[\underbrace{L - \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\dot{q}_i}_{E}\Big] = 0$$

4

$$E = mc^2\sqrt{1-\left(\frac{v}{c}\right)^2} + \frac{m}{\sqrt{1-\left(\frac{v}{c}\right)^2}}\underbrace{(v_x^2+v_y^2+v_z^2)}_{v^2}$$

$$= \frac{mc^2\left(1-(v/c)^2\right) + mv^2}{\sqrt{1-(v/c)^2}}$$

> 0

$$\boxed{E = \frac{mc^2}{\sqrt{1-(v/c)^2}}} \simeq mc^2 + \frac{1}{2}mv^2 + \ldots$$

$$\frac{dt}{d\tau} = \frac{1}{\sqrt{1-(v/c)^2}} \qquad \vec{p} = \frac{m\vec{v}}{\sqrt{1-\left(\frac{v}{c}\right)^2}} = m\frac{dt}{d\tau}\frac{d\vec{r}}{dt}$$

$$= m\frac{d\vec{r}}{d\tau}$$

5

$$E = mc^2 \frac{dt}{d\tau}$$

$$= mc \frac{d(ct)}{d\tau}$$

$$\left(\frac{E}{c}, \vec{p}\right) = m \underbrace{\frac{d}{d\tau}(ct, \vec{r})}_{\text{4-vector by construction}}$$

4-vector by construction, proper velocity

$$\boxed{\eta^\alpha := \frac{dx^\alpha}{d\tau}}, \quad \boxed{p^\alpha = m\eta^\alpha} \leftarrow \text{momenergy}$$

invariant.

Lorentz invariant

$$\eta_\alpha \eta^\alpha = -c\frac{dt}{d\tau}\, c\frac{dt}{d\tau} + \left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2 + \left(\frac{dz}{d\tau}\right)^2$$

$$= -c^2\left(\frac{dt}{d\tau}\right)^2 + \left(\frac{dt}{d\tau}\right)^2 v^2$$

6

$$\eta_\alpha \eta^\alpha = \frac{v^2 - c^2}{1-\left(\frac{v}{c}\right)^2}$$

$$\boxed{\eta_\alpha \eta^\alpha = -c^2}$$

$$p_\alpha p^\alpha = -m^2c^2$$

Drop tensor notation.

$$-\left(\frac{E}{c}\right)^2 + p^2 = -m^2c^2$$

$$E^2 = c^2p^2 + m^2c^4$$

$$\boxed{E = \sqrt{(cp)^2 + (mc^2)^2}}$$

F.6 Video pages for the Lorentz force law from an invariant action

①

The Lorentz Force Law from an invariant action

$$\int A_\alpha \, dx^\alpha$$

$$S = \underbrace{-mc^2 \int d\tau}_{\text{free particle}} + q \int A_\alpha \, dx^\alpha$$

$$\frac{d\tau}{dt} = \sqrt{1-\left(\frac{v}{c}\right)^2} \quad , \quad S = \int \mathcal{L} \, dt$$

$$\Rightarrow \mathcal{L} = -mc^2 \sqrt{1-(v/c)^2}$$

②

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

$$\int A_\alpha \, dx^\alpha = -\int A^0 c \, dt + \underbrace{\int \vec{A} \cdot d\vec{r}}_{\int \vec{A} \cdot \vec{v} \, dt}$$

$$\mathcal{L} = -mc^2 \sqrt{1-(v/c)^2} - qcA^0 + q\vec{A} \cdot \vec{v}$$

$$p_x = \frac{\partial \mathcal{L}}{\partial v_x} = \frac{mv_x}{\sqrt{1-(v/c)^2}} + qA_x$$

$$\vec{P} = \underbrace{\frac{m\vec{v}}{\sqrt{1-(v/c)^2}}} + q\vec{A}$$

$$\vec{P} = \vec{p} + q\vec{A}$$

canonical / generalized

ordinary / mechanical.

3

$$\frac{d}{dt}\left(\underbrace{\frac{\partial \mathcal{L}}{\partial \dot{q}_i}}\right) = \frac{\partial \mathcal{L}}{\partial q_i}$$

$$\frac{d}{dt}\left(\underbrace{\frac{m\vec{v}}{\sqrt{1-(v/c)^2}}}_{\vec{p}} + q\vec{A}\right) = -qc\,\nabla A^0 + q\nabla(\vec{A}\cdot\vec{v})$$

$$\frac{d\vec{p}}{dt} = -q\frac{d\vec{A}}{dt} - qc\,\nabla A^0 + q\nabla(\vec{A}\cdot\vec{v})$$

For arb. $\vec{a}$ and $\vec{b}$, we have:

$$\nabla(\vec{a}\cdot\vec{b}) = \vec{a}\times(\nabla\times\vec{b}) + \vec{b}\times(\nabla\times\vec{a}) + (\vec{a}\cdot\nabla)\vec{b} + (\vec{b}\cdot\nabla)\vec{a}$$

$$\nabla(\vec{A}\cdot\vec{v}) = \vec{v}\times(\nabla\times\vec{A}) + (\vec{v}\cdot\nabla)\vec{A}$$

4

$$\frac{d\vec{p}}{dt} = -q\frac{d\vec{A}}{dt} - qc\,\nabla A^0 + q\,\vec{v}\times(\nabla\times\vec{A}) + q\,\underbrace{(\vec{v}\cdot\nabla)\vec{A}}$$

convective derivative.

$$\frac{d\vec{A}}{dt} = \frac{\partial\vec{A}}{\partial t} + \underbrace{(\vec{v}\cdot\nabla)\vec{A}}$$

$$\frac{d\vec{p}}{dt} = -q\frac{\partial\vec{A}}{\partial t} - qc\,\nabla A^0 + q\,\vec{v}\times(\nabla\times\vec{A})$$

$$= q\left[\underbrace{\left(-c\nabla A^0 - \frac{\partial\vec{A}}{\partial t}\right)}_{\vec{E}\ (?)} + \vec{v}\times\underbrace{(\nabla\times\vec{A})}_{\vec{B}\ (?)}\right]$$

$$\boxed{\frac{d}{dt}\left[\frac{m\vec{v}}{\sqrt{1-(v/c)^2}}\right] = q\left[\vec{E} + \vec{v}\times\vec{B}\right]}$$

Bucherer's expt.

F.7 Video pages for the transformation of electric and magnetic fields between inertial frames

1

Transformations of the $\vec{E}$ and $\vec{B}$ fields between inertial frames

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

$$F_{01} = -E_x/c \;,\; F_{02} = -E_y/c \;,\; F_{03} = -E_z/c$$

$$\text{indep components} = \frac{4\times4-4}{2} = 6$$

$$F_{12} = \partial_1 \partial A_2 - \partial_2 A_1 \qquad F_{23} = B_x \;,\; F_{31} = B_y$$

$$= \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}$$

$$= \hat{z}\cdot(\nabla\times\vec{A})$$

$$= B_z$$

2

$$F_{\mu\nu} \equiv \begin{array}{cc} & \nu \\ \mu & \begin{Bmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{Bmatrix} \end{array}$$

$$F^{\mu\nu} \equiv \begin{Bmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{Bmatrix}$$

③

$$\tilde{F}^{\mu\beta} = \Lambda^{\mu}{}_{\alpha}\,\Lambda^{\beta}{}_{\kappa}\,F^{\alpha\kappa}$$

$$\Lambda = \begin{bmatrix} c\theta & -s\theta & & 0 \\ -s\theta & c\theta & & \\ & & 1 & \\ 0 & & & 1 \end{bmatrix}$$

standard configuration
$\vec{v} = \hat{x}\, c\tanh\theta$
← velocity of ~ frame

$$\begin{aligned}
\tilde{E}_x &= c\,\tilde{F}_{01} \\
&= c\,\Lambda^{0}{}_{\alpha}\Lambda^{1}{}_{\kappa}\,F^{\alpha\kappa} \\
&= c\,\Lambda^{0}{}_{\alpha}\left(\Lambda^{1}{}_{0}F^{\alpha 0} + \Lambda^{1}{}_{1}F^{\alpha 1} + \cancel{\Lambda^{1}{}_{2}}^{=0}F^{\alpha 2} + \cancel{\Lambda^{1}{}_{3}}^{=0}F^{\alpha 3}\right) \\
&= \cancel{c\Lambda^{0}{}_{0}\Lambda^{1}{}_{0}F^{00}} + c\Lambda^{0}{}_{1}\Lambda^{1}{}_{0}F^{10} + \cancel{c\Lambda^{0}{}_{2}\Lambda^{1}{}_{0}F^{20}} + \cancel{c\Lambda^{0}{}_{3}\Lambda^{1}{}_{0}F^{30}} \\
&\quad + c\Lambda^{0}{}_{0}\Lambda^{1}{}_{1}F^{01} + \cancel{c\Lambda^{0}{}_{1}\Lambda^{1}{}_{1}F^{11}} + \cancel{c\Lambda^{0}{}_{2}\Lambda^{1}{}_{1}}F^{21} + \cancel{c\Lambda^{0}{}_{3}\Lambda^{1}{}_{1}}F^{31} \\
&= c(-\sinh\theta)(-\sinh\theta)F^{10} + c(\cosh\theta)(\cosh\theta)(-F^{10})
\end{aligned}$$

④

$$\begin{aligned}
\tilde{E}_x &= c\,F^{10}\,\underbrace{(\sinh^2\theta - \cosh^2\theta)}_{-1} \\
&= c(-E_x/c)(-1) \\
&= E_x
\end{aligned}$$

$$\tilde{B}_x = B_x$$

$$\begin{aligned}
\tilde{E}_y &= c\,\hat{F}^{02} \\
&= c\,\Lambda^{0}{}_{\alpha}\Lambda^{2}{}_{\kappa}\,F^{\alpha\kappa} \\
&= c\,\Lambda^{0}{}_{\alpha}\left(\cancel{\Lambda^{2}{}_{0}}F^{\alpha 0} + \cancel{\Lambda^{2}{}_{1}}F^{\alpha 1} + \Lambda^{2}{}_{2}F^{\alpha 2} + \cancel{\Lambda^{2}{}_{3}}F^{\alpha 3}\right) \\
&= c\,\Lambda^{0}{}_{\alpha}F^{\alpha 2} \\
&= c\,\Lambda^{0}{}_{0}F^{02} + c\Lambda^{0}{}_{1}F^{12} + \cancel{c\Lambda^{0}{}_{2}F^{22}} + \cancel{c\Lambda^{0}{}_{3}}F^{32} \\
&= c\,\cosh\theta\,E_y/c + c(-\sinh\theta)\,B_z \\
&= \cosh\theta\,E_y - \sinh\theta\,cB_z
\end{aligned}$$

5

$$\tilde{E}_z = \cosh\theta\, E_z + \sinh\theta\, cB_y$$
$$\tilde{B}_y = \cosh\theta\, B_y + \sinh\theta\, E_z/c$$
$$\tilde{B}_z = \cosh\theta\, B_z - \sinh\theta\, E_y/z\ .$$

$$\vec{E} = \underbrace{(\hat{v}\cdot\vec{E})\,\hat{v}}_{\vec{E}_\parallel} + \underbrace{(\vec{E} - (\hat{v}\cdot E)\,\hat{v})}_{\vec{E}_\perp}$$

$$\vec{B} = \vec{B}_\parallel + \vec{B}_\perp$$

$$\tilde{\vec{E}}_\perp = \hat{y}\,(\cosh\theta\, E_y - \sinh\theta\, cB_z) + \hat{z}\,(\cosh\theta\, E_z + \sinh\theta\, cB_y)$$
$$= \cosh\theta\,(\hat{z}\,E_z + \hat{y}\,E_y) + \sinh\theta\, c\,(\hat{z}\,B_y - \hat{y}\,B_z)$$
$$= \cosh\theta\,(\hat{z}\,E_z + \hat{y}\,E_y) + \sinh\theta\, c\,(\hat{z}\,(\vec{B}\cdot\hat{y}) - \hat{y}\,(\vec{B}\cdot\hat{z}))$$

recall identity $\vec{a}\times(\vec{b}\times\vec{c}) = \vec{b}\,(\vec{a}\cdot\vec{c}) - \vec{c}\,(\vec{a}\cdot\vec{b})$

6

$$\tilde{\vec{E}}_\perp = \cosh\theta\, \vec{E}_\perp + \sinh\theta\, \hat{x}\times c\vec{B}$$
$$= \cosh\theta\,\left(\vec{E}_\perp + \frac{v}{c}\,\hat{x}\times c\vec{B}\right)$$

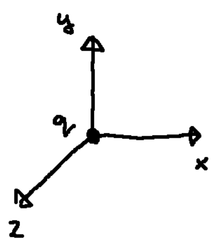
$$\tilde{\vec{E}}_\perp = \cosh\theta\,(\vec{E}_\perp + \vec{v}\times\vec{B})$$
$$\tilde{\vec{B}}_\perp = \cosh\theta\,\left(\vec{B}_\perp - \frac{\vec{v}}{c}\times\frac{\vec{E}}{c}\right)$$
$$\tilde{\vec{E}}_\parallel = \vec{E}_\parallel$$
$$\tilde{\vec{B}}_\parallel = \vec{B}_\parallel$$

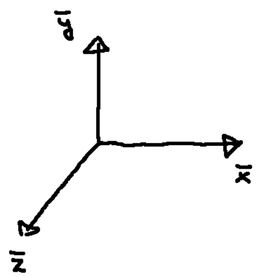
*

F.8 Video pages for fields due to a uniformly moving point charge

1

Fields due to a uniformly moving point charge





Frame 1

t, x, y, z

Frame 2

$\bar{t}, \bar{x}, \bar{y}, \bar{z}$

$$V_{21} = \hat{x}\, V_{21,x} = c\vec{\beta} \quad , \quad \beta = \frac{V_{21,x}}{c}$$

2

In frame 1 (charge stationary)

$$\vec{B} = 0$$

$$\vec{E} = \frac{q}{4\pi\epsilon_0} \frac{\hat{r}}{r^2}$$

$$= \frac{q}{4\pi\epsilon_0} \frac{\overbrace{(x\hat{x} + y\hat{y} + z\hat{z})}^{\vec{r}}}{r^3}$$

$$r^3 = (x^2 + y^2 + z^2)^{3/2}$$

Apply field transformation rules to obtain $\bar{\vec{E}}, \bar{\vec{B}}$ (the fields observed in frame 2)

3

$$\bar{\vec{E}}_\parallel = \vec{E}_\parallel$$

$$= \frac{q}{4\pi\epsilon_0}\,\frac{x}{r^3}\,\hat{x} \qquad (*)$$

Use coordinates in frame 2.

$$x = \gamma\,(\bar{x} + \beta c \bar{t})$$
$$ct = \gamma\,(c\bar{t} + \beta \bar{x})$$
$$y = \bar{y}$$
$$z = \bar{z}$$

$$\bar{\vec{E}}_\parallel = \frac{q}{4\pi\epsilon_0}\,\frac{\gamma\,(\bar{x} + \beta c \bar{t})\,\hat{x}}{\left[(\gamma(\bar{x}+\beta c\bar{t}))^2 + \bar{x}^2 + \bar{z}^2\right]^{3/2}}$$

Now consider transverse components.

$$\bar{\vec{E}}_\perp = \gamma\,(\vec{E}_\perp + \underbrace{\cancel{\vec{\beta}\times\vec{B}c}}_{=0})$$

$$= \frac{q}{4\pi\epsilon_0}\,\frac{\gamma\,(y\hat{y} + z\hat{z})}{r^3}$$

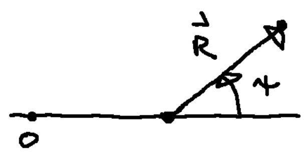
4

$$\bar{\vec{E}}_\perp = \frac{q}{4\pi\epsilon_0}\,\gamma\,\frac{(\bar{y}\hat{y} + \bar{z}\hat{z})}{\cdots}$$

$$\bar{\vec{E}} = \bar{\vec{E}}_\parallel + \bar{\vec{E}}_\perp$$

$$= \frac{q}{4\pi\epsilon_0}\,\gamma\,\frac{\left[(\bar{x} + \widetilde{\beta c\bar{t}})\,\hat{x} + \bar{y}\hat{y} + \bar{z}\hat{z}\right]}{\left[\,(\gamma(\bar{x}+\beta c\bar{t}))^2 + \bar{y}^2 + \bar{z}^2\right]^{3/2}}$$

Frame 2



$$\vec{R} := (\bar{x} + \beta c\bar{t})\,\hat{x} + \bar{y}\hat{y} + \bar{z}\hat{z}$$

$$\hat{x}\cdot\vec{R} = R\cos\psi$$

$$\frac{y^2+z^2}{R^2} = \sin^2\psi$$

(5)

$$\vec{E} = \frac{q}{4\pi\epsilon_0}\,\gamma\,\frac{\vec{R}}{\left[(\gamma R\cos\psi)^2 + R^2\sin^2\psi\right]^{3/2}}$$

$$= \frac{q}{4\pi\epsilon_0}\,\gamma\,\frac{\hat{R}}{R^2}\,\frac{1}{\left[(\gamma\cos\psi)^2 + \sin^2\psi\right]^{3/2}}$$

$$\gamma^2\cos^2\psi + \sin^2\psi = \gamma^2(1-\sin^2\psi) + \sin^2\psi$$

$$= \gamma^2 + (1-\gamma^2)\sin^2\psi$$

$\gamma = \cosh\theta$

$$= \cosh^2\theta\left(1 + \frac{-\sinh^2\theta}{\cosh^2\theta}\sin^2\psi\right)$$

$\cosh^2\theta - \sinh^2\theta = 1$

$$= \gamma^2(1-\beta^2\sin^2\psi)$$

$$\vec{E} = \frac{q}{4\pi\epsilon_0}\,\frac{\hat{R}}{R^2}\,\frac{\gamma}{\gamma^3}\,\frac{1}{\left[1-\beta^2\sin^2\psi\right]^{3/2}} \qquad \gamma^2 = \frac{1}{1-\beta^2}$$

(6)

$$\boxed{\vec{E} = \frac{q}{4\pi\epsilon_0}\,\frac{\hat{R}}{R^2}\,\frac{(1-\beta^2)}{(1-\beta^2\sin^2\psi)^{3/2}}}$$

If $\psi = \pi/2$ (90°), $\vec{E}$ field is γ factor higher than in electrostatics.

Consider $\vec{B}$ field.

$$\bar{\vec{B}}_{\parallel} = \vec{B}_{\parallel}, \quad \bar{\vec{B}}_{\perp} = \gamma\left[\vec{B}_{\perp} - \vec{\beta}\times\frac{\vec{E}}{c}\right]$$

Consider inverse transform: $\underbrace{\vec{B}_{\perp}}_{=\vec{0}} = \gamma\left[\bar{\vec{B}}_{\perp} + \vec{\beta}\times\frac{\bar{\vec{E}}}{c}\right]$

7

$$\vec{B}_\perp = -\vec{\beta} \times \frac{\vec{E}_\perp}{c}$$ for already determined.

$$(\vec{v} \times \vec{E}) \cdot \vec{v} = 0$$

$$\vec{B} = -\vec{\beta} \times \frac{\vec{E}}{c}$$

frame motion, 2 wrt 1

$$\boxed{\vec{B} = \frac{\vec{v}}{c} \times \frac{\vec{E}}{c}}$$

velocity of particle in frame 2

F.9 Video pages for Maxwell's equations in a manifestly covariant form

1

Maxwell's equations in a manifestly covariant form

$(\rho c, \vec{J})$ is a four vector

rest frame of a moving charge density ρ_0, $\vec{J}_0 = 0$

in any other inertial reference frame $\rho = \dfrac{\rho_0}{\sqrt{1-(v/c)^2}}$

$$\vec{J} = \frac{\vec{v}\,\rho_0}{\sqrt{1-(v/c)^2}}$$

$$(\rho c, \vec{J}) = \left(\rho_0 c \frac{1}{\sqrt{1-(v/c)^2}}, \frac{\vec{v}\,\rho_0}{\sqrt{1-(v/c)^2}} \right)$$

$$= \rho_0 \frac{1}{\sqrt{1-(v/c)^2}} (c, \vec{v})$$ ← four-velocity

②

$$\left\{\begin{array}{l} \nabla\cdot\vec{E} = \rho/\epsilon_0, \quad \cancel{\nabla\times\vec{E} = -\frac{\partial\vec{B}}{\partial t}}, \quad \cancel{\nabla\cdot\vec{B} = 0} \\ \nabla\times\vec{B} = \mu_0\vec{J} + \mu_0\epsilon_0\frac{\partial\vec{E}}{\partial t} \end{array}\right\}$$

$$\boxed{\vec{B} = \nabla\times\vec{A}} \qquad \boxed{\vec{E} = -\nabla\varphi - \frac{\partial\vec{A}}{\partial t}} \quad *$$

$$\Downarrow$$

$$\nabla\cdot\vec{B} = 0$$

$$\nabla\times\vec{E} = -\frac{\partial}{\partial t}\nabla\times\vec{A}$$

$$\nabla\times\vec{E} = -\nabla\times\left(\frac{\partial\vec{A}}{\partial t}\right)$$

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

$$\vec{E} + \frac{\partial\vec{A}}{\partial t} = -\nabla\varphi$$

$$\vec{E} = -\nabla\varphi - \frac{\partial\vec{A}}{\partial t}$$

③

$$\nabla\cdot\vec{E} = \rho/\epsilon_0 \qquad \nabla\cdot\vec{A} = -\frac{1}{c^2}\frac{\partial\varphi}{\partial t}$$

$$\nabla\cdot\left(-\nabla\varphi - \frac{\partial\vec{A}}{\partial t}\right) = \rho/\epsilon_0$$

$$\boxed{-\nabla^2\varphi - \frac{\partial}{\partial t}(\nabla\cdot\vec{A}) = \rho/\epsilon_0}$$

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

$$\nabla\times(\nabla\times\vec{A}) = \mu_0\vec{J} + \frac{1}{c^2}\frac{\partial}{\partial t}\left(-\nabla\varphi - \frac{\partial\vec{A}}{\partial t}\right)$$

$$\nabla(\nabla\cdot\vec{A}) - \nabla^2\vec{A} = \mu_0\vec{J} - \frac{1}{c^2}\nabla\frac{\partial\varphi}{\partial t} - \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2}$$

$$\boxed{\nabla(\nabla\cdot\vec{A}) - \nabla^2\vec{A} + \frac{1}{c^2}\nabla\frac{\partial\varphi}{\partial t} + \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = \mu_0\vec{J}} \quad *$$

$$\nabla\cdot\vec{A} = 0$$

$$-\nabla^2\vec{A} = \mu_0\vec{J}$$

$$\nabla^2\varphi = -\rho/\epsilon_0$$

(4)

Coulomb gauge : $\nabla\cdot\vec{A} = 0$ ←

(very useful in magnetostatics)

$$\vec{A} \to \vec{A} + \nabla g$$ ← scalar function

$$\vec{E} = -\nabla\varphi - \frac{\partial \vec{A}}{\partial t} \to \vec{E} = -\nabla\left(\varphi - \frac{\partial}{\partial t} g\right) - \frac{\partial}{\partial t}\left(\vec{A} + \nabla g\right)$$

$$= -\nabla\varphi - \frac{\partial \vec{A}}{\partial t}$$

$$\varphi \to \varphi - \frac{\partial g}{\partial t}$$

$$\nabla\cdot\vec{A} = f \qquad \nabla\cdot\left(\overbrace{\vec{A} + \nabla g}\right) = t$$

$$\boxed{\nabla^2 g = t - f}$$ ← we can solve this for the g that gives

(5)

Lorenz Gauge : $\underbrace{\nabla\cdot\vec{A} = -\frac{1}{c^2}\frac{\partial\varphi}{\partial t}}$

$$* \quad -\nabla^2\varphi + \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = \rho/\epsilon_0$$

$$* \quad -\nabla^2\vec{A} + \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = \mu_0\vec{J}$$

$$\to -\nabla^2\varphi/c + \frac{1}{c^2}\frac{\partial^2(\varphi/c)}{\partial t^2} = \frac{\rho}{c\epsilon_0}$$

$$c^2 = \frac{1}{\mu_0\epsilon_0} \qquad \frac{1}{\epsilon_0} = \mu_0 c^2 \quad \Big) = \mu_0 \rho c$$

$$A^\alpha \equiv (\varphi/c, \vec{A}) \qquad J^\beta \equiv (\rho c, \vec{J})$$

$$\boxed{-\partial_\alpha\partial^\alpha A^\beta = \mu_0 J^\beta}$$!!! Maxwell's equations !!!

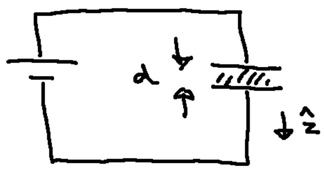
$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right) \begin{matrix} \varphi/c \\ A_x \\ A_y \\ A_z \end{matrix} = \mu_0 \begin{matrix} \rho c \\ J_x \\ J_y \\ J_z \end{matrix}$$

F.10 Video pages for Poynting's theorem

1

Poynting's Theorem

$$\frac{d}{dt}\left(\frac{mc^2}{\sqrt{1-(v/c)^2}}\right) = \underline{q\vec{v}\cdot\vec{E}}$$



$Q = VC$

$$dW = V' dQ$$

$$W = \int_0^V V' C \, dV'$$

$$= \frac{1}{2} C V^2 \qquad (*)$$

$$C = \frac{\epsilon_0 A}{d}$$

$$\vec{E} = \frac{Q}{A\epsilon_0}\hat{z} = \frac{VC}{A\epsilon_0}\hat{z}$$

$$V = \frac{E_z A \epsilon_0}{C}$$

$$W = \frac{1}{2} C \frac{E_z^2 A^2 \epsilon_0^2}{C^2} = \frac{1}{2}\frac{E_z^2 A^2 \epsilon_0^2}{\epsilon_0 A/d}$$

2

$$W = \frac{1}{2}\epsilon_0 E_z^2 \, Ad$$

Suggests that

$$\rho_E \stackrel{?}{=} \frac{1}{2}\epsilon_0 E^2$$

energy density in E field

Similar argument using infinite solenoid

$$\rho_B \stackrel{?}{=} \frac{B^2}{2\mu_0}$$

3

$$P = q\vec{v}\cdot\vec{E}$$

going into the particle's energy

convert to using $\vec{J} = \rho\vec{v}$

volume

$$P = \int d\tau\, \vec{J}\cdot\vec{E}$$

rewrite using Maxwell's eqns.

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

$$P = \int d\tau \left[\frac{\nabla\times\vec{B}}{\mu_0} - \epsilon_0\frac{\partial\vec{E}}{\partial t}\right]\cdot\vec{E}$$

$$= \int d\tau \left[\frac{1}{\mu_0}\underbrace{(\nabla\times\vec{B})\cdot\vec{E}} - \epsilon_0\frac{\partial\vec{E}}{\partial t}\cdot\vec{E}\right]$$

4

$$\nabla\cdot(\vec{v}\times\vec{w}) = \vec{w}\cdot(\nabla\times\vec{v}) - \vec{v}\cdot(\nabla\times\vec{w})$$

$$-\nabla\cdot(\vec{E}\times\vec{B}) = \vec{E}\cdot(\nabla\times\vec{B}) - \vec{B}\cdot(\nabla\times\vec{E})$$

$$P = \int d\tau \left[\frac{1}{\mu_0}\left\{\nabla\cdot(\vec{E}\times\vec{B}) + \vec{B}\cdot\underbrace{(\nabla\times\vec{E})}_{-\frac{\partial\vec{B}}{\partial t}}\right\} - \epsilon_0\frac{\partial\vec{E}}{\partial t}\cdot\vec{E}\right]$$

$$P = \int d\tau\left[\frac{1}{\mu_0}\nabla\cdot(\vec{E}\times\vec{B}) - \frac{1}{2\mu_0}\frac{\partial}{\partial t}B^2 - \frac{1}{2}\epsilon_0\frac{\partial}{\partial t}E^2\right]$$

$$P = \underbrace{\int d\tau\,\frac{1}{\mu_0}\nabla\cdot(\vec{E}\times\vec{B})} - \frac{d}{dt}\int d\tau\left(\frac{B^2}{2\mu_0} + \frac{\epsilon_0}{2}E^2\right)$$

5

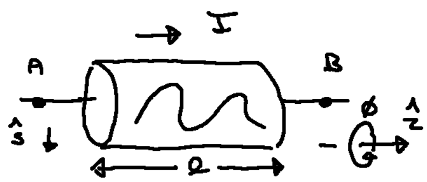
$$P = -\oint \underbrace{\frac{1}{\mu_0}(\vec{E}\times\vec{B})}_{=:\ \vec{S},\ \text{Poynting vector}}\cdot d\vec{a} \;-\; \frac{d}{dt}\int d\tau \underbrace{\left[\frac{B^2}{2\mu_0}+\frac{\epsilon_0}{2}E^2\right]}_{=:\ u_{em}}$$

$=: u_{em}$ ← electromagnetic energy density

$$0 = \underbrace{P}_{\text{power going into particles.}} + \underbrace{\oint \vec{S}\cdot d\vec{a}}_{\text{power that's leaving region}} + \underbrace{\frac{d}{dt}\int d\tau\, u_{em}}_{\text{power change of stored field energy}} \quad \Big\}\ \text{energy conservation.}$$

6

Example of Poynting vector usage



$V_{AB} = R I$

$P = I^2 R$ ← *

$$\frac{du_{em}}{dt} = 0 \qquad \text{“steady state”}$$

$$P = -\oint \frac{1}{\mu_0}(\vec{E}\times\vec{B})\cdot d\vec{a}$$

take surface as surface of resistor

What are $\vec{E}$ and $\vec{B}$:

Ampere's law: $2\pi s B_\phi = \mu_0 I$

$\vec{B} = \hat{\phi}\,\frac{\mu_0 I}{2\pi s}$ on surface

Gauss' law: $\vec{E} = V/\ell\ \hat{z}$

(7)

$$\frac{1}{\mu_0}\vec{E}\times\vec{B} = \frac{V}{\ell}\,\hat{z}\times\hat{\phi}\,\frac{\mu_0 I}{2\pi s}$$

$$= \frac{V}{\ell}\,(-\hat{s})\,\frac{\mu_0 I}{2\pi s}$$

End caps make no contribution since $\hat{s}\cdot\hat{n} = 0$ on end caps

$$-\oint \frac{1}{\mu_0}\vec{E}\times\vec{B}\cdot d\vec{a} = \frac{V\cancel{\mu_0}I}{\cancel{\ell}\,\cancel{2\pi s}}\;\frac{\cancel{2\pi s}\,\cancel{\ell}}{\cancel{\mu_0}}$$

$$= VI$$

$$\underline{P = RI^2}$$

F.11 Video pages for momentum conservation and Maxwell's stress tensor

①

Momentum conservation and Maxwell's stress tensor

$$\vec{P} = \sum_i \vec{p}_i$$

$$\frac{d\vec{P}}{dt} = \sum_i \dot{\vec{p}}_i$$

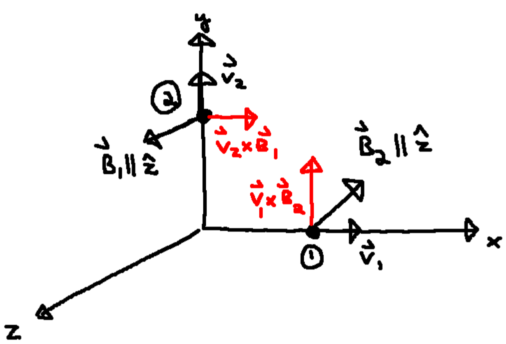
$$= \sum_i \vec{F}_i$$

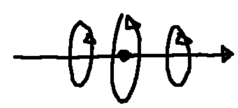
$$= \sum_{i \neq j} \vec{F}_{ij}$$ force on particle i due to particle j

If $\vec{F}_{ij} = -\vec{F}_{ji}$ (Newton's 3rd law)

$$\frac{d\vec{P}}{dt} = \vec{0}$$

②





The Lorentz force does not obey Newton's 3rd law!
Problem with momentum conservation (?).

Fields can store momentum!!!

③

$$\frac{d\vec{p}}{dt} = \underbrace{q}_{d\tau\,\rho}(\vec{E} + \vec{v}\times\vec{B})$$

generalize to: $$\frac{d\vec{p}}{dt} = \int d\tau\,[\rho\vec{E} + \vec{J}\times\vec{B}] \qquad (*)$$

Similar to Poynting's theorem derivation, write ρ and $\vec{J}$ in terms of fields.

$$\nabla\cdot\vec{E} = \rho/\epsilon_0 \;*\; \Rightarrow \quad \rho = \epsilon_0\nabla\cdot\vec{E}$$

$$\nabla\times\vec{B} = \mu_0\vec{J} + \mu_0\epsilon_0\frac{\partial\vec{E}}{\partial t} \;*\; \Rightarrow \quad \vec{J} = \frac{1}{\mu_0}\nabla\times\vec{B} - \epsilon_0\frac{\partial\vec{E}}{\partial t}$$

Substitute into (*):

$$\frac{dp}{dt} = \int d\tau\,\Big[\underset{\frac{1}{\mu_0}(\nabla\cdot\vec{B})\vec{B}}{\epsilon_0(\nabla\cdot\vec{E})\vec{E}} + \Big(\frac{1}{\mu_0}\nabla\times\vec{B} - \underbrace{\epsilon_0\frac{\partial\vec{E}}{\partial t}\Big)\times\vec{B}}\Big]$$

④

$$\frac{\partial}{\partial t}(\vec{E}\times\vec{B}) = \underbrace{\frac{\partial\vec{E}}{\partial t}\times\vec{B}} + \vec{E}\times\underset{\vec{E}\times(-\nabla\times\vec{E})}{\frac{\partial\vec{B}}{\partial t}}$$

$$\frac{d\vec{p}}{dt} = \int d\tau\,\Big[\epsilon_0(\nabla\cdot\vec{E})\vec{E} - \epsilon_0\vec{E}\times(\nabla\times\vec{E})$$
$$+ \frac{1}{\mu_0}(\nabla\times\vec{B})\times\vec{B} + \frac{1}{\mu_0}(\nabla\cdot\vec{B})\vec{B}\Big)$$
$$- \epsilon_0\frac{\partial}{\partial t}(\vec{E}\times\vec{B})\Big]$$

$$\vec{v}\times(\nabla\times\vec{v}) = \frac{1}{2}\nabla(v^2) - (\vec{v}\cdot\nabla)\vec{v} \qquad *$$

by

$$\underbrace{\nabla(\vec{A}\cdot\vec{B})}_{\text{generic}} = \vec{A}\times(\nabla\times\vec{B}) + \vec{B}\times(\nabla\times\vec{A}) + (\vec{A}\cdot\nabla)\vec{B} + (\vec{B}\cdot\nabla)\vec{A}$$

$$\nabla(v^2) = \underbrace{2\vec{v}\times(\nabla\times\vec{v})} + 2(\vec{v}\cdot\nabla)\vec{v}$$

(5)

$$(\nabla\cdot\vec{v})\vec{v} + (\vec{v}\cdot\nabla)\vec{v} - \frac{1}{2}\nabla v^2$$

Can we use the divergence theorem?
Try to look at each component individually.

$$\hat{x}\cdot\left[(\nabla\cdot\vec{v})\vec{v} + (\vec{v}\cdot\nabla)\vec{v} - \frac{1}{2}\nabla v^2\right]$$

$$= (\nabla\cdot\vec{v})v_x + (\vec{v}\cdot\nabla)v_x - \frac{1}{2}\hat{x}\cdot\nabla v^2$$

$$= (\nabla\cdot\vec{v})v_x + \vec{v}\cdot(\nabla v_x)$$

$$\nabla\cdot(f\vec{A}) = f\,\nabla\cdot\vec{A} + \nabla f\cdot\vec{A}$$ ← general identity

$$= \nabla\cdot(v_x\vec{v}) - \frac{1}{2}\nabla\cdot(\hat{x}\,v^2)$$

$$= \nabla\cdot\left(v_x\vec{v} - \frac{1}{2}v^2\hat{x}\right)$$

suitable for div theorem $\int d\tau\,\nabla\cdot\vec{A} = \oint \vec{A}\cdot d\vec{a}$

(6)

$$\hat{x}\cdot\frac{d\vec{p}}{dt} = \epsilon_0\oint\left(E_x\vec{E} - \frac{1}{2}E^2\hat{x}\right)\cdot d\vec{a} + \frac{1}{\mu_0}\oint\left(B_x\vec{B} - \frac{1}{2}B^2\hat{x}\right)\cdot d\vec{a} - \int d\tau\,\hat{x}\cdot\epsilon_0\frac{\partial}{\partial t}(\vec{E}\times\vec{B})$$

$$\frac{d\vec{p}}{dt} = \underbrace{\hat{x}_i\oint T_{ij}\,da_j}_{\overleftrightarrow{T}\cdot d\vec{a}} - \int d\tau\,\epsilon_0\frac{\partial}{\partial t}(\vec{E}\times\vec{B})$$

where

Maxwell's stress tensor $\quad T_{ij} := \epsilon_0\left(E_iE_j - \frac{1}{2}E^2\delta_{ij}\right) + \frac{1}{\mu_0}\left(B_iB_j - \frac{1}{2}B^2\delta_{ij}\right)$

$$0 = \underbrace{\frac{d\vec{p}}{dt}}_{\text{change in momentum of particles.}} - \underbrace{\oint\overleftrightarrow{T}\cdot d\vec{a}}_{\text{momentum flow out of region.}} + \underbrace{\int d\tau\,\epsilon_0\frac{\partial}{\partial t}(\vec{E}\times\vec{B})}_{\text{momentum going into fields}}$$

$\epsilon_0\vec{E}\times\vec{B}$ momentum density

F.12 Video pages for wave equations

1

Wave equations

Maxwell's equations with no sources:

$$\nabla\cdot\vec{E} = 0 \quad (1) \qquad \nabla\times\vec{E} = -\frac{\partial\vec{B}}{\partial t} \quad (2)$$

$$\nabla\cdot\vec{B} = 0 \quad (3) \qquad \nabla\times\vec{B} = \mu_0\epsilon_0\frac{\partial\vec{E}}{\partial t} \quad (4)$$

Take $\nabla\times$ of both sides of Eq. (2):

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

$$\nabla(\underbrace{\cancel{\nabla\cdot\vec{E}}}_{=0}) - \nabla^2\vec{E} = -\frac{\partial}{\partial t}\nabla\times\vec{B}$$

2

Use Eq. (4) on RHS:

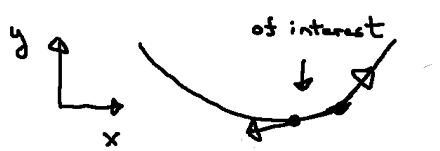
$$-\nabla^2\vec{E} = -\frac{\partial}{\partial t}\overbrace{\left(\mu_0\epsilon_0\frac{\partial\vec{E}}{\partial t}\right)}^{\nabla\times\vec{B}}$$

$$\boxed{\nabla^2\vec{E} = \mu_0\epsilon_0\frac{\partial^2\vec{E}}{\partial t^2}}$$

By similar argument: $\nabla^2\vec{B} = \mu_0\epsilon_0\dfrac{\partial^2\vec{B}}{\partial t^2}$

these are wave equations

A more familiar context: a string under tension.



two forces on segment don't cancel if there is curvature, so segment will accelerate.

③

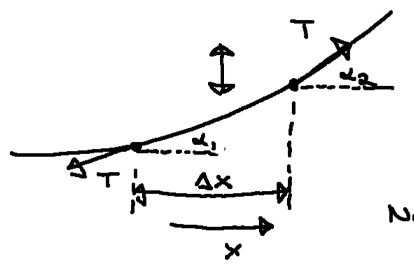
Some approximations:

① tension is contant T

② string remains close to parallel with horizontal axis *

③ segments just move in y-direction

④ no friction, no stiffness



vertical force on segment

$$F_y = T\sin\alpha_1 - T\sin\alpha_2$$

$$= T\left(\frac{\partial y_2}{\partial x} - \frac{\partial y_1}{\partial x}\right)$$

Newton's 2nd law

$$m\frac{\partial^2 y}{\partial t^2} = T\left(\frac{\partial y_2}{\partial x} - \frac{\partial y_1}{\partial x}\right)$$

④

mass of segment $= m = \lambda\,\Delta x$

mass per unit length

$$\boxed{\lambda\frac{\partial^2 y}{\partial t^2} = T\frac{\partial^2 y}{\partial x^2}}$$

also a wave equation

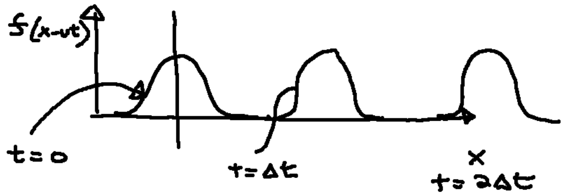
wave: "a wave is a disturbance of a continuous medium that propagates with a fixed shape at a constant velocity" GITE 4

Consider any twice differentiable function of a single variable: $f(\phi)$

$\phi = x - vt$

v constant

5



$$\frac{\partial y}{\partial t} = \frac{\partial f}{\partial \phi}\frac{\partial \phi}{\partial t} \qquad \text{chain rule}$$

$$= \frac{\partial f}{\partial \phi}(-v)$$

$$\left.\begin{aligned} \frac{\partial^2 y}{\partial t^2} &= \frac{\partial^2 f}{\partial \phi^2}(v^2) \\ \frac{\partial^2 y}{\partial x^2} &= \frac{\partial^2 f}{\partial \phi^2} \end{aligned}\right\}$$

$$\lambda v^2 \cancel{\frac{\partial^2 f}{\partial \phi^2}} = T \cancel{\frac{\partial^2 f}{\partial \phi^2}}$$

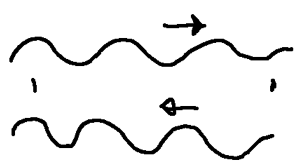
$$\boxed{v = \sqrt{\frac{T}{\lambda}}}$$

$$\sqrt{\frac{kg \cdot m/s^2}{kg/m}} = m/s \quad \checkmark$$

6

$$\boxed{\frac{1}{v^2}\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2}} \qquad *$$

To satisfy BC's of a string fixed at two points, we can use a superposition of two "travelling waves" moving in opposite directions



\+

allowed by linearity of wave equation

not all solutions to the wave equation are waves.

=

standing waves

F.13 Video pages for electromagnetic plane waves

1

Electromagnetic plane waves

$$\nabla^2 \vec{E} = \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} \qquad (*)$$

$$\vec{E} = \underset{\text{constant}}{\vec{E}_0} \cos(kz - \omega t) \quad \text{satisfies } (*).$$

$$f(x \mp vt)$$

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$$

$$\vec{E} = \vec{E}_0 \cos\left(k \underbrace{\left[z - \frac{\omega}{k} t\right]}_{\phi}\right)$$

$$\vec{B} = \vec{B}_0 \cos\left(k \left[z - \frac{\omega}{k} t\right]\right)$$

2

Use complex numbers to simplify algebra:

$$\vec{E}(\vec{r}, t) = \vec{E}_0 e^{i(kz - \omega t)} \qquad (*)$$

$\vec{E}_0$ could be a complex number

Similarly for $\vec{B}(\vec{r}, t)$.

Meaning? We should interpret the real part of $\vec{E}(\vec{r}, t)$ as being the actual field.

Recall for any complex number z: $\mathrm{Re}(z) = \frac{z + z^*}{2}$

Suppose $(*)$ together with a $\vec{B}(\vec{r}, t)$ satisfies Maxwell's eqns, then $\vec{E}^*(\vec{r}, t)$, $\vec{B}^*(\vec{r}, t)$ also satisfy Maxwell's equations.

3

Double check that $\vec{E} = \vec{E}_0 e^{i(kz-\omega t)}$

satisfies wave equation $\nabla^2 \vec{E} = \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2}$

$$\vec{E}_0 (+ik)^2 \cancel{e^{i(kz-\omega t)}} = \mu_0 \epsilon_0 (-i\omega)^2 \cancel{e^{i(kz-\omega t)}} \vec{E}_0$$

$$\frac{\omega}{k} = \sqrt{\frac{1}{\mu_0 \epsilon_0}} = c \quad ✓$$

Apply $\nabla \cdot \vec{E} = 0$: $\nabla \cdot \left(\vec{E}_0 e^{i(kz-\omega t)} \right) = 0$ *

Remember vector identity:

$$\nabla \cdot (f \vec{v}) = \nabla f \cdot \vec{v} + f \, \nabla \cdot \vec{v}$$

$$ik\hat{z} e^{i(kz-\omega t)} \cdot \vec{E}_0 = 0$$

$$\boxed{\vec{E}_0 \cdot \hat{z} = 0} \quad \text{similarly} \quad \boxed{\vec{B}_0 \cdot \hat{z} = 0}$$

4

What is the relationship between $\vec{E}_0$ and $\vec{B}_0$?

$$* \quad \nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

$$\nabla \times \left(\vec{E}_0 \underbrace{e^{i(kz-\omega t)}} \right) = -\frac{\partial}{\partial t} \vec{B}_0 e^{i(kz-\omega t)}$$

Vector identity: $\nabla \times (f \vec{v}) = \nabla f \times \vec{v} + f \, \nabla \times \vec{v}$

$$ik\hat{z} e^{i(kz-\omega t)} \times \vec{E}_0 = \vec{B}_0 \, i\omega t$$

$$\boxed{\vec{B}_0 = \frac{1}{c} \hat{z} \times \vec{E}_0} \quad (1)$$

$$\nabla \times \vec{B} = \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} \Rightarrow \boxed{\vec{E}_0 = -\hat{z} \times c\vec{B}_0} \quad (2)$$

5

Show that (2) ⇒ (1) :

$$\hat{z} \times \vec{E}_0 = \hat{z} \times (-\hat{z} \times c\vec{B}_0)$$

$$= -[\hat{z}\,\underbrace{(\hat{z}\cdot c\vec{B}_0)}_{=0} - c\vec{B}_0\,(\hat{z}\cdot\hat{z})]$$

$$= c\vec{B}_0$$

$$\vec{B}_0 = \frac{1}{c}(\hat{z} \times \vec{E}_0)$$ //

Also (1) ⇒ (2) by similar argument.

6

$$\vec{E} = \vec{E}_0\, e^{i(kz-\omega t)}$$

$$\vec{B} = \frac{\hat{z} \times \vec{E}_0}{c}\, e^{i(kz-\omega t)}$$

with $\vec{E}_0 \cdot \hat{z} = 0$ and $\omega/k = \frac{1}{\sqrt{\mu_0\epsilon_0}} = c$

Generalize to arb. direction of propagation $\hat{k}$

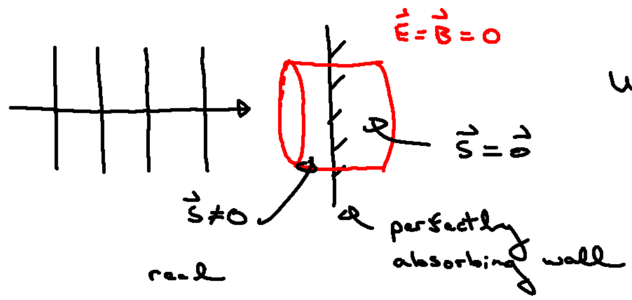
$$\boxed{\begin{aligned} \vec{E} &= \vec{E}_0\, e^{i(\vec{k}\cdot\vec{r}-\omega t)} \\ \vec{B} &= \frac{\hat{k} \times \vec{E}_0}{c}\, e^{i(\vec{k}\cdot\vec{r}-\omega t)} \\ \vec{E}_0 \cdot \hat{k} &= 0\,, \quad \frac{\omega}{|\vec{k}|} = \frac{1}{\sqrt{\mu_0\epsilon_0}} = c \end{aligned}}$$

plane wave solutions.

F.14 Video pages for energy and momentum transport in electromagnetic waves

1

Energy and momentum transport by electromagnetic waves



What is power / unit area dissipated in wall?

real

$$\vec{E} = \vec{E}_0 \cos(\vec{k}\cdot\vec{r} - \omega t)$$

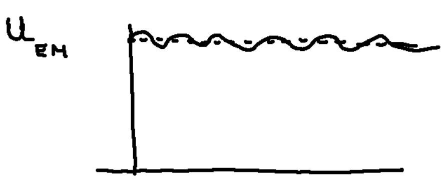
$$\vec{B} = \hat{k} \times \frac{\vec{E}_0}{c} \cos(\vec{k}\cdot\vec{r} - \omega t)$$

linearly polarized.

2

Apply Poynting's theorem

$$P = -\frac{d}{dt} \overbrace{\int d\tau \left[\frac{\epsilon_0 E^2}{2} + \frac{1}{\mu_0}\frac{B^2}{2} \right]}^{=:\, u_{EM}} - \frac{1}{\mu_0} \oint \underbrace{(\vec{E}\times\vec{B})}\cdot d\vec{a}$$



$\langle \dots \rangle$ =: time average.

$$\left\langle \frac{du_{EM}}{dt} \right\rangle = 0$$

③

We only need to consider the Poynting vector term.

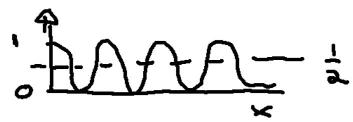
$$\vec{S} = \frac{1}{\mu_0} \vec{E} \times \vec{B}$$

$$= \frac{1}{\mu_0} \vec{E}_0 \cos(\vec{k}\cdot\vec{r} - \omega t) \times \hat{k} \times \frac{\vec{E}_0}{c} \cos(\vec{k}\cdot\vec{r} - \omega t)$$

$$\langle \vec{S} \rangle = \frac{1}{\mu_0 c} \vec{E}_0 \times \left(\hat{k} \times \frac{\vec{E}_0}{c}\right) \langle \cos^2(\vec{k}\cdot\vec{r} - \omega t) \rangle \quad \frac{1}{2}$$

$$= \frac{1}{\mu_0 c} \left(\hat{k} E_0^2 - \vec{E}_0 \underbrace{(\vec{E}_0 \cdot \hat{k})}_{=0} \right) \frac{1}{2}$$

$\nabla \cdot \vec{E} = 0$



$$\cos^2 x = \frac{1 + \cos 2x}{2}$$

$$\boxed{\langle \vec{S} \rangle = \frac{1}{2\mu_0 c} E_0^2 \hat{k}}$$

④

$$P = A \frac{E_0^2}{2\mu_0 c} \qquad \boxed{\frac{P}{A} = \frac{E_0^2}{2\mu_0 c}}$$

intensity

What units does $\mu_0 c$ have?

$$\frac{N}{A^2} \frac{m}{s} = \frac{N}{A \frac{C}{\cancel{s}}} \frac{m}{\cancel{s}} = \frac{V}{A} = \Omega$$

$$\mu_0 c = \underline{377\,\Omega}$$

impedance of free space

$$P = \frac{V^2}{R} \quad ✓$$

5

How intense would have to be to "pull" an electron off a hydrogen atom?

$$\underset{\text{at Bohr radius}}{E} \simeq \frac{q}{4\pi\epsilon_0}\frac{1}{a_0^2}$$

$$\simeq \frac{1.6*10^{-19}\,C}{4\pi * 8.85*10^{-12}\frac{1}{\frac{N}{C^2}\cdot m^2}*(0.5*10^{-10}\,m)^2}$$

$$\simeq 6*10^{11}\,\frac{N}{C} \qquad \left(\frac{V}{m}\right)$$

$$I = \frac{1}{2*377\Omega}\,|E|^2$$

$$\simeq 4.4*10^{20}\,\frac{W}{m^2}$$

6

Consider a pulsed laser beam $\Delta t \simeq 10\,fs$ duration

$A \simeq (10^{-6}\,m)^2$ cross-section

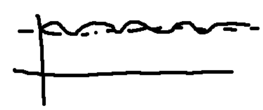
$$E = I\,\Delta t\,A$$

energy per pulse

$$= 4.4*10^{20}\,\frac{W}{m^2}*(10^{-6}\,m)^2 * 10^{-14}\,s$$

$$= 4.4*10^{-6}\,J \simeq 4.4\,\mu J$$

$$\frac{d\vec{p}}{dt} = -\oint \overleftrightarrow{T}\cdot d\vec{a} + \underbrace{\frac{d}{dt}\underbrace{\int d\tau\, \epsilon_0(\vec{E}\times\vec{B})}}_{=0 \text{ when time averaged.}}$$



7

What is $\overleftrightarrow{T}$?

$$T_{ij} = \epsilon_0 \left(E_i E_j - \frac{1}{2} E^2 \delta_{ij}\right) + \frac{1}{\mu_0}\left(B_i B_j - \frac{1}{2} B^2 \delta_{ij}\right)$$

$$\vec{E} = \hat{x}\, E_0 \cos(kz - \omega t)$$

$$\vec{B} = \underbrace{\hat{z} \times \hat{x}}_{\hat{y}}\ \frac{E_0}{2} \cos(kz - \omega t)$$

again we use $\langle \cos^2 x \rangle = \frac{1}{2}$

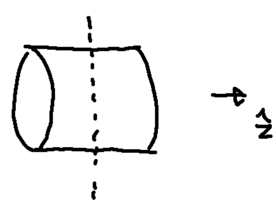
$$\langle T \rangle = \frac{1}{2} \left\{ \underbrace{\frac{\epsilon_0 E_0^2}{2}} \begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix} + \underbrace{\frac{1}{2\mu_0}\left(\frac{E_0}{c}\right)^2} \begin{bmatrix} -1 & & \\ & 1 & \\ & & -1 \end{bmatrix} \right\}$$

$$\frac{1}{\mu_0 \epsilon_0} = c^2 \Rightarrow \frac{1}{\mu_0 c^2} = \epsilon_0$$

8

$$\langle T \rangle = \frac{1}{2} \epsilon_0 E_0^2 \begin{bmatrix} 0 & & \\ & 0 & \\ & & -1 \end{bmatrix}$$

$$\underbrace{-\oint \overleftrightarrow{T} \cdot d\vec{a}}$$



$$= A\, \frac{1}{2} \epsilon_0 E_0^2\, \hat{z} \quad , \quad \epsilon_0 = \frac{1}{\mu_0 c^2}$$

$$\boxed{\frac{1}{A} \left\langle \frac{d\vec{p}}{dt} \right\rangle = \hat{z}\, \frac{E_0^2}{2\mu_0 c^2}} = \hat{z} \left\langle \frac{dE}{dt} \right\rangle \frac{1}{c}$$

$$E = \sqrt{(pc)^2 + (\underbrace{mc^2}_{=0})^2} \qquad \underline{E = pc}$$

F.15 Video pages for solving Maxwell's equations in the Lorenz gauge

1

Solving Maxwell's equations in the Lorenz gauge

$$\partial_\alpha A^\alpha = 0$$

$$-\partial_\alpha \partial^\alpha A^\beta = \mu_0 J^\beta \quad *$$

$$\left(-\nabla^2 + \frac{1}{c}\frac{\partial^2}{\partial t^2}\right) \begin{bmatrix} \varphi/c \\ A_x \\ A_y \\ A_z \end{bmatrix} = \mu_0 \begin{bmatrix} \rho c \\ J_x \\ J_y \\ J_z \end{bmatrix} \quad *$$

$$\left(-\nabla^2 + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) \psi(\vec{r}, t) = f(\vec{r}, t) \quad *$$

2

Take a Green's function approach:

$$\left(-\nabla^2 + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) G(\vec{r}, t) = \underbrace{\delta^3(\vec{r})}_{\delta(x)\delta(y)\delta(z)} \delta(t)$$

$$\psi(\vec{r}, t) = \int d^3r' \, dt' \; G(\vec{r}-\vec{r}', t-t') \, f(\vec{r}', t') \quad *$$

Use Fourier transform technique:

$$\tilde{G}(\vec{r}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dt \; G(\vec{r}, t) \, e^{i\omega t}$$

$$G(\vec{r}, t) = \int_{-\infty}^{\infty} d\omega \; \tilde{G}(\vec{r}, \omega) \, e^{-i\omega t} \quad *$$

③

$$\tilde{\delta}(\omega) = \frac{1}{2\pi}$$

because $\displaystyle\int_{-\infty}^{\infty} d\omega \; \frac{1}{2\pi} \; e^{-i\omega t} = \frac{1}{\cancel{2\pi}} \cancel{2\pi}\, \delta(t)$

$$\left[-\nabla^2 + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right] \int_{-\infty}^{\infty} d\omega \; \tilde{G}(\vec{r},\omega)\, e^{-i\omega t} = \delta^3(\vec{r})\, \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \; e^{-i\omega t}$$

$$\int_{-\infty}^{\infty} d\omega \; \left[-\nabla^2 + \frac{(-i\omega)^2}{c^2}\right] \tilde{G}(\vec{r},\omega)\, e^{-i\omega t} = \quad ''$$

$$\int_{-\infty}^{\infty} d\omega \; \underbrace{\left\{ \left[-\nabla^2 - \frac{\omega^2}{c^2}\right] \tilde{G}(\vec{r},\omega) - \frac{\delta^3(\vec{r})}{2\pi} \right\}}_{=0} e^{-i\omega t} = 0$$

④

$$(+\nabla^2 + k^2)\, \tilde{G}(\vec{r},\omega) = -\frac{\delta^3(\vec{r})}{2\pi} \qquad *$$

$$k := \frac{\omega}{c}$$

$\tilde{G}(\vec{r},\omega)$ can be written $\tilde{G}(r,\omega)$, and use spherical coordinates.

$$\frac{1}{r}\frac{d^2}{dr^2}\left(r\,\tilde{G}(r,\omega)\right) + k^2\, \tilde{G}(r,\omega) = -\frac{\delta^3(\vec{r})}{2\pi}$$

Except at origin:

$$\frac{d^2 \left(r\,\tilde{G}(r,\omega)\right)}{dr^2} + k^2 \underbrace{r\,\tilde{G}(r,\omega)}_{=:\, x(r)} = 0$$

$$\frac{d^2 x(r)}{dr^2} = -k^2 x(r) \;\Rightarrow\; x(r) = \underset{\text{arb. constant}}{\cdot}\, e^{\pm ikr}$$

(5)

$$\tilde{G}(r,\omega) = A\,\frac{e^{ikr}}{r}\cancel{\alpha} + B\,\frac{e^{-ikr}}{r}\cancel{\alpha}$$

where A and B are given by B.C.'s.

Suppose $k=0$ then $\underbrace{-\nabla^2 \tilde{G}(r,0) = \delta^3(\vec{r})/2\pi}_{\text{Poisson's eqn for a point charge at origin.}}$

$$-\nabla^2 \varphi = \frac{q}{\epsilon_0}\,\delta^3(\vec{r})$$

$$\varphi = \frac{1}{4\pi\epsilon_0}\,\frac{q}{r}$$

$A + B = \frac{1}{2\pi}\frac{1}{4\pi}$. Assume $A = \frac{1}{2\pi}\frac{1}{4\pi}$

$B = 0$

(6)

$$\tilde{G}(r,\omega) = \frac{1}{4\pi}\,\frac{1}{2\pi}\,\frac{e^{ikr}}{r}$$

Inverse Fourier transform:

$$G(r,t) = \int_{-\infty}^{\infty} d\omega\, \frac{1}{4\pi}\,\frac{1}{2\pi}\,\frac{e^{i\frac{\omega}{c}r}}{r}\, e^{-i\omega t}$$

$$= \frac{1}{4\pi}\,\frac{1}{2\pi}\,\frac{1}{r}\int_{-\infty}^{\infty} d\omega\; e^{i\omega\left[\frac{r}{c}-t\right]}$$

$$= \frac{1}{4\pi}\,\frac{1}{\cancel{2\pi}}\,\frac{1}{r}\,\cancel{2\pi}\,\delta\!\left(\frac{r}{c}-t\right)$$

Now this Green's function for an arb. source field $f(\vec{r},t)$.

7

$$G(\vec r - \vec r', t - t') = \frac{1}{4\pi|\vec r - \vec r'|}\,\delta\left(\frac{|\vec r - \vec r'|}{c} - (t - t')\right)$$

$$\psi(\vec r, t) = \int \underbrace{d^3\vec r'\,dt'}_{4}\,\frac{1}{4\pi|\vec r - \vec r'|}\,\delta\left(\frac{|\vec r - \vec r'|}{c} - (t - t')\right) f(\vec r', t')$$

$$\int_{-\infty}^{\infty} dx'\,\delta(x - x') f(x') = f(x)$$

$$\psi(\vec r, t) = \frac{1}{4\pi}\int d^3\vec r'\,\frac{1}{|\vec r - \vec r'|}\,f\left(\vec r', t - \frac{|\vec r - \vec r'|}{c}\right)$$

$$= \frac{1}{4\pi}\int d\tau'\,\frac{1}{\mathscr{r}}\,f\Big(\vec r', \underbrace{t - \frac{\mathscr{r}}{c}}_{\text{retarded time}}\Big)$$

GITEY notation

$$\ast\ \vec A(\vec r, t) = \frac{\mu_0}{4\pi}\int d\tau'\,\frac{1}{\mathscr{r}}\,\vec J\left(\vec r', t - \frac{\mathscr{r}}{c}\right)$$

$V \qquad \rho$

and similarly $\varphi(\vec r, t)$ can be written in terms of ρ

G Acknowledgements

The Learn course page emoji 🧲 is from: https://openmoji.org/library/#emoji=1F9F2, licensed CC BY-SA 4.0. Thank you to Professor Grindlay who taught me graduate electricity and magnetism many years ago. I am grateful for errata and suggestions from: Finn Dodgson.

References

[1] D. J. Griffiths, *Introduction to electrodynamics*, Fourth edition (Cambridge University Press, Cambridge, United Kingdom ; New York, NY, 2018).

[2] L. D. Landau and E. M. Lifshitz, *The classical theory of fields*, 4. rev. Engl. ed., repr (Elsevier Butterworth Heinemann, Amsterdam Heidelberg, 2010).

[3] M. L. Boas, *Mathematical methods in the physical sciences*, 3rd ed (Wiley, Hoboken, NJ, 2006).

[4] N. D. Mermin, "Relativistic addition of velocities directly from the constancy of the velocity of light", Am. J. Phys. **51**, 1130–1131 (1983).

[5] T. Lahaye, P. Labastie, and R. Mathevet, "Fizeau's "aether-drag" experiment in the undergraduate laboratory", Am. J. Phys. **80**, 497–505 (2012).

[6] A. Pais, *Subtle is the Lord: the science and the life of Albert Einstein* (Oxford University Press, Oxford ; New York, 2005).

[7] D. H. Frisch and J. H. Smith, “Measurement of the Relativistic Time Dilation Using mu-Mesons”, Am. J. Phys. **31**, 342–355 (1963).

[8] A. Zangwill, *Modern electrodynamics* (Cambridge University Press, Cambridge, 2013).

[9] J. D. Jackson, *Classical electrodynamics*, 3rd ed (Wiley, New York, 1999).

[10] E. M. Purcell, *Electricity and magnetism*, Third edition (Cambridge University Press, Cambridge, 2013).

[11] J. R. Taylor, *Classical mechanics* (University Science Books, Sausalito, Calif, 2005).

[12] K. S. Thorne and R. D. Blandford, *Modern classical physics: optics, fluids, plasmas, elasticity, relativity, and statistical physics* (Princeton University Press, Princeton, 2017).

[13] A. M. Steane, *Relativity made relatively easy* (Oxford University Press, Oxford, United Kingdom, 2012).

[14] D. C. Kay, *Schaum’s outline of theory and problems of tensor calculus*, Schaum’s Outline Series. Schaum’s Outline Series in Mathematics (McGraw-Hill, New York, 1988).

[15] R. N. Zare, *Angular momentum: understanding spatial aspects in chemistry and physics* (Wiley, New York, 1988).

[16] J. Bailey et al., “Measurements of relativistic time dilatation for positive and negative muons in a circular orbit”, Nature **268**, 301–305 (1977).

[17] L. D. Landau and E. M. Lifshitz, *Mechanics*, 3. ed. (Elsevier, Butterworth Heinemann, Amsterdam, 2007).

[18] E. F. Taylor and J. A. Wheeler, *Spacetime physics: introduction to special relativity*, 2nd ed (W.H. Freeman, New York, 1992).

[19] R. Shankar, *Principles of quantum mechanics*, 2nd ed (Plenum Press, New York, 1994).

[20] M. S. Livingston, “Part i, History of the cyclotron”, Phys. Today **12**, 18–23 (1959).

[21] L. W. Alvarez, *Alvarez: adventures of a physicist* (Basic Books, New York, 1987).

[22] M. A. Heald and J. B. Marion, *Classical electromagnetic radiation*, Third edition (Dover Publications, Inc, Mineola, New York, 2012).

[23] L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskij, *Electrodynamics of continuous media*, 2. ed., rev. and enlarged, repr, Course of Theoretical Physics / L. D. Landau and E. M. Lifshitz 8 (Elsevier Butterworth-Heinemann, Amsterdam Heidelberg, 2009).

[24] W. S. C. Williams, *Introducing special relativity* (Taylor & Francis, London, 2002).

[25] F. W. J. Olver and National Institute of Standards and Technology (U.S.), eds., *NIST handbook of mathematical functions* (Cambridge University Press : NIST, Cambridge ; New York, 2010).

[26] M. Abramowitz and I. A. Stegun, eds., *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, Dover Books on Mathematics (Dover Publ, New York, NY, 2013).

[27] E. Tiesinga, P. J. Mohr, D. B. Newell, and B. N. Taylor, "CODATA recommended values of the fundamental physical constants: 2018", Rev. Mod. Phys. **93**, 025010 (2021).

[28] Particle Data Group et al., "Review of Particle Physics", Progress of Theoretical and Experimental Physics **2020** (2020) 10.1093/ptep/ptaa104.