UCR-TR-2024-FLIP-P165

UC Riverside Physics 165, Spring 2024

Introduction to Particle Physics

Kinematics and Dynamics of the Standard Model

Flip Tanedo
flip.tanedo@ucr.edu,

Department of Physics & Astronomy, University of California, Riverside, CA 92521, USA

An introduction to elementary particle physics: the study of the fundamental constituents of matter and the forces that dictate their interactions. We focus on building a theoretical understanding of the Standard Model of particle physics based on Feynman diagrams.

Last Compiled: April 9, 2024
Image: P5 Logo, Sandbox Studio

Contents

Chapter 1

The Course

1.1 Our Goal

The goal of this course is to teach the *theoretical framework* of particle physics. The underlying structure of this discipline is called **quantum field theory** and is the union of special relativity and quantum mechanics. Over one-quarter course we tell the story of leptons, quarks, and various gauge bosons—but what I *really* want to convey to you is how quantum field theory works.

Ordinarily, quantum field theory is a graduate-level course that you take after taking not only upper division quantum mechanics, but graduate-level quantum mechanics, electrodynamics, statistical mechanics/field theory, with smatterings of courses that hammer home special relativity (perhaps in general relativity course) and bits of complex analysis. Even then, the course has a reputation for being challenging because it demands a level of physical sophistication to appreciate.[a]

[a] Usually we tell our graduate students to expect to take quantum field theory a few times in order to prepare to spend the rest of their research careers continuing to chip away at the frontier of human knowledge in this field. I am reminded of the [lightly paraprhased] quote by theorist Nima Arkani-Hamed: "You can learn quantum mechanics in a few weeks. But you need to dedicate a lifetime of research to hope to understand quantum field theory."

This course is an attempt to give a working knowledge of the big picture. You will see how Feynman diagrams are both a perturbative expansion of a transition amplitude *and* a useful mnemonic for the story of particle scattering. You will see how indices are a physicist's crutch to mathematically implement symmetries. Along the way, you may come to appreciate corners of quantum mechanics and relativity that otherwise may slip the usual undergrad curriculum.

Most of call, *this course is a bridge*. For those who are interested in theoretical physics of any type[b] or those interested in particle physics of any type[c]: I want you to leave this course knowing how one *uses* quantum field theory in particle physics so that if you go to graduate school, you will already have the bird's-eye-view of how different technical ideas come together. Alternatively, for those who know that their passions are not in this discipline: I want you to appreciate what the scaffolding of quantum field theory buys us as physicists, so that when you go off to become your future self, you are an informed ambassador of physics.

[b] In some sense, all theoretical physicists are quantum field theorists. Yes, even condensed matter theorists.

[c] Including experimental particle physics and astro-particle physics.

1.2 What we miss

This is *not* a 'modern physics course—what I mean by that are courses that are glorified "physics for poets" courses that offer ideas without mathematical rigor. Instead, this course is *all about* understanding the mathematical rigor, even though we will not derive every step (leaning instead on analogies as appropriate) and even though the purpose of the course is not to see who can calculate the most tedious cross section.

A course like this can span multiple terms and focus on many different aspects of particle physics. Given that we have *one* term and that we want to start by assuming the bare minimum, we have to make deliberate cuts to what we investigate. Rather than trying to do a little bit of everything, we will dive deeply into the theory and sacrifice the following:

- The experimental foundation of particle physics. This is the biggest sacrifice because physics is ultimately an *empirical* science and practitioners need to be grounded in experiment.[a] We excise this aspect in part because I would struggle to do it justice, but also because we are in a moment where the types of experiments that particle physicists do has evolved rapidly over the last decade.[b]
- The history of particle physics. I did not appreciate this as a student, but good physicists understand how those who came before them had their key *aha!* moments: what what puzzles where they thinking about, how did they make progress?[c]
- Computational tools. Particle physicists were the original 'big data' scientists with a huge throughput of data in particle colliders. Experience with some of the computational tools are a great way to get into undergraduate research in this field. I regret that this is something that we cannot fit into this course, but I encourage those interested in this field to be prepared to do computational work.

[a] Even theorists.

[b] UCR students are encouraged to reach out to members of our experimental particle cosmology group to learn more about this.

[c] This is different from a hagiography of physics heroes. Most of our physics heroes are flawed human beings and most of the heroism is rooted in a broader collective of people than our stories usually tell.

1.3 Prerequisites

I have done my best to minimize the prerequisite knowledge for this course. At the bare minimum, we require the following:

1. This means you have had first-year physics and did well.[a]
2. Analytical mechanics at the level of having some familiarity with Lagrangian mechanics and variational principles.
3. At least two quarters of quantum mechanics so that you are familiar with bras, kets, operators, superposition, and amplitudes. You should be comfortable with angular momentum and spin.
4. Some introduction to special relativity so that you are familiar with the principles of length contraction and time dilation.
5. Linear algebra at the level of Physics 17 at UCR.[b]

[a] At the very least, if you took first-year physics now you would ace it.

[b] The course notes for Physics 17 may be a useful reference.[1]

In a perfect world, you would also have the following background:

- A solid background in linear algebra and some familiarity with how this relates to the representation of groups. It would help if you do not cringe at the word *tensor*. It would be even better if you knew what makes a tensor a tensor as opposed to a multi-dimensional array of numbers.
- Some familiarity with the notion of a generator of a transformation and its exponentiation, for example from a quantum mechanics course.
- An idea of what a Green's function is.[c] We will not use this word too often, but understanding what it means usually correlates to a mathematics and physics background that will be helpful.
- Some introduction to relativistic quantum mechanics; for example, knowing the Dirac and the Klein-Gordon equations.
- Know what a cross section is from a mechanics course or, even better, Fermi's Golden Rule from a quantum mechanics course.

[c] The course notes for Physics 231 may be a useful reference.[2]

We do not live in a perfect world and it is much better to take the pioneering spirit of jumping in. As one of my undergraduate mentors told me, *enthusiasm makes up for a lot of things.*[d]

[d] Enthusiasm can be measured by how much time do you set aside to learn the *extra* stuff that you need to best appreciate this course.

1.4 What it takes to succeed in this course

Do your homework I tell the Physics 39 classes that the best advice I can give is to *do your homework.*[a] Trust me, I get no particular joy assigning or reviewing your homework—unlike many other topics I may teach, I actually know this shit. The value of homework is *practice.*[4] This is my commitment to you to cobble together an opportunity for *you* check your understanding, to learn more (and more meaningfully), and to develop your own style of *doing physics*. Part of my commitment to you is to create a space—our class—where we can work together on these meaningful exercises. I encourage you to do the homework, this is where you *become a physicist.*

[a] Whenever I do this, I think about the Mary Schmich essay and Baz Luhrmann song, "Wear Sunscreen." The song plays in the back of my head every time I talk about homework.[3]

Ask questions This course can be technical, it draws on different branches of physics, and even then it is just scratching the surface of some of the biggest open questions in science. If you engage with the material deeply enough, I *expect you to be confused.* I am often confused. Do everyone a favor and ask during class. If you are embarrassed, I suggest phrasing your question as follows:

> Is it obvious that...?

[1] https://sites.google.com/ucr.edu/physics017/
[2] https://sites.google.com/ucr.edu/p231/
[3] https://en.wikipedia.org/wiki/Wear_Sunscreen
[4] https://www.youtube.com/watch?v=p-BR1mXwtB0

When you ask it this way, the question is no longer about you being confused, it is about what is the most insightful way to think about a topic.

How to answer questions The best part of a class is the community that we build together. I may be the person in the class who is most familiar with this material, but you and your classmates are the ones who are familiar with the journey that you're on—and there is a lot that I learn from your journey. As such, expect me to *ask you questions*. This can be scary, but know that when I ask a question, I am not testing *you* and what you have learned, I am testing *me* and what I need to teach. If you find yourself on the business end of a question during class, consider the following options:

- If you know the answer, give the answer and a justification.
- If you are not sure but think you know the answer, give the answer and a justification. You can even say that you are not sure and give reasons to be skeptical.[b] This is useful for me as well.
- If you have no idea what the answer is, then say so and explain what is confusing about the question. If you do not understand the question, then say that you do not understand the question and then ask a specific question for clarification.

[b] This is part of the scientific method.

1.5 Significant figures

Most of the big ideas in this course do not require precise numbers. Usually we can get by with one significant figures—the mass of the proton is 1 GeV. Sometimes we can get by with *zero* significant figures—the Planck scale is on the order[a] of $\mathcal{O}(10^{19}\,\text{GeV})$. In fact, you should start every problem by thinking about the order of magnitude before you think about any significant figures.

[a] The 'big-O' notation $\mathcal{O}$ means 'order of magnitude. Because our study of particle physics spans the very small to the very large, you would be wise to be comfortable with this notation.

The only time where we will have to get into the weeds and sort one more than one significant figure is when we need to take the difference of two numbers and it matters whether the difference is positive or negative.[b]

[b] Of course, this is simply saying that we care about the difference to one significant figure.

Example 1.5.1. A good example of this is the mass of the proton versus the mass of the neutron. These, in turn, are due to small differences in the mass of the up versus the down quarks. The masses are:

$$m_\text{p} = 938.3\text{ MeV} \qquad\qquad m_\text{n} = 939.6\text{ MeV} \tag{1.5.1}$$

and their mass difference is on the order of a percent of the masses. If proton were heavier than the neutron—a change in the at the *third* significant figure—then the universe as you know it would be radically different. How different? For starters, instead of hydrogen atoms we would have neutrons—and we can say goodbye to chemistry.[c]

[c] This and related observations are described beautifully in Robert Cahn's article "The eighteen arbitrary parameters of the standard model in your everyday life," which I strongly recommend that everyone read. Cahn is also the author of a great book on representation theory for those learning the subject on their own[5].

[5] Robert N. Cahn. "The Eighteen arbitrary parameters of the standard model in your

1.6 These notes

This is a first draft of these lecture notes. You can expect both outright errors and explanations that may not be fully satisfying. Please bring up any questions to me at your soonest convenience.

I also have a tendency towards marginalia—sidenotes and footnotes. It is a habit I pick up from my favorite textbook authors (Tony Zee in particular); I feel like it brings a bit of the flavor of the course into the text. They also reflect my occasionally scatter-brained enthusiasm for this subject. If you are in a particular hurry, you can safely skip the side notes. I have tried to include references on the page that we use them rather than in a comprehensive bibliography in the back. Students who are especially dedicated to this subject are encouraged to pursue as many of these references as they are able to.

1.7 References

Unfortunately there is not a perfect book on particle physics at this level.[a] Here are a few that you may consider.

[a] Jeff Richman at UCSB is working on one that may be close. At the time of this writing that book is not yet complete.

1.7.1 Particle Physics

- David Griffiths. *Introduction to elementary particles*. 2008. ISBN: 978-3-527-40601-2
- Andrew J. Larkoski. *Elementary Particle Physics: An Intuitive Introduction*. Cambridge University Press, June 2019. ISBN: 978-1-108-49698-8, 978-1-108-57940-7
- Michael E. Peskin. *Concepts of Elementary Particle Physics*. Oxford Master Series in Physics. Oxford University Press, Sept. 2019. ISBN: 978-0-19-881218-0, 978-0-19-881219-7. DOI: 10.1093/oso/9780198812180.001.0001
- R. N. Cahn and G. Goldhaber. *The experimental foundations of particle physics*. Cambridge: Cambridge Univ. Press, 2009. ISBN: 978-0-521-52147-5. DOI: 10.1017/CBO9780511609923
- Dave Goldberg. *The Standard Model in a Nutshell*. Princeton University Press, Feb. 2017. ISBN: 978-0-691-16759-6
- Alessandro Bettini. *Introduction to elementary particle physics*. 2008. ISBN: 978-0-521-88021-3

everyday life". In: *Rev. Mod. Phys.* 68 (1996), pp. 951–960. DOI: 10.1103/RevModPhys.68.951; and R.N. Cahn. *Semi-Simple Lie Algebras and Their Representations*. Dover Books on Mathematics. Dover Publications, 2006. ISBN: 9780486449999

1.7.2 Introductions to Quantum Field Theory

Recently there has been an explosion in the number of "quantum field theory for undergraduates" textbooks. Here are a few that I think are particularly effective for further study. The books all start at the level assumed for this course, but each goes much deeper into the subject and wold be well suited for anyone interested in pursuing theoretical research in graduate school.

- Jakob Schwichtenberg. *Physics from Symmetry*. Undergraduate Lecture Notes in Physics. Cham: Springer International Publishing, 2018. ISBN: 978-3-319-66630-3, 978-3-319-66631-0. DOI: 10.1007/978-3-319-66631-0
- J. Schwichtenberg. *No-Nonsense Quantum Field Theory: A Student-Friendly Introduction*. No-Nonsense Books, 2020
- Tom Lancaster and Stephen J. Blundell. *Quantum Field Theory for the Gifted Amateur*. Oxford University Press, 2014. ISBN: 978-0-19-969933-9
- J. Donoghue and L. Sorbo. *A Prelude to Quantum Field Theory*. Princeton University Press, 2022. ISBN: 9780691223506
- Anthony Zee. *Quantum Field Theory in a Nutshell: Second Edition*. Princeton University Press, Feb. 2010. ISBN: 978-0-691-14034-6
- Richard P. Feynman. *QED: The Strange Theory of Light and Matter*. Princeton University Press, Oct. 2014. ISBN: 978-0-691-16409-0, 978-0-691-02417-2
- M. J. G. Veltman. *Diagrammatica: The Path to Feynman rules*. Vol. 4. Cambridge University Press, May 2012. ISBN: 978-1-139-24339-1, 978-0-521-45692-0

1.7.3 Bird's eye view of particle physics

- The "Pathways to Innovation and Discovery in Particle Physics" Report of the 2023 Particle Physics Project Prioritization Panel is a great reference for the present state of the discipline.[6][b]
- Fermilab and SLAC, two of our flagship national laboratories in particle physics, have a for-the-public online magazine, *Symmetry: Dimensions of Particle Physics*. This is a great starting point to dig a bit deeper into what is going in in particle physics.[7] You may also find a discussion of more mathematical topics in some of the articles in *Quanta Magazine*[8].
- There are a few well-known histories of particle physics ostensibly written for the general public. These include *The Rise of the Standard*

[b] For those pursuing graduate study in particle physics: this document outlines scientific priorities of US particle physics funding agencies. These priorities tend to align with the topics where research groups are looking for new graduate students. You may want to read the relevant parts carefully.

[6] https://www.usparticlephysics.org/2023-p5-report/
[7] https://www.symmetrymagazine.org
[8] https://www.quantamagazine.org/physics/

Model[9] and *Inward Bound*.[10]

- In many ways, this field still exists in the shadow of the Superconducting Supercollider. The definitive history of this unfortunate saga is the book *Tunnel Visions*.[11] YouTube documentarian BobbyBroccoli has an excellent movie-length three-part synopsis.[12]
- There have been a few celebratory conference on the Standard Model. Steven Weinberg's 2003 talk[c] is a succinct history of both the successes and false directions—it is a nice chronological scaffolding to see how some of the big ideas in particle physics came to be. There is a great video archive of talks from the SM50 conference at Case Western in 2018[13] and the 50 Years of QCD conference at UCLA in 2023[14].

[c] arXiv:hep-ph/0401010

Exercise 1.7.1. Perhaps it is a bit silly that the celebration of 50 years of quantum chromodynamics—a *part* of the Standard Model—was a few years after the celebration of the Standard Model. As an exercise, explain the chronology of what specific events were being celebrated at each event.

[9] L. Hoddeson et al. *The Rise of the Standard Model: A History of Particle Physics from 1964 to 1979*. Cambridge University Press, 1997. ISBN: 9780521578165.

[10] A. Pais. *Inward Bound: Of Matter and Forces in the Physical World*. Oxford paperbacks. Clarendon Press, 1988. ISBN: 9780198519973.

[11] M. Riordan, L. Hoddeson, and A.W. Kolb. *Tunnel Visions: The Rise and Fall of the Superconducting Super Collider*. University of Chicago Press, 2015. ISBN: 9780226305837.

[12] https://www.youtube.com/watch?v=3xSUwgg1L4g

[13] https://www.youtube.com/playlist?list=PLBELrG1nZ2U6H3I1il4NhNVNcdUVUE5Ye

[14] https://www.youtube.com/playlist?list=PLjqOpQgpjtKoy21wrZ9hPnYM1KN97WYJH

Chapter 2

Introduction

This is a course on the theory of the Standard Model of particle physics. It is a theory so successful that it is called the *Standard Model* with capital letters. Among its crowning achievements is the discovery of the Higgs boson[1] and the verification of the electron's anomalous magnetic moment[2]—the most precisely predicted and measured number in nature.[a]

[a] The error is less than a part per trillion.

This chapter is an *amouse bouche* for the course. We review some of the key themes and big ideas to prepare you for a systematic study in subsequent chapters. This chapter may dip into some jargon and hint at some of the deeper undercurrents of quantum field theory. I include them here not to discourage you, but to whet your appetite for the journey ahead.

2.1 What is Particle Physics

Particle physics has had a few different names. It used to be called *subatomic physics* because it is the study of things that are smaller than the atom. But then nuclear physics split off as a discipline[a] and what does it mean for a particle to be 'smaller' anyway?

[a] By the end of this course, you may come to appreciate why this is the case.[3]

> **Exercise 2.1.1.** If someone who is not a physicist asked you what the size of the electron is, what would you say? Go ahead and look up the radius of an electron. What does this number mean? What about a photon? What is the 'size' of a

[1] Serguei Chatrchyan et al. "Combined Results of Searches for the Standard Model Higgs Boson in pp Collisions at $\sqrt{s} = 7$ TeV". in: *Phys. Lett. B* 710 (2012), pp. 26–48. DOI: 10.1016/j.physletb.2012.02.064. arXiv: 1202.1488 [hep-ex]; Georges Aad et al. "Combined search for the Standard Model Higgs boson using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC". in: *Phys. Lett. B* 710 (2012), pp. 49–66. DOI: 10.1016/j.physletb.2012.02.044. arXiv: 1202.1408 [hep-ex].

[2] X. Fan et al. "Measurement of the Electron Magnetic Moment". In: *Phys. Rev. Lett.* 130.7 (2023), p. 071801. DOI: 10.1103/PhysRevLett.130.071801. arXiv: 2209.13084 [physics.atom-ph].

[3] Nuclear physics works with *effective theory* of hadrons at the GeV scale. This theory is distinct from the Standard Model because it is precisely the regime where the theory of quarks is non-perturbative and alternative descriptions are necessary.

quantum of light?[b]

[b] Confused? Good. You're thinking.

Another historical name for particle physics is *high-energy physics*. This is because our experimental apparatuses are traditionally colliders where we smash together particles at high energies.

Exercise 2.1.2. What does it mean for a particle to have a "high energy"? What is this energy being compared to? The kinetic energy of a mosquito flapping its wings is negligible for most human-scale activities, but it is quite substantial when smashing together two electrons.

This is because high energy colliders are actually microscopes. Go ahead and review Rutherford scattering in your favorite quantum mechanics book. Rutherford scattering is the famous experiment that lead to the discovery of nuclei by shining an electron beam onto a gold foil target.

Exercise 2.1.3. How is the energy of the electron beam related to the spatial resolution of the experiment?

In this way, one may be forgiven for conflating particle physics with *collider* physics. Indeed, particle physics is still a field that looks to the Large Hadron Collider as one of its Meccas. However, unlike the case even twenty years ago, modern particle physics is far more diverse than 'just' high-energy [collider] physics.[c] As of the time of this writing, a pretty good summary of the field is the 2023 Particle Physics Project Prioritization Panel (P5) report.[4] The report is the guiding document for the Department of Energy and National Science Foundation—the primary funding agencies for particle physics in the US—for the key questions that particle physics seeks to answer over the coming decade.

[c] Theorist Hitoshi Murayama once said that this field stretches *"from the smallest scales to the largest, from the heavens to the hell."* By this he was referring to the most fundamental particles, the cosmological evolution of the entire universe, space-based telescopes, and underground detectors.

So I am calling this field **particle physics**. Even this name is debatable: while the objects that we *observe* are *particles*—usually energy and momentum eigenstates, perhaps described by some sort of creation operator $a^\dagger$—the underlying mathematical objects in our theory are *quantum fields*. Even then, some phenomena in particle physics seem to probe the decidedly 'wavey' limit of a quantum field.[5] You could call this "applied quantum field theory," but applications of quantum field theory are much broader than particle physics. So for our purposes, let's call it particle physics and not get too hung up on minutiae.

So what is particle physics?

> In this course, particle physics is the branch of science that seeks to build a model quantum field theory that describes the elementary constituents of nature and their fundamental interactions.

[4] `https://www.usparticlephysics.org/2023-p5-report/`

[5] See, for example, wave dark matter, `arXiv:2101.11735`.

Figure 2.1: What is the difference between these two? One is a screenshot of the Millennium Falcon from *Wookiepedia*, the other is a picture of a *model* of the Millenium Falcon from Lego. Physicists build models of nature. We hope that the model may illuminate principles of how nature works, but we do not confuse the model as a prescription for how nature must behave.

We should parse a few key words here:

- Quantum field theory is a theory-of-theories. The Standard Model of particle physics is *a* quantum field theory in the same way that the hydrogen atom is a model quantum mechanical system.
- We are building *models*. A model is a mathematical description of an actual system. The mathematical description is usually idealized and hence simplified. The model is a way to learn about how underlying principles (usually symmetries) can lead to particular phenomena. See Fig. 2.1.
- We talk about the *elementary constituents* of nature. In our model, these are quantum fields. Excitations of these fields are what we call elementary particles. The name is meant to evoke some sense of indivisibility, but this is not at all necessarily the case.[d] For example, it can be convenient to write models where protons and neutrons are elementary, even though we know that they are not.
- Finally, the *fundamental interactions* between particles is conventionally what are called forces. We shall see in this class in particle physics, there is no strict distinction between particle and force—they are all described by quantum fields whose excitations are particles.[e]

[d] I once read a story about a woman on a plane who was bragging about how brilliant her son is. When she saw her seatmate reading a book, *Elementary Particle Physics*, she said "hmph! Well, my son studies *advanced* particle physics."

[e] In particle physics, everything is a particle. In contrast, to astronomers, subatomic scientists are either particle physicists or a metallurgists.[6]

These ingredients come together through a tool called a Feynman diagram.

2.2 Diagrammar

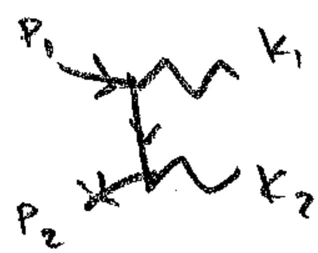


Figure 2.2: Example of a Feynman diagram. Here an electron and a positron annihilate into a pair of photons.

Fig. 2.2 is an example of a Feynman diagram. Feynman diagrams are a perturbative expansion of the quantum mechanical amplitude for a something to happen. You may recall that the quantum amplitude is a "square root of the probability[7]." The statement that diagrams are a *perturbative expansion* means that there is some small parameter for which we are performing a Taylor expansion[a]. In the standard case, this perturbative expansion is

[a] Formally the amplitudes are complex and may contain singularities. It is thus more appropriate to say that this is a *Laurent expansion*. One of the 'deep' ideas in quantum field theory is the intimate relationship of *analyticity* (complex differentiability) to physical properties. To dig deeper, see my Physics 231 notes.[8]

[6] https://en.wikipedia.org/wiki/Metallicity

[7] This notion comes from quantum mechanics where the probability of observing a state $|\Psi\rangle$ is $\langle\Psi|\Psi\rangle = |\Psi|^2$. In quantum field theory we typically talk about *cross sections*, σ. The relation between the amplitude $\mathcal{M}$ and cross section is $\sigma \sim |\mathcal{M}|^2$ and contains kinematic factors. We determine these kinematic factors in subsequent chapters.

usually an expansion in couplings.

By **coupling** I mean a parameter of the theory that determines how much some particles interact with each other. When the coupling is large, the interaction is very strong. When the coupling is small, the interaction is very weak. One coupling that you may be familiar with is the electrodynamic coupling, e. You are probably most familiar seeing e as an ingredient in the fine structure constant,

$$\alpha = \frac{1}{\hbar c \varepsilon_0} \frac{e^2}{4\pi} \approx \frac{1}{137} \ . \tag{2.2.1}$$

The first factor of $(\hbar c \varepsilon_0)^{-1}$ are relics of using silly units. When we use natural units—see Sec. 2.4—these are set to one. You can see that $1/137$ is a small number, so it at least makes sense that if we had some amplitude $\mathcal{M}(\alpha)$ that is a function of the electrodynamic couplings through α that we could imagine doing the perturbative expansion

$$\mathcal{M} = \mathcal{M}_0 + \alpha \mathcal{M}_1 + \alpha^2 \mathcal{M}_2 + \cdots \ , \tag{2.2.2}$$

and then dropping any subleading terms since we expect them to be percent-level corrections. If the couplings are large, then this expansion breaks down because subsequent terms are not small. In fact, this is what happens with the strong interactions (quantum chromodynamics), the force that holds nuclei together. Thus there are regimes where the usual Feynman expansion fails: it seems like we should not use these diagrams to describe the interactions of the quarks and gluons that are 'inside' a proton.

> **Example 2.2.1.** I seem to have implied that Feynman diagrams do not work for the strong interaction. Despite this, collider physicists working on the Large Hadron Collider 'speak' the language of Feynman diagrams. They'll even draw diagrams that involve the strong force. What gives? Apparently I haven't told you the whole story...

Note that I wrote *couplings* not the common phrase *coupling constants*. That is because—brace yourselves—these couplings are generally *not* constant. In fact, they depend on the energy scale at which you probe them. If I smash together color-charged particles[b] at high energies, the analogous fine structure parameter for the strong force, $\alpha_s = g_s^2/4\pi$ depends on the characteristic energy scale[10] $\sqrt{Q^2}$ at which one probes the interaction, see Fig. 2.3. This idea should be shocking the first time that you see it. The 'coupling constants' are not constant at all. The meaning of this oddity is an idea called *renormalization* and is rooted in making sense of what the *actual* small parameter is in perturbation theory: it turns out not to be the

[b] In this class and in this field, we write *colored* to mean color-charged, or charged with respect to the strong force. See Chapter 5 of *The Disordered Cosmos* for an anthropological discussion.[9]

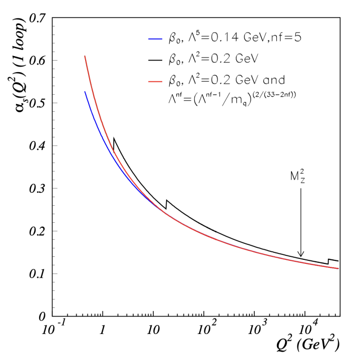


Figure 2.3: Approximate value of the strong force fine structure parameter, α_s. Lines correspond to slightly different calculations. The horizon-

[8] https://sites.google.com/ucr.edu/p231/

[9] C. Prescod-Weinstein. *The Disordered Cosmos: A Journey into Dark Matter, Spacetime, and Dreams Deferred*. PublicAffairs, 2021. ISBN: 9781541724693

[10] We write $\sqrt{Q^2}$ rather than E because Q^2 is a Lorentz-invariant quantity, as we explain below in our review of special relativity.

fine structure constant, but the fine structure constant times a function of the kinematics of a process. This means that at sufficiently high energies, we can meaningfully talk about Feynman diagrams of quarks exchanging gluons. However, at low energies, these diagrams lose their meaning. If you were paying attention in the previous sub-section, this is the regime where particle physics becomes nuclear physics.

Example 2.2.2. This reminds me of the difference between a *physicist* and a *physics fan*. A physics fan is someone who thinks that $1/137$ is a fundamental constant of the universe and so should be tattooed on their body. A physicist stops to think: *Where does this number come from?* and proceeds to learn about the scale-dependence of the electric coupling. In fact, in this course you will find that the electric coupling is not even a fundamental parameter but a combination of more fundamental parameters. I do not care whether or not you are a physics fan, but my goal is to train you as a physicist.

Exercise 2.2.1. On the subject of couplings: is the gravitational coupling *large* or *small* relative to the electromagnetic coupling? You may recall from popular physics that one of these is [surprisingly?] much stronger than the other. Try to make this quantitative by comparing two numbers. HINT: this is a bit of a trick question. Try writing out the gravitational fine structure constant and argue why it appears that you cannot meaningfully compare it to the electromagnetic coupling.

Feynman diagrams are graphs that represent a mathematical expression for a complex number. The graphs are trajectories in spacetime that we read from left to right. Each line in a diagram represents a particle. The lines may have decorations or labels that indicate their identity. In Fig. 2.2 the two lines on the left are an electron and a positron. The two wiggly lines on the right are each photons. Each vertex (intersection of lines) represents a factor of the coupling. In Fig. 2.2 the two vertices mean that this diagram contains two powers of the electric coupling, e. Thus this diagram contributes to the $\mathcal{O}(e^2) = \mathcal{O}(\alpha)$ term in the expansion of the amplitude $\mathcal{M}(e^+e^- \to \gamma\gamma)$ in (2.2.2). The notation $\mathcal{M}(e^+e^- \to \gamma\gamma)$ simply means the amplitude for an electron e^- and a positron e^+ to turn into two photons, $\gamma\gamma$.

Something else that may be familiar from quantum mechanics: amplitudes sum together. This is the origin of quantum interference and all of the fun parts of quantum physics. Our perturbative expansion (2.2.2) is one example of such a sum. However there are usually multiple diagrams that contribute at a given order in perturbation theory. One of the brilliant things about these diagrams is that they have a complementary interpretation: as a *sum over histories*. This is an idea that we emphasize in our review of quantum mechanics, but the idea is this:

> The amplitude to go from some initial state to some final state is represented by the sum of all Feynman diagrams that connect the initial state to the final state. Each individual Feynman

diagram in the sum represents a possible history that the initial state could have taken to reach the final state.

This sum over histories interpretation is outrageous the first time you see it but is ultimately an extension of the principle of least time that underlies Snell's law in optics. The generalization to a principle of least action is the core of Lagrangian mechanics and was in fact Feynman's Ph.D thesis.[11]

The diagrams are simply a tool. Feynman's long time physics rival and co-Nobel prize winner, Julian Schwinger, was a master of calculating amplitudes in quantum field theory *without* any diagrams. In one interpretation of history, Feynman's main contribution at this stage of physics history was bringing that calculational technology to the masses by giving each term an intuitive meaning.[c] Veltman has a nice physics-oriented history of the spread of Feynman diagrams in his book *Diagrammatica*.[13] This book should not be confused with a set of lecture notes he co-wrote with 't Hooft called "Diagrammar,[14]" which inspired the name of this section.

[c] It is no surprise that he is also (mis-)attributed as the spokesperson of the "shut up and calculate" school of quantum physics.[12]

As a final note: there is far more to quantum field theory than Feynman diagrams. As a perturbative expansion, Feynman diagrams represent the *easiest* part of quantum field theory. But just as there are functions that cannot be meaningfully Taylor expanded[d], there are vast swaths of field theory that are intrinsically *non-perturbative*. Quantum chromodynamics at low energies—where we must turn to methods in nuclear physics—is one example. We make this caveat to emphasize that although we lean heavily on the diagrammatic interpretation of particle physics, we are still just scratching the surface.

[d] Try expanding $\exp(-x^{-2})$ around the point $x_0 = 0$. Every term at every order vanishes, even though the function is well defined and non-zero away from the origin. The quantum field theory analog is something called an instanton and is beyond the scope of this course.

2.3 Kinematics and Dynamics

I may be the only person who makes a big deal about this, but we can separate our study of particle physics into two parallel tracks: kinematics and dynamics. You probably know that these words "have to do with physics," but the distinction between them is rarely delineated. In fact, I suspect I may be making it up—in which case, here are the working definitions for this class.

Kinematics has to do with the motion of particles through space and time. The kinematics of a scattering process relates to the momentum and energy of those particles. The conservation of energy and momentum are also kinematical facts[a]. The thread of physics that is most relevant to this

[a] Though their derivation through Noether's theorem is arguably a statement about dynamics.

[11] R.P. Feynman and L.M. Brown. *Feynman's Thesis: A New Approach to Quantum Theory*. World Scientific, 2005. ISBN: 9789812563668.

[12] N. David Mermin. "Could Feynman Have Said This?" In: *Physics Today* 57.5 (May 2004), pp. 10–11. ISSN: 0031-9228. DOI: 10.1063/1.1768652. URL: https://doi.org/10.1063/1.1768652

[13] M. J. G. Veltman. *Diagrammatica: The Path to Feynman rules*. Vol. 4. Cambridge University Press, May 2012. ISBN: 978-1-139-24339-1, 978-0-521-45692-0.

[14] Gerard 't Hooft and M. J. G. Veltman. "Diagrammar". In: *NATO Sci. Ser. B* 4 (1974), pp. 177–322. DOI: 10.1007/978-1-4684-2826-1_5.

study is *relativity*, and in particular the flat-spacetime version known as *special relativity*.[b]

[b] In contrast, general relativity is the study non-flat spacetimes. One can argue that it is the *dynamics* of spacetime itself.

Dynamics, on the other hand, has to do with the rules for how particles interact. When we talk about a theory or a model of particle physics, we usually mean a description of the dynamics. These are encoded in the action or Lagrangian of a theory. The dynamics of a theory draws primarily from the rules of *quantum mechanics*.

Feynman diagrams are an output of the dynamics of a theory. However, a Feynman diagram is only physically meaningful if it obeys the rules of the kinematics of a theory. Kinematics are an additional consideration when we convert the squared amplitude into something physically measurable.

Example 2.3.1. You can draw a Feynman diagram for a process like $\gamma \to e^+e^-$ by which a photon decays into an electron–positron pair. You could even calculate the amplitude for this process to happen. However, this process is kinematically forbidden because it cannot simultaneously conserve energy and momentum. The amplitude is nonzero, but the decay rate is forced to be zero by kinematics.

Exercise 2.3.1. Show that both energy and momentum cannot be conserved in $\gamma \to e^+e^-$. There are many ways to do this, including some that are more slick than others. If you are stuck, come back to this exercise after our review of special relativity.

In this course, you can think of dynamics as the set of rules[c] that tell us how we may construct Feynman diagrams. These rules are an encoding of the Lagrangian of the theory. You can think of kinematics as conditions on the energies and momenta of the initial and final states of an amplitude—these are the external lines of a diagram. Notably, kinematic constraints do not apply to the particles on the *inside* of a Feynman diagram. Lines that obey the kinematic constraints are called **on shell**, while those that do not are called **off shell**.[d] With this jargon, we say that external lines on a Feynman diagram represent parts of the initial or final states of a process and must be on shell. Internal lines are, in general, off shell.

[c] Called *Feynman rules*.

[d] What is the *shell*? It is the hypersurface the four-dimensional space of energy and momentum that satisfies the Einstein relation, $E^2 = m^2c^4 + p^2c^2$ for a particle of mass m, energy E, and three-momentum $\boldsymbol{p}$.

2.3.1 Symmetry

The notion of *symmetry* plays a central role for both the kinematics and the dynamics. The mathematical description of symmetry is called group theory and the way in which symmetries act on objects is called representation theory.[e] The tables of Clebsch–Gordan coefficients that you may have invoked in your study of addition of angular momentum in quantum mechanics is an output of the representation theory of the group of three-dimensional rotations. In this course we are specifically interested in the representation theory of continuous groups—symmetries like rotations where you can transform by an arbitrary amount—called *Lie groups*.[f] As humble physicists[g] the way we work with symmetries is to introduce indices. Objects that carry these indices are called **tensors**.

[e] I am name-dropping subjects because I am often asked what subjects should an aspiring theoretical physicist master.

[f] Pronounced 'lee.'

[g] Oppenheimer: Well, if that's how you treat a lieutenant colonel than I hate to see how you treat a humble physicist.
Leslie Groves: If I ever meet one I'll let you know. (From *Oppenheimer*, 2023)

In special relativity (kinematics) the simplest tensors are four-vectors. They are called four-vectors because they are vectors that have four components. For example, the four-momentum of a particle may be written p^μ. The index is μ. There is a related object called p_μ with a lower index. These are related by a tensor called the metric tensor, which for our purposes we write $\eta_{\mu\nu}$. The metric gives us a way to define an inner product. This should all sound familiar from linear algebra because this *is* linear algebra. In representation theory, the objects that get rotated[h] are vectors in a vector space. If you ever wonder why I teach Physics 17 the way that I do, it is because I want students to be primed to understand representation theory as it appears in physics.

[h] By 'rotate' I mean a general symmetry transformation. For the case of special relativity, one can have boosts in addition to rotations.

Maybe the phrase 'inner product' caught your ear. This is the same idea that comes up in quantum mechanics. In fact, now you may recall that quantum mechanics really boils down to complex linear algebra. In fact, many ideas in quantum mechanics are ultimately group theoretical. For example, the commutator is the natural multiplication operation between elements of a group.[i] Furthermore, finite transformations are the exponentiation of infinitesimal transformations. For example, the Hamiltonian $\hat{H}$ is the generator of translations in time. A finite translation in time is

[i] Check that while this may be surprising, it is sensible. The commutator of two operators is another operator in the same way that the multiplication of two things of a given type should be another thing of the same type.

$$\hat{U}(t) = e^{-i\hbar t\hat{H}} \ . \tag{2.3.1}$$

[**Flip**: check the $\hbar$] In particle physics we will meet several **internal symmetries** that mathematically describe the rotation of an object in different vector spaces. These do *not* correspond to spacetime rotations or boosts. Instead, they may represent a rotation between quarks of different color charges.[j] The infinitesimal generators of these transformations take the form[15]

[j] We define these carefully below where we discuss quantum chromodynamics. For this introduction just humor me and go with the flow to appreciate the big idea.

$$(T^A)^i{}_j \ , \tag{2.3.2}$$

where we recognize three indices. The index A is called an adjoint index and tells you which direction you are rotating. For the rotation group, A takes values from 1 to 3 corresponding to rotations about the x, y, and z axes. All other rotations are combinations of these. The other two indices, i and j depend on the *representation* of the object that we are rotating.

Example 2.3.2. In quantum chromodynamics we there are eight generators of so-called color symmetry. This means A takes values from 1 through 8. This group is called $\mathrm{SU}(3)$, which stands for the set of 3×3 special unitary matrices.[k]A quark has indices that I conventionally write with lowercase letters from the middle of the Roman alphabet that take values m from 1 to 3 corresponding to red, blue, and green. The matrix $(T^4)^m{}_n$ represents a particular rotation around

[15] I am using 'physicist' shorthand here and referring to a tensor by its components. Supercilious mathematicians sneer at us for this. Formally, $A^i{}_j$ is not a matrix, it is the *i–j component* of a matrix A. Physicists justify our sloppiness because anyone who is paying attention should understand what we mean and, more importantly, by keeping the indices explicit we can see how the tensor transforms.

the $A = 4$ axis. A finite transformation by angle θ takes the form

$$q^m \to \sum_n e^{i\theta(T^4)^m{}_n} q^n \equiv \sum_n U(\theta)^m{}_n q^n \ , \tag{2.3.3}$$

where the sum over n is what we expect from matrix multiplication.

[k] Special means unit determinant. Unitary, as you may recall from quantum mechanics, means that the hermitian conjugate is its inverse.

This is all to say that indices feature front-and-center in this course. They are a crutch for us to talk about the underlying symmetries that govern both the kinematics and dynamics of particle physics. We will spend a good chunk of this course building familiarity with how to interpret and manipulate indices. This is a mathematical skill that is far more general than particle physics itself.

2.4 Natural Units

By this stage of your physics career you are an expert at converting units. The trick is to multiply be one in different forms. Suppose you have some unit x that is related to unit y by some prefactor,

$$x = ay \ . \tag{2.4.1}$$

Then you can derive that

$$1 = \frac{ay}{x} = xay \ . \tag{2.4.2}$$

Then if some quantity is, say, $3.4\,x$, you know that you can write it out in terms of y simply by multiplying by one, cleverly written:

$$3.4\,x = 3.4 \times 1 \times x = 3.1 \times \frac{ay}{\cancel{x}} \times \cancel{x} = (3.4a)\,y \ . \tag{2.4.3}$$

(2.4.2) tells us that there is a universal, unambiguous constant ratio that relates unit x to unit y.

Example 2.4.1. Suppose someone tells you the number of feet in a mile,

$$1 \text{ mile} = 5280 \text{ feet} \ . \tag{2.4.4}$$

This number just so happens to be the mass of the B meson in MeV. You can derive that

$$1 = 5280 \ \frac{\text{feet}}{\text{mile}} = \frac{1}{5280} \ \frac{\text{mile}}{\text{feet}} \ . \tag{2.4.5}$$

From this you can deduce that a distance of 1.5 miles is

$$1.5 \,\text{mile} = 1.5 \cancel{\text{mile}} \times \left(5280 \ \frac{\text{feet}}{\cancel{\text{mile}}} \right) = 7920 \text{ feet} \approx 8000 \text{ feet} \ . \tag{2.4.6}$$

If this all looks completely simple then *good*, it is supposed to. There is nothing deep or mysterious about changing units. Let us really put it to work. **Natural units** are a convenient choice that boils down to the following identifications:

$$c = 1 \qquad\qquad \hbar = 1 \ . \tag{2.4.7}$$

That's right. The speed of light c and reduced Planck's constant $\hbar$ are set to one. This may bother you. After all, you know from past coursework that these are *not* dimensionless quantities:

$$c = 3 \times 10^8 \, \frac{\text{m}}{\text{s}} \qquad\qquad \hbar = 1 \times 10^{-34} \, \frac{\text{m}^2\text{kg}}{\text{s}} \ . \tag{2.4.8}$$

Setting $c = 1$ would then mean that there is an unambiguous way conversion between length and time, as if these were measuring the "same thing." But length is measured by rulers and time is measured by clocks: how are these the same? They are the same *precisely* because nature[a] gives us a universal, unambiguous constant ratio that relates units of length into units of time. This constant is the speed of light.

[a] All our observations since the Michelson–Morley experiment are consistent with a constant speed of light and this is built into our theory of special relativity. Theoretically this is an aesthetic unification of space and time that laid the foundation of general relativity, which in turn has passed every experimental prediction.

Example 2.4.2. A lightyear is a unit of distance. It is defined to be the distance traversed by a particle traveling at the speed of light,

$$\text{lightyear} = c \, \text{year} \ , \tag{2.4.9}$$

where we see that the speed of light in natural units $c = 1$ plays the role of a conversion factor in (2.4.1).

Identifying the speed of light as a conversion factor ends up relating another set of dimensionful quantities. All velocities in natural units are dimensionless. This is because we can simply write any velocity in units of the speed of light.

Example 2.4.3. The tangential speed of the Earth around the solar system is around $v = 200\,\text{km/s}$. In natural units this is a dimensionless number:

$$v = 200 \, \frac{\text{km}}{\text{s}} = 2 \times 10^5 \, \frac{\text{m}}{\text{s}} \times \left(\frac{1}{3 \times 10^8} \frac{\text{s}}{\text{m}} \right) = 7 \times 10^{-3} \ . \tag{2.4.10}$$

In natural units, any sensible velocity has magnitude less than one. Otherwise something is traveling faster than the speed of light.

Exercise 2.4.1. What goes wrong in physics if a particle can travel faster than the speed of light? Hint: review the relativity of simultaneity.

Velocities are dimensionless in natural units. Recall that energy has the units of mass times velocity squared. You may recall this from from

$$E_{\text{kinetic}} = \frac{1}{2} m v^2 \ . \tag{2.4.11}$$

You may argue that this formula is only true for kinetic energy. That is true, energy—no matter what the form—is carries the same type of dimension. Because velocities are dimensionless, then the dimensions of energy and the dimensions of mass must be the same. In other words, mass and energy are "the same thing." Given a particle of some mass—say the mass of a proton, m_p—there is an associated energy that is $m_\mathrm{p} \times 1 = m_\mathrm{p}c^2$. This looks remarkably like the non-relativistic limit of the Einstein relation,

$$E = mc^2 \ . \tag{2.4.12}$$

Indeed, in that limit, the Einstein relation just tells us that mass and energy are the same thing. The square of the speed of light plays the role of a conversion factor between them. It is conventional for particle physicists using natural units to measure everything in units of energy. A particularly useful energy scale is

$$m_\mathrm{p} = 1\,\mathrm{GeV} \ . \tag{2.4.13}$$

To one significant figure, the mass of the proton happens to be about one billion times an electron volt.

Exercise 2.4.2. How much do you weigh in GeV?

Sometimes we lapse into other powers of electron volt. Some useful values are the mass of the electron and the mass of the Higgs boson[b], and the center-of-mass energy of proton-proton collisions at the Large Hadron Collider:

[b] We write m_h to three significant figures because the Higgs is a big deal.

$$m_e = 0.5\,\mathrm{MeV} \qquad m_h = 125\,\mathrm{GeV} \qquad E_\mathrm{cm} = 14\,\mathrm{TeV} \ . \tag{2.4.14}$$

What about $\hbar = 1$? Planck's constant carries units of angular momentum[c], or energy times time. Using the just-established equivalence of mass and energy in natural units, this tells us that

[c] These are also the units of action, $S = \int dt\, L$.

$$\hbar = 7 \times 10^{-22}\,\mathrm{MeV\,s} \equiv 1 \ . \tag{2.4.15}$$

This means that the fundamental unit of "quantum-ness" tells us that time and inverse energy are "the same thing."

Example 2.4.4. How would you measure $\hbar$ to determine that it is a universal constant? Like measuring c, there are many options. You could look at the deflection of particles in a Stern–Gerlach experiment, measure distributions of position uncertainties given multiple measurements of a momentum eigenstate, the spectrum of hydrogen (and corrections thereof), etc.

At this point, you can multiply and divide by c and $\hbar$ as needed to write all dimensionful parameters in units of GeV to some power.

Example 2.4.5. An additional conversion is to set the Boltzmann factor, $k_\mathrm{B} = 1$. This is the observation that thermal energy is energy and can be measured in GeV.

We provide a useful table for conversions to one significant figure.

	GeV	g	K	cm^{-1}	sec^{-1}	M$_\mathrm{Pl}$
GeV		1×10^{-24}	1×10^{13}	5×10^{13}	2×10^{24}	8×10^{-20}
g	6×10^{23}		7×10^{36}	3×10^{37}	9×10^{47}	5×10^{4}
K	9×10^{-14}	2×10^{-37}		4	1×10^{11}	7×10^{-33}
cm^{-1}	2×10^{-14}	4×10^{-38}	2×10^{-1}		3×10^{10}	2×10^{-33}
sec^{-1}	7×10^{-25}	1×10^{-48}	8×10^{-12}	3×10^{-11}		5×10^{-44}
M$_\mathrm{Pl}$	1×10^{19}	2×10^{-5}	1×10^{32}	6×10^{32}	2×10^{43}	

Table 2.1: Conversion of units using $\hbar = c = k_\mathrm{B} = 1$. The row heading is equal to the table entry times the column heading so that a GeV is a small number of Planck masses, M_Pl. Adapted from Palash Pal's website.[16]

Example 2.4.6 (Mnemonics)**.** You can use your favorite equations in physics as mnemonics for natural units. We already saw how $E = mc^2$ reminds us that energy and mass both carry the same units. You can also invoke Heiseinberg's uncertainty relations

$$\Delta x\,\Delta p \sim \hbar \qquad\qquad \Delta E\Delta t \sim \hbar \tag{2.4.16}$$

to remind us that momentum and distance carry reciprocal units, as do energy and time. Because $c = 1$ tells us that distance and time are the same, we find

$$\text{length} \sim \text{time} \sim \frac{1}{\text{mass}} \sim \frac{1}{\text{energy}} \ . \tag{2.4.17}$$

The great thing about natural units is that we just have to keep track of one unit, say GeV. Dimensional analysis is very simple and we introduce the bracket notation:

$$[x] \coloneqq \text{“dimensions of } x \text{ .”} \tag{2.4.18}$$

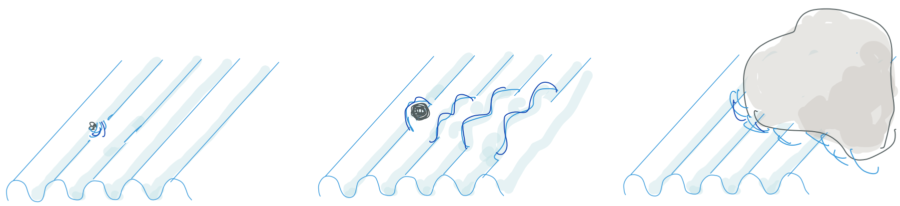
$[x]$ means: what power of energy is the unit of quantity x?

Example 2.4.7. For a distance ℓ, time t, mass m, and energy E:

$$[\ell] = -1 \qquad [t] = -1 \qquad [m] = 1 \qquad [E] = 1 \ . \tag{2.4.19}$$

Exercise 2.4.3. What are the dimensions of the gravitational constant, $[G_\mathrm{N}]$? Hint: you can use the force law for gravity to figure out the SI units of G_N.

[16] https://www.saha.ac.in/theory/palashbaran.pal/conv.html



(a) $\Delta x \ll \lambda$ (b) $\Delta x \sim \lambda$ (c) $\Delta x \gg \lambda$

Figure 2.4: The ability to probe features (like a rock) of typical scale Δx depends on having a probe whose wavelength λ is roughly on the order of Δx. If the scales are mismatched, then the waves are unable to resolve the $\mathcal{O}(\Delta x)$ features.

Exercise 2.4.4. Show that the units of action are the same as the units of angular momentum. HINT: use the expression for the action with your favorite choice of Lagrangian.

Example 2.4.8 (Collider physics as microscopy)**.** The Large Hadron Collider is a microscope. The center of mass energy of the proton–proton collisions is $E = 10$ TeV. To convert this into a length scale, we divide by .

$$\hbar c = 10^{21} \frac{1}{\text{MeV s}} \times 10^{-8} \frac{\text{s}}{\text{m}} \tag{2.4.20}$$

This gives us

$$E = 10^7 \text{ MeV} \times 10^{21} \frac{1}{\text{MeV s}} \times 10^{-8} \frac{\text{s}}{\text{m}} \tag{2.4.21}$$

$$= \frac{1}{10^{-20}\,\text{m}} \,. \tag{2.4.22}$$

The length scale associated with 10 TeV is tiny: 10^{-20} m. Compare this to the typical atomic length, Å $= 0.1$ nm $= 10^{-10}$ m. The inverse relation $\ell \sim E^{-1}$ makes it clear that increasing the energy decreases the length scale.

We can understand the $E \sim \ell^{-1}$ relation from ordinary optical microscopy. As we increase the energy of a photon, we increase its frequency and therefore decrease its wavelength, λ. We need the wavelength of the probe to be *smaller* that the characteristic size features that we are studying, $\lambda \ll \Delta x$. Figure 2.4 demonstrates this idea by showing rocks of different sizes in a stream with water waves of uniform scale λ.

Chapter 3

Rapid Review of Relativity

We begin our study of particle physics with a lightning review of selected topics in special relativity and quantum mechanics.

3.1 Kinematics

The popular Einstein relation, $E = mc^2$, is actually the low [kinetic] energy limit of the 'full' relation:

$$E^2 = m^2c^4 + \mathbf{p}^2c^2 \ . \tag{3.1.1}$$

Equations that relate energy to momentum show up all over physics and have a special name: **dispersion relations**.[a]

[a] As a student I always found this name intimidating because it would keep showing up in very different and increasingly advanced corners of physics. I felt like I must be missing something deep, especially since the word *dispersion* did not seem to obviously fit. Historically, these relate wavelength to frequency. Recall that wave velocity is the product of wavelength and frequency—but wave velocity is purely a property in medium. Wavelength (or wave number) is directly related to momentum—think $\sin(kx)$—while frequency is directly related to energy—think $E = \hbar\omega$. These of these parameters are related to absorption (or decay) through complex analysis; these are the celebrated Kramers–Krönig relations. We mention all of this to encourage you to look these ideas up and see how they connect; they are one of the deep threads in physics.

> **Exercise 3.1.1.** It is obvious that (3.1.1) reproduces $E = mc^2$ when $\mathbf{p} = 0$. Show that the leading order correction in the small-$\mathbf{p}$ limit is simply the non-relativistic kinetic energy of the particle. HINT: start by identifying the *dimensionless* small parameter and Taylor expand.

In natural units we set $c = 1$. In this relation, m is the mass of a particle while E and $\mathbf{p}$ are the energy and three-momentum respectively. Let us write this with the kinematic quantities on the same side of the equation:

$$m^2 = E^2 - \mathbf{p}^2 \ . \tag{3.1.2}$$

A particle that satisfies this relation is said to be **on shell** or *physical*. Anticipating quantum mechanics, another way of saying this is that on shell particles are *observable states*. Personally, I think of 'on shell' states as being *nice* states that are relatable to my ordinary human experiences. This is in contrast to **off shell** particles, which states that are *not* on shell and are intrinsically quantum. Off shell states are not observable and do not make sense classically.

A scattering process is one where some number of *observed* initial state particles interact quantum mechanically and produce and *observed* number of final state particles. These initial and final states must each be on shell and conserve energy and momentum.

Rule 3.1.1 (Kinematics)**.** A **physical scattering process** is one with an on shell initial state and an on shell final state. This just means that each particle in the initial and final state are on shell. Furthermore, the total energy E and total three-momentum $\mathbf{p}$ are conserved through the process. In equations:

$$E_{\text{in}} = E_{\text{out}} \qquad \mathbf{p}_{\text{in}} = \mathbf{p}_{\text{out}} \qquad m_i^2 = E_i^2 - \mathbf{p_i}^2 \ , \tag{3.1.3}$$

where i labels each of the external (initial or final) particles. Technically, we should also specify that $E_i > 0$, but for us we can take this as an "obvious" fact.[b] Note that in this notation, E_{in} is the sum of the energies of all the initial state particles, and similarly for the other in/out quantities.

[b] From a group theory perspective, positive energy means that we are restricting to the *orthochronus* Lorentz group. This means that particles always move forward in time, no matter what reference frame we are in. The relation between energy positivity and direction in time should be clear from the time evolution operator, (2.3.1) with $\hat{H} \to E$.

Suppose you have a particle detector that measures the energy of a particle passing through it—this is called a *calorimeter*. If you also know the mass of the particle, then you can also unambiguously determine the magnitude of the momentum, $|\mathbf{p}|$. Alternatively, if you could separately measure the energy and the momentum of a particle, then you can unambiguously infer its mass. This is all obvious because you are using one Einstein relation to relate three variables—mass, energy, and momentum.

Exercise 3.1.2. Suppose you have a process where one particle decays into two other particles; these particles do not necessarily have the same masses, but suppose you know all of the masses. You know the energy and momentum of the initial particle, E_{in} and $\mathbf{p}_{\text{in}}$. You do not know, *a priori*, the energies or momenta of the final state particles. How many unknown scalar quantities are there? How many constraint equations are there? Hint: recall that $\mathbf{p}$ is a vector with three separate quantities. Argue that there is generically *no* solution to this system unless some parameters (e.g. the masses) are just right.

Exercise 3.1.3. Suppose you have a process where two particles come in, and two particles come out. The particles do not necessarily have the same mass, but you know all of the masses. If you know the energies of the initial particles, how many unknowns are there and how many constraint equations are there? Do you expect this system to have a solution? Note: for the purposes of this problem, assume that all energies are positive. This corresponds to satisfying the orthochronus constraint. If you are nervous about this, check that by increasing the energies of the initial particles you can make sure that the final state particles have positive energy.

3.2 Time Dilation and Length Contraction

Flip Comment

This section is a very brief review of the main points. For a more systematic derivation, please see my Physics 17 lecture notes.

My favorite refernece to appreciate the geometric structure of special relativity is the book *Very Special Relativity.*[1] It looks like a miniature coffee table book, but it is a perfect book for physics students who have completed their lower-level coursework.[a,b] Once you have mastered this, you can pick up your favorite general relativity textbook for a bit more of the mathematical formalism. Some suggestions: Hartle,[2] Schutz,[3] and Carroll.[4] I would be remiss not to also mention the beautiful and elegant tome known lovingly as MSW;[5] a book so beloved that it had its own 50th anniversary celebration.[6]

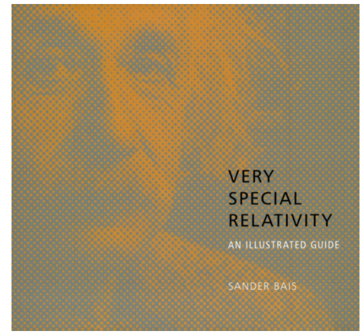


Figure 3.1: How to learn special relativity.

[a] If you can derive every result in the book then you are ready to take general relativity. You should be able to do this over a winter break.

[b] Do not confuse the title of this book with the Cohen–Glashow hypothesis, arXiv:hep-ph/0601236 , which is a rather different thing.

The key tenet of special relativity is that the speed of light is constant. You already know from (2.4) that this constant means that there is a natural conversion between space and time. Indeed, you should already be familiar with the two primary manifestations of this in mechanics:

- Time dilation: We measure time to pass more slowly for objects moving relative to our reference frame.
- Length contraction: We measure the distance along the direction of motion to be shorter for objects moving relative to our reference frame.

This is formalized with respect to the relative velocity[c] $\beta = v$ and the factor

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \ . \tag{3.2.1}$$

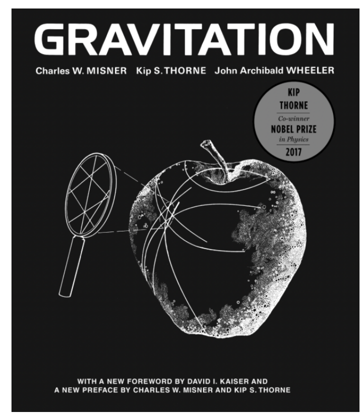


Figure 3.2: Published 50 years ago—right around when the Standard Model was established—MSW is still one of the most insightful places to learn and re-learn relativity.

> **Exercise 3.2.1.** What is the range of allowed values of β? What is the range of allowed values for γ? What is the allowed range of the product $\gamma\beta$?

We, as observers, define a stationary reference frame with coordinates t and x. For our purposes, let us assume a (1+1)-dimensional spacetime.[d] Suppose that there is another references frame that is moving at constant velocity relative to ours. Then an observer in that reference frame has a coordinate system t' and x'. Fig. 3.3 plots the axes of this coordinate system on our coordinate system.

[c] In SI units we would say $\beta = v/c$.

[d] This notation means one dimension of time and one dimension of space.

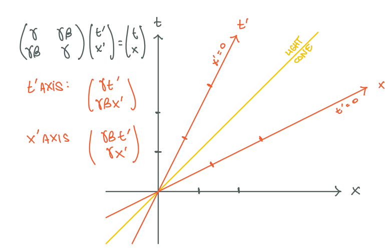


Figure 3.3: Coordinates of a boosted observer relative to our coordinates.

> **Exercise 3.2.2.** Compare Fig. 3.3 to the spatial x–y plane and the x'–y' plane where the primed coordinates are those of an observer rotated by angle θ relative to the unprimed observer.

[1] S. Bais. *Very Special Relativity: An Illustrated Guide*. Harvard University Press, 2007. ISBN: 9780674026117.

[2] J. B. Hartle. *Gravity: An introduction to Einstein's general relativity*. 2003. ISBN: 978-0-8053-8662-2.

[3] B. Schutz. *A First Course in General Relativity*. Cambridge University Press, 2009. ISBN: 9780521887052.

[4] Sean M. Carroll. *Spacetime and Geometry: An Introduction to General Relativity*. Cambridge University Press, July 2019. ISBN: 978-0-8053-8732-2, 978-1-108-48839-6, 978-1-108-77555-7. DOI: 10.1017/9781108770385.

[5] C.W. Misner et al. *Gravitation*. Princeton University Press, 2017. ISBN: 9781400889099.

[6] https://www.youtube.com/watch?v=a-4-IPBNV60

the coordinates are related by

$$t' = \gamma t - \gamma\beta x \qquad\qquad x' = \gamma x - \gamma\beta t \tag{3.2.2}$$
$$t = \gamma t' + \gamma\beta x' \qquad\qquad x = \gamma x' + \gamma\beta t' \ . \tag{3.2.3}$$

We realize that this is a linear system of equations,[e] so we can write this as a matrix multiplication:

[e] Recall that system of equations is linear if a *linear combination* of solutions is also a solutions.

$$\begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta \\ \gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} t' \\ x' \end{pmatrix} \ . \tag{3.2.4}$$

Exercise 3.2.3. Show that

$$\begin{pmatrix} t' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} \ . \tag{3.2.5}$$

You can do this either "read it off' (3.2.3) or apply the matrix inverse to both sides of (3.2.4). Show that the two matrices are, indeed, inverses of each other.

Exercise 3.2.4. Confirm that (3.2.4) maps x' and t' axes to the x and t plane as shown in Fig. 3.3.

3.3 Indexology of Special Relativity

3.3.1 Upper Index

The unification of space and time is manifest in the formalism of **four vectors**. At one level, this is a simple generalization of three-component vectors to four-component vectors. For example, an *event* is a position $\mathbf{x}$ and a time t and we can write it as a four-vector with an upper index:

$$x^\mu \equiv (x^0, x^1, x^2, x^3) = (t, x, y, z) \ . \tag{3.3.1}$$

There are several caveats that we should make at this point since they are not often stated out loud in textbooks:

- We use the usual physicist's abuse of notation where we identify an object x with a generic component, x^μ.
- It is conventional to index time with $\mu = 0$. Some old texts use the antiquated notation x^4. For this reason in theories of extra dimensions it is conventional to label the extra dimension as x^5.
- There is no such thing as a *position vector*.[a] This is because vector spaces have a special point, the origin $x^\mu = 0$. There is no such special point in spacetime. However, relative positions, $\Delta x^\mu = x^\mu_A - x^\mu_B$, are well defined vectors.

[a] I know a mathematician who was confused when the physics students in his calculus class kept talking about position vectors, as if the students were talking about unicorns. This may seem like nit picking, but it this insight is part of the underlying geometric picture built on fiber bundles. See arXiv:hep-th/0611201 for a pedagogical introduction that should be accessible to students of this class with a bit of work.

In particle physics we do not often deal with (relative) position four vectors because our states are typically momentum eigenstates. Just as time and space are unified by $c = 1$, so too are energy and momentum. These are combined into a four-momentum,

$$p^\mu \equiv (p^0, p^1, p^2, p^3) = (E, p^x, p^y, p^z) \ . \tag{3.3.2}$$

p^μ is the Fourier transform of x^μ. It should not surprise you that we work with momentum eigenstates: these are the states that are eigenstates of the Hamiltonian by virtue of having a well defined energy. The on shell condition (3.1.2) then gives a well defined three-momentum magnitude.

Like any other vector space, linear combinations of vectors are also vectors. Recall that a linear combination of vectors is a sum of vectors that each may have some rescaling by a real number. For two vectors, this is:

$$(\alpha\mathbf{v} + \beta\mathbf{w})^i = \alpha v^i + \beta w^i \ . \tag{3.3.3}$$

We read this relation as: the i^{th} component of the vector $(\alpha\mathbf{v} + \beta\mathbf{w})$ is equivalent to the number given by the sum of two rescaled vector components: α times v^i and β times w^i. The coefficients α and β are any real numbers. In this example we used the index i to show that this is true for *any* vector space, not just the space for four-vectors. We call the μ index[b] a **Lorentz index** because it indexes the four directions of spacetime.

[b] In our convention, lowercase letters from the middle of the Greek alphabet are Lorentz indices.

Example 3.3.1. For example, this means that conservation of energy and momentum in an interaction

$$\sum_{i=\text{incoming}} p^\mu_{(i)} - \sum_{i=\text{outgoing}} p^\mu_{(i)} = 0 \tag{3.3.4}$$

where $p^\mu_{(i)}$ is the μ component of the four-momentum of the i^{th} particle in an interaction. By convention, incoming particles have positive signs and outgoing particles have negative signs. Setting $\mu = 0$ in (3.3.4) tells us that the sum of all energies going into an interaction must match the sum of all energies coming out of the interaction. Similarly, setting $\mu = 1$ tells us that momentum in the x-direction is conserved.

Vectors with upper indices have equivalent names: column vectors, contravariant vectors, kets. To be very precise mathematically: the vector is the object $\mathbf{v}$, while v^μ is the μ^{th} component of the vector. Physicists like to blur the line here and will write a *generic component* of a vector v^μ to interchangeably mean the vector as a whole or the specific component.[7][c]

[c] What is implicit here is an understood set of basis vectors $\mathbf{e}_\mu$ such that $\mathbf{v} = v^\mu \mathbf{e}_\mu$. These basis vectors carry the abstract vector-ness of $\mathbf{v}$ while v^μ can be thought of as a set of numbers. Refer to my Physics 17 notes to see this exhaustively.

3.3.2 Lower Index

Upper index vectors also have lower-index counterparts. These also have many equivalent names: row vectors, covariant vectors, bras, one-forms.

[7] Roger Penrose proposes formalizing this shorthand in what is called *abstract index notation*. A good online discussion is this one: https://math.stackexchange.com/questions/455478/

There is a mathematical machine called a **metric** that can raise or lower indices. For now let us assume that lower indexed objects exist. In fact, often momentum is written with a lower index:

$$p_\mu = (E, p_x, p_y, p_z) \ . \tag{3.3.5}$$

As introduced, this as a separate object than the upper index four-momentum (3.3.2). However, the metric gives a concrete relation between the two:

$$p_\mu \equiv (E, -p^x, -p^y, -p^z) \qquad p^\mu \equiv (E, -p_x, -p_y, -p_z) \ . \tag{3.3.6}$$

To go from upper to lower Lorentz index, evidently there is simply a minus sign on the spatial ($\mu = 1, 2, 3$) components.

3.3.3 Summation Convention

We adopt **summation convention** where:

> **Rule 3.3.1** (Summation convention)**.** In an expression where an upper and a lower index are written with the same character, it is *understood* (left unwritten) that one should sum over these indices. We say that these indices are **contracted**.

The key example for us is the contraction of an upper-index four momentum and its lower-index counterpart:

$$p^\mu p_\mu = p_\mu p^\mu = E^2 - p_x^2 - p_y^2 - p_z^2 = E^2 - \mathbf{p}^2 \ . \tag{3.3.7}$$

Using this convention, the on shell condition (3.1.2) is[d]

$$p^2 \equiv p^\mu p_\mu = m^2 \ . \tag{3.3.8}$$

[d] Note that $p^2 = p_\mu p^\mu$ while $\mathbf{p}^2 = p_x^2 + p_y^2 + p_z^2$.

Contracted indices are effectively not indices: $p^2 = p^\mu p_\mu$ is a *number* and not a four-vector or other type of tensorial object.

Evidently the way we are writing our theory tells us that the unification of space and time (energy and momentum) is also a statement about the on shell condition for particles.

The following examples are completely general for any type of linear transformation from a vector space to itself.

> **Example 3.3.2.** A linear transformation ("matrix") $M^i{}_j$ is a *linear map* that takes vectors into vectors. For example, the transformation $M = M^i{}_j$ acts on a vector $\mathbf{v} = v^i$ as $M\mathbf{v}$ with components given by
>
> $$(M\mathbf{v})^i = \sum_k M^i{}_k v^k = M^i{}_1 v^1 + M^i{}_2 v^2 + \cdots \ . \tag{3.3.9}$$
>
> Observe that the matrix must have an upper index and a lower index. Summation convention tells us that the lower index allows the matrix to contract with the index of the vector it acts on. The upper index of the matrix corresponds is 'leftover' and is the free index of the resulting vector, $M\mathbf{v}$.

Example 3.3.3. In (3.3.9) the index k is contracted. We may rewrite this equation using summation convention as

$$(M\mathbf{v})^i = M^i{}_k v^k \ . \tag{3.3.10}$$

3.4 Lorentz Transformations

How do you relate vector quantities between different reference frames? A particle may have some four momentum[a] p. The *components* of p depend on the reference frame of the observer. We measure the components to be $p^\mu = (p^0, p^1, p^2, p^3)$. Another observer in another reference frame measures components $p'^\mu = (p'^0, p'^1, p'^2, p'^3)$. These components are related by a **Lorentz transformation** (3.2.4):

[a] We write $\boldsymbol{p}$ to mean the abstract four-vector without specifying its components. We skip the 'physics notation' of writing $\boldsymbol{p}^\mu$ for clarity.

$$p^\mu = \Lambda^\mu{}_\nu p'^\nu \ . \tag{3.4.1}$$

If we align the spatial parts of each observer's coordinate system so that their relative motion is in the z direction, then the components of the Lorentz transformation are

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & & & \beta\gamma \\ & 1 & & \\ & & 1 & \\ \beta\gamma & & & \gamma \end{pmatrix} \ . \tag{3.4.2}$$

This is simply what we showed in (3.2.3), extended to the 4×4 case. You should think about this as an analog of the rotation matrix in Euclidean space.[b] Here β is the relative velocity between the two frames.

[b] The relation is as follows. Rotations are the isometries of Euclidean space: they are the transformations that preserve the dot product. Lorentz transformations are the isometries of Minkowski space: they are the transformations that preserve the metric. In both cases, the isometries are the allowed transformations that preserve the spacetime structure as encoded by the metric.

Exercise 3.4.1. Check that (3.4.2) is a transformation from frame $\mathcal{O}'$ to frame $\mathcal{O}$, where $\mathcal{O}'$ is moving in the $+\hat{z}$ direction relative to the $\mathcal{O}$. Show that the inverse of Λ simply swaps the sign of β.

Example 3.4.1. A nice way to see the relation between Lorentz transformations and boosts is to note that we may write the boost parameter as **rapidity** η defined through the identification:

$$\Lambda[\eta] = \begin{pmatrix} \cosh\eta & \sinh\eta \\ \sinh\eta & \cosh\eta \end{pmatrix} = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} \ . \tag{3.4.3}$$

Compare this to a rotation matrix,

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \ . \tag{3.4.4}$$

Observe the key relations:

$$\cosh^2\eta - \sinh^2\eta = 1 \tag{3.4.5}$$
$$\cos^2\theta + \sin^2\theta = 1 \ . \tag{3.4.6}$$

Here we appreciate the critical minus sign that is the difference between space and time.[c]

What about the transformation of lower index objects, like p_μ? We can motivate based on what we know physically. But first, it is useful to review the case of simple rotations:

[c] The connection between Lorentz transformations and rotations is further explicit in the so called Pauli-metric formalism where the time-component of a four-vector is *imaginary*. In that formalism the metric is proportional to the identity and both Lorentz transformations and rotations take the form (3.4.4).

Example 3.4.2 (Transformation of row vectors)**.** Consider a Euclidean three-vector $\mathbf{v} = (v^x, v^y, v^z)^\mathsf{T}$ and a Euclidean row vector, $\underset{\sim}{\mathbf{w}} = (w_x, w_y, w_z)$. You know that the contraction these two objects is a number that is *invariant*.

$$\underset{\sim}{\mathbf{w}}\mathbf{v} = w_x v^x + w_y v^y + w_z v^z. \tag{3.4.7}$$

In fact, if $\underset{\sim}{\mathbf{w}} = \mathbf{w}^\mathsf{T}$, then $\underset{\sim}{\mathbf{w}}\mathbf{v}$ is simply the inner product $|\mathbf{w}||\mathbf{v}|\cos\theta$. At any rate, the contraction $w_i v^i$ carries no indices and is a pure scalar quantity. That means it does not transform under rotations. However, we know that $\mathbf{v}$ *does transform under rotations*,

$$\mathbf{v} \to R(\theta)\mathbf{v} \ , \tag{3.4.8}$$

with $R(\theta)$ given in (3.4.4). That means $\underset{\sim}{\mathbf{w}}$ should also transform in a way to compensate the $\mathbf{v}$ transformation and keep $\underset{\sim}{\mathbf{w}}\mathbf{v} = w_i v^i$ constant. Call this transformation $S(\theta)$. By the rules of matrix multiplication, $S(\theta)$ must act from the right:

$$\underset{\sim}{\mathbf{w}} \to \underset{\sim}{\mathbf{w}} S(\theta) \tag{3.4.9}$$

but from the index perspective the order does not matter:

$$w_i \to w_i S(\theta)^i{}_j = S(\theta)^i{}_j w_i \ . \tag{3.4.10}$$

This is because each term in the sums on the right-hand sides of (3.4.10) are just multiplication of two numbers. *Make sure you understand this.*

The statement that $\underset{\sim}{\mathbf{w}}\mathbf{v}$ is constant under rotations is

$$\underset{\sim}{\mathbf{w}}\mathbf{v} \to \underset{\sim}{\mathbf{w}} S(\theta)R(\theta)\mathbf{v} = R(\theta)^i{}_\ell S(\theta)^k{}_i w_k v^\ell \tag{3.4.11}$$

so that we requre:

$$R(\theta)^i{}_\ell S(\theta)^k{}_i = S(\theta)^k{}_i R(\theta)^i{}_\ell = \delta^k{}_l \ , \tag{3.4.12}$$

or in other words, as matrices, $S(\theta)R(\theta) = 1$. Note that while it turns out to also true that $R(\theta)S(\theta) = 1$, (3.4.12) is *only* telling us that $S(\theta)R(\theta) = 1$ because when we write this index contraction was a matrix multiplication, the order matters: we have to make sure contracted indices are consecutive.

We thus conclude that

$$S(\theta) = R(\theta)^{-1} = R(-\theta) = R(\theta)^\mathsf{T} \ . \tag{3.4.13}$$

So we demonstrate something very important: row vectors like $\underset{\sim}{\mathbf{w}}$ 'rotate oppositely' from column vectors like $\mathbf{v}$.

The lesson from the above example is that lower indexed objects transform with the *inverse transformation* as upper indexed objects. Thus if

$p^\mu \to \Lambda^\mu{}_\nu p^\nu$, we must have

$$p_\mu \to (\Lambda^{-1})^\nu{}_\mu p^\nu \ . \tag{3.4.14}$$

Check to make sure that you agree with the position of the indices. It is $(\Lambda^{-1})^\nu{}_\mu$, not $(\Lambda^{-1})_\mu{}^\nu$. Refer back to Example 3.4.2 if that is not clear.

3.5 Example: Muon Decay

Cosmic rays can produce muons when they hit the upper atmosphere about 10 kilometers above the surface of the Earth.[a] These muons are highly relativistic, with a velocity of $\beta = 0.9999$. See Fig. 3.4. We know from laboratory experiments that the lifetime of a muon *at rest* is 2 microseconds. Based on the simple estimate $d = c\tau \approx 600$ meters, we would not expect any muons to reach the surface of the Earth. However, not only do large cosmic ray telescopes have dedicated muon detectors, but you can make your own citizen science muon detector[b]. What gives? Here are the facts:

[a] This sub-section is copied from my Physics 17 (2023) notes and is adapted from an example in Griffiths.

[b] https://muonpi.org

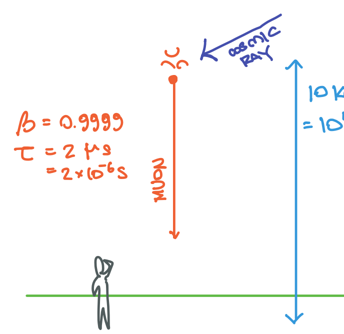


Figure 3.4: Very technical sketch of a muon produced in the upper atmosphere heading towards earth.

- Both the observer on Earth and the muon agree that their relative velocity is $|\beta| = 0.9999$.
- The muon's lifetime is known in the muon's rest frame.
- The distance from the surface of the Earth to the upper atmosphere is known in the Earth's rest frame.

The Lorentz transformation between the muon frame and the observe frame mix up space and time separations. We can only express the distance or time that the muon travels by calculating in the same reference frame.

First let us consider calculating everything in the Earth's reference frame. This is shown on the left of Fig. 3.5. This means we have to take the muon's lifetime in the muon's rest frame and convert it into a lifetime in the Earth's frame. In the muon's frame the lifetime is simply

$$\Delta x'^\mu = \begin{pmatrix} \tau \\ 0 \end{pmatrix} \qquad\qquad \Delta x^\mu = \begin{pmatrix} \gamma\tau \\ \gamma\beta\tau \end{pmatrix} \ , \tag{3.5.1}$$

where we have noticed that the muon lifetime is a displacement along the t' axis used our results in (3.2.4). We find that the lifetime in the Earth's frame is actually larger than the lifetime in the muon rest frame: $t = \gamma\tau$. There is also a spatial component, but this is no surprise: in the Earth frame the muon is moving, so when the muon decays it is in a different position.

How much is the muon's lifetime *time dilated*? We plug β into the expression for $\gamma = (1-\beta^2)^{-1/2}$. We find[8] that to one significant figure, $\gamma = 100$. This means that the time for the muon decay in the Earth's frame is $2 \times 10^{-}4$ seconds, which means that it travels approximately $d = 60$ km,

[8] There's a cute trick here: $\beta^2 = (1-\epsilon)^2 \approx 1 - 2\epsilon$ by Taylor expanding in the small number $\epsilon = 10^{-4}$.

which is larger than the distance from the upper atmosphere to the surface. As a sanity check: this is exactly the value in the spatial component of Δx^μ.

Great: so one story is that relativistic muons have lifetimes that are much larger than their at-rest lifetime. This means that to the observer on earth, the muon simply lives longer than we would expect from measurements of muons at rest. There are, however, (at least) two sides to every good story. What does the muon see?

In the muon's rest frame, the muon *knows* that it goes *kaput* in 2 microseconds. It sees the surface of the Earth approaching it with velocity $\beta = 0.9999$. By now you can guess that must change: the measurement in the Earth's frame—that the height of the upper atmosphere is 10 km—must be different in the muon's frame. And in fact, the muon must measure the distance of the rapidly approaching Earth to be much smaller than 10 km. How does this *length contraction* work?

We show this on the right side of Fig. 3.5. Note that now we have a measurement in the Earth frame (a vertical line denoting a fixed distance) that we want to project onto the muon frame (red axes). In the Earth frame, we denote the distance by two unit ticks in the spatial direction. In the muon frame, this line intersects the x' axis with *less* than two ticks.

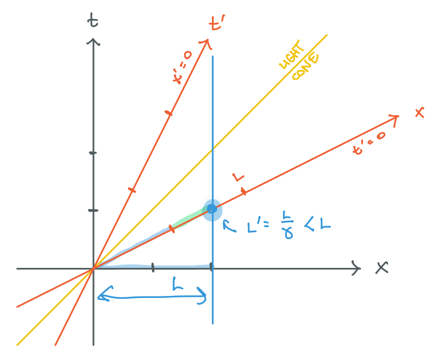


Figure 3.5: The distance from the surface of the Earth to the upper atmosphere is length contracted in the muon's frame.

> **Exercise 3.5.1.** Using the Lorentz transformation laws, show that the distance from the muon to the surface of the Earth at the moment of the muon's creation is $L' = L/\gamma$ where $L = 10$ km is the distance in the Earth frame.

3.6 The metric tensor

The **inner product** (or dot product) is a machine that takes two vectors and outputs a number. It is manifested by a tensor called the metric, which has two lower indices:

$$\langle p, q\rangle = p \cdot q = g_{\mu\nu} p^\mu q^\mu \ . \tag{3.6.1}$$

In special relativity the metric is conventionally written as $\eta_{\mu\nu}$ and has a simple form in Cartesian coordinates:

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} = \text{diag}(1, -1, -1, -1) \ . \tag{3.6.2}$$

Some physics tribes[a] use a different convention, $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$. The choice of whether the spatial or temporal pieces pick up the minus sign is a convention—while the intermediate steps of any calculation differ by these annoying signs, any physical result is independent of the convention.[b]

[a] Particle physicists use the 'West coast' or 'mostly minus' notation. It is usually relativists and formal theorists who use the East coast/-mostly plus convention. A third convention, the 'Pauli convention' uses a metric proportional to the identity but with the timelike component imaginary $x^4 = ix^0$. In that notation, boosts look like complex rotations. See Appendix F of *Diagrammatica*[9] by Veltman for a discussion.

[b] I am hopelessly entrenched in the mostly minus tribe. If I were being honest, I think the mostly-plus metric is probably easier to start learn if you were starting from scratch. But I

[9] Veltman, *Diagrammatica: The Path to Feynman rules*

The metric has an inverse, $g^{\mu\nu}$ or $\eta^{\mu\nu}$ in special relativity. It has two upper indices and satisfies

$$g_{\mu\nu} g^{\mu\nu} = g^{\mu\nu} g_{\mu\nu} = \delta^\mu_\nu \ . \tag{3.6.3}$$

We do not bother writing $(g^{-1})^{\mu\nu}$ because the height of the indices indicates precisely whether you are using the metric or inverse metric. In special relativity, the components of $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$ are identical. Also observe that the Kronecker-δ has no distinction between first and second indices:

$$\delta^\mu_\nu = \begin{cases} 1 & \text{if } \mu = \nu \\ 0 & \text{otherwise} \end{cases} \ . \tag{3.6.4}$$

The Kronecker-δ represents the components of the unit matrix, $\mathbb{1}^\mu{}_\nu = \delta^\mu_\nu$.

3.7 Example: What does someone else measure?

In special relativity there is an object called the four-velocity.[a] In our rest frame, our four velocity is

[a] This subsection borrows from my Physics 17 (2023) notes.

$$v^\mu = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ . \tag{3.7.1}$$

This literally means that when we are at rest, we are moving one second per second in the time direction. Objects moving relative to us have four velocities that are Lorentz transformations of the v^μ above.

We notice that we can write the energy of a particle in a way that uses the inner product:

$$\langle v, p \rangle = p \cdot v = g_{\mu\nu} p^\mu v^\nu \ . \tag{3.7.2}$$

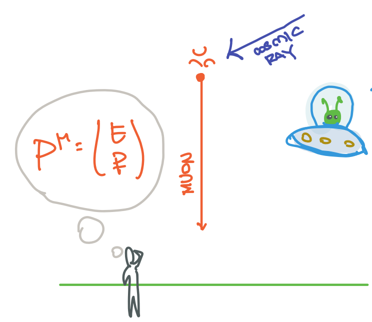


Figure 3.6: You measure the four-momentum of a particle. What is the energy that an alien moving at some velocity β relative to you measures?

Of course, all this does is pick out the p^0 component, as we knew it had to. However, unlike writing p^0, the inner product $p \cdot v$ has no indices. It is a pure number and so it does not transform under Lorentz transformations.

At this point you wonder if we are simply reciting random facts that we have developed. Consider the following scenario illustrated in Fig. 3.6. While you are measuring the particle energy $p^0 = E$, you notice an alien traveling relativistically with velocity β relative to you. The alien has sophisticated equipment to measure the particle energy, and you know that the alien measures a different energy E'. How can you determine what the alien measures?

One way to do this is to calculate the full Lorentz transformation between your frame and the alien frame. That is tedious. It turns out that we can use the four-velocity as a useful trick. All objects with mass have a four-velocity equal to (3.7.1) in their rest frame. This means that the alien measures the particle to have energy $v_{\text{alien}} \cdot p_{\text{alien}}$, where the subscript 'alien' means that these are all calculated in the alien's frame.

We now remember that $v_{\text{alien}} \cdot p_{\text{alien}}$ is a number. It does not matter what frame we calculate it in. Thus it is equivalent to the same dot product measured in our frame:

$$E' = v_{\text{alien}} \cdot p_{\text{alien}} = v \cdot p \ , \tag{3.7.3}$$

where the right-hand side is the alien four-velocity and the particle four-momentum as measured in our frame. The alien four-velocity is simply a Lorentz transformation of (3.7.1). More practically, it is something that you can measure in your own frame.

> **Exercise 3.7.1.** Rephrase everything in this example in terms of the inner product in Minkowski space. Bonus if you use the word 'projection.'

3.8 Tensors: the meaning of indices

The lightning-quick review of special relativity here is an example of tensor analysis in physics.[a] Tensors show up *all over the place* in physics. Even in your lower division physics courses: did you ever wonder why it is called the moment of inertia *tensor* and not the moment of inertia *matrix*? Yes, there is a difference.[10][b] For our purposes, we can think of a **tensor** as an object that has an ordered number of indices. The indices each have a height, either upper or lower.

[a] Where can you learn more? I recommend my Physics 17 lecture notes. Want a more advanced version? Check out my Physics 231 lecture notes.

[b] I am frustrated that few textbooks take a moment to explain the significance this difference.

The order of the indices matters; if you want it takes precedence over the height of the index:

$$T^{ij}{}_{k}{}^{\ell} \neq T_{k}{}^{ij\ell} \ . \tag{3.8.1}$$

If you have a metric, this is more clear since you can lower (or raise) all the indices, so that

$$T^{ij}{}_{k}{}^{\ell} = T^{ijm\ell} g_{mk} \tag{3.8.2}$$

$$T_{k}{}^{ij\ell} = T^{mij\ell} g_{mk} \tag{3.8.3}$$

and you can see that $T^{mij\ell} \neq T^{ijm\ell}$.

There are some equivalent ways of thinking about tensors. The most formally correct way is to think about them as *multilinear maps* between vector spaces (or tensor products[c] thereof). Perhaps more practically, a tensor is an object that encodes information that transforms according to the height of the indices. Generalizing our rules from Section 3.4:

[c] Given a vector space V, a tensor product $V \otimes V$ is two copies of the vector space.

> **Rule 3.8.1** (Tensor transformation)**.** If T is a tensor with upper indices $i_1, \cdots, i_N$ and lower indices $j_1, \cdots, j_M$, then under a symmetry transformation $R(\theta)$, T

[10] See e.g. https://hepweb.ucsd.edu/ph110b/110b_notes/node24.html

transforms as

$$T^{\cdots}_{\cdots} \to R(\theta)^{i_1}{}_{k_1} \cdots R(\theta)^{i_N}{}_{k_N} (R(\theta)^{-1})^{\ell_1}{}_{j_1} \cdots (R(\theta)^{-1})^{\ell_M}{}_{j_M} T^{\cdots}_{\cdots} \ , \tag{3.8.4}$$

where we have written $\cdots$ on the left hand side to mean 'some arrangement of upper i and lower j indices,' while on the right the $\cdots$ mean 'some arrangement of upper k and lower ℓ indices.'

In other words:

- For each upper index, multiply by a factor of $R(\theta)$ and contract with the lower index of $R(\theta)$.
- For each lower index, multiply by a factor of $R(\theta)^{-1}$ and contract with the upper index of $R(\theta)^{-1}$. Note that this is contraction with the *first* index of $R(\theta)^{-1}$.

To apply the rule for Lorentz transformations, simply replace R with Λ and convert the indices into μs and νs.

We ultimately care about quantities that are *invariant* under Lorentz transformations. These are objects that everyone agrees upon, no matter what their reference frame. If you were to build a theory of nature, you would want the physical laws from that theory to be invariant with respect to reference frame—so the objects you would build that theory out of are naturally invariants.

3.9 Isometry

How does one know that rotations are important symmetries in Euclidean space? This certainly comes from our own experience with physics—the laws of physics are rotationally invariant. In fact, we learn very quickly in our physics training to use rotational invariance to simplify problems. What about in special relativity? Historically, the Michelson–Morley non-observation[a] of aether—that is, the observation that the speed of light is constant—led us to realize that requirement that physics cannot depend on reference frame leads to a different class of symmetries. What characterizes these symmetries?

[a] Perhaps the most significant null result in science.

From a top-down perspective, rotations and their generalizations[b] are *isometries*. An **isometry** is a symmetry of *the form of the metric*. Specifically, it is a transformation—enacted by a matrix acting on vectors, for example—for which the *components* of the metric do not change. With our conventions, special relativity requires that the components of the metric are $\eta_{\mu\nu} = \text{diag}(+,-,-,-)$; this means that the metric takes this form in *any* reference frame.

[b] Such as Lorentz transformations in special relativity.

Ah! But we know from Rule 3.8.1 that the metric is an object with two lower indices. That means that it has a prescribed transformation rule. If $v^\mu \ toR^\mu{}_\alpha v^\alpha$ for some transformation R, then the metric transforms as

$$g_{\mu\nu} \to R^\alpha{}_\mu R^\beta{}_\nu g_{\alpha\beta} \stackrel{?}{=} g_{\mu\nu} \ . \tag{3.9.1}$$

For R to be an isometry, the last expression (marked '?') must be true:

$$R^{\alpha}{}_{\mu} R^{\beta}{}_{\nu} g_{\alpha\beta} = g_{\mu\nu} \ . \tag{3.9.2}$$

We call the isometries of spacetime Lorentz transformations and typically write $R = \Lambda$. The isometries of Euclidean space are the rotations, which we usually keep writing as R. In this course we also meet the complex versions of rotations,[c] which we write $R = U$.

[c] This should familiar from quantum mechanics. Symmetries are described by a mathematical class called a group. Rotations in N dimensions are the group $\mathrm{SO}(N)$ for "special orthogonal" $N \times N$ matrices. Special means unit determinant and orthogonal means $R^{\mathsf{T}}R = \mathbb{1}$. The complex version is the group $\mathrm{SU}(N)$, for special *unitary* matrices. Unitary means $U^{\dagger}U = 1$.

Exercise 3.9.1 (Definition of Lorentz Tranform)**.** The relation (3.9.2) in special relativity is

$$\Lambda^{\alpha}{}_{\mu} \Lambda^{\beta}{}_{\nu} \eta_{\alpha\beta} = \eta_{\mu\nu} \ . \tag{3.9.3}$$

This is of course just a change of symbols. When this is introduced in some textbooks, they often write it in matrix form rather than index form:

$$\Lambda^{\mathsf{T}} \eta \Lambda = \eta \ . \tag{3.9.4}$$

Show that these two forms are indeed equivalent. HINT: the matrix formalism suppresses indices because it assumes that consecutive indices are contracted.

Example 3.9.1. In Exercise 3.9.1, you were faced with something that is rarely explained: if a matrix has indices $M^{i}{}_{j}$, what is the index structure of the transpose, M^{T}? More generally, given a linear transformation M, what is the index structure of the *adjoint*, $M^{\dagger}$?

Exercise 3.9.1 gives you a hint for this. The correct identification is

$$M^{\dagger\, i}{}_{j} = M^{j}{}_{i} \ , \tag{3.9.5}$$

but this does not seem to make sense because the heights of the indices do not match. In fact, what we really mean is

$$M^{\dagger\, i}{}_{j} = g_{\ell j} g^{ki} A^{\ell}{}_{k} \ . \tag{3.9.6}$$

The way to see this is to go from the root definition of the adjoint with respect to the inner product. Recall that the inner product is a (bi-)linear function that takes two vectors and returns a number. It is defined with respect to the metric as

$$\langle \mathbf{v}, \mathbf{w} \rangle = g_{ij} v^{i} w^{j} \ . \tag{3.9.7}$$

Then the adjoint $M^{\dagger}$ is defined relative to M to be the matrix that satisfies

$$\langle \mathbf{v}, M^{\dagger}\mathbf{w} \rangle = \langle M\mathbf{v}, \mathbf{w} \rangle \tag{3.9.8}$$

for every choice of $\mathbf{v}$ and $\mathbf{w}$. Writing out this expression with explicit indices gives

$$g_{ab} v^{a} M^{\dagger\, b}{}_{c} w^{c} = g_{de} M^{d}{}_{f} v^{f} w^{e} \ . \tag{3.9.9}$$

Observe that all indices are contracted—they are all dummy indices and can be relabeled for convenience. Let us do this for the right-hand side so that the indices on $\mathbf{v}$ and $\mathbf{w}$ match those on the left-hand side:

$$g_{ab} v^{a} M^{\dagger\, b}{}_{c} w^{c} = g_{dc} M^{d}{}_{a} v^{a} w^{c} \ . \tag{3.9.10}$$

Usually one avoids repeating the use of indices since this can lead to ambiguities about which indices are contracting; here there is no such issue because each side of the equal sign is evaluated independently. *However*, we make a clever observation: the equality holds for *any* vectors $\mathbf{v}$ and $\mathbf{w}$. This means we can choose basis vectors where only the a^{th} component of $\mathbf{v}$ and the c^{th} component of $\mathbf{w}$ are non-zero—for any specific choice of $a = \hat{a}$ and $c = \hat{c}$. In that case, the sum over these dummy variables collapses to those specific choices of hatted index:

$$g_{\hat{a}b} v^{\hat{a}} {M^{\dagger}}^{b}{}_{\hat{c}} w^{\hat{c}} = g_{d\hat{c}} M^{d}{}_{\hat{a}} v^{\hat{a}} w^{\hat{c}} \qquad \text{no sum over hatted indices} \tag{3.9.11}$$

$$g_{\hat{a}b} \cancel{v^{\hat{a}}} {M^{\dagger}}^{b}{}_{\hat{c}} \cancel{w^{\hat{c}}} = g_{d\hat{c}} M^{d}{}_{\hat{a}} \cancel{v^{\hat{a}}} \cancel{w^{\hat{c}}} \qquad \text{no sum over hatted indices} \tag{3.9.12}$$

$$g_{\hat{a}b} {M^{\dagger}}^{b}{}_{\hat{c}} = g_{d\hat{c}} M^{d}{}_{\hat{a}} \qquad . \tag{3.9.13}$$

Because there was no sum over hatted indices, we can cancel out the common factor of $v^{\hat{a}} w^{\hat{c}}$ on both sides. The last line gives (3.9.6) upon contracting with the inverse metric $g^{\hat{a}e}$ on both sides and then re-naming the indices appropriately.[11]

Example 3.9.2. Let us also clarify the phrase *the components of the metric are unchanged* as related to isometries. This is different from saying *the inner product is preserved*. For any transformation—not necessarily an isomsetry—$\mathbf{v} \to A\mathbf{v}$, the inner product is preserved *if you also allow the metric to transform*:

$$\langle \mathbf{v}, \mathbf{w} \rangle = g_{ij} v^i w^j \to {A^{-1}}^{k}{}_{i} \, {A^{-1}}^{\ell}{}_{j} g_{ij} \, A^{i}{}_{m} v^m \, A^{j}{}_{n} v^j \tag{3.9.14}$$

$$= (A^{-1}A)^{k}{}_{m} (A^{-1}A)^{\ell}{}_{n} g_{k\ell} v^m w^n \tag{3.9.15}$$

$$= g_{k\ell} v^k w^\ell = \langle \mathbf{v}, \mathbf{w} \rangle \ . \tag{3.9.16}$$

The point though, is that we do *not* let the metric transform. A transformation where the metric transforms cannot be interpreted as a chance of reference frame since all reference frames are supposed to see the same components of the metric.

This is obvious when considering transformations like $A = \text{diag}(3, 3)$. This stretches the length of vectors by a factor of 3. This is only meaningful as a physical transformation if we do not simultaneously transform the metric.

Example 3.9.3. It is clear that rotations (Lorentz transformations) are a leave the metric unchanged. What about translations? We know that translation invariance is a symmetry of spacetime—the homogeneity of space leads to momentum conservation. Why do we not have a matrix R for translations?

This is related to the point in Section 3.3.1 that there is no such thing as a 'position vector' and the definition of *linear* in (3.C.1). There is simply no such thing as translation in a vector space. In a vector space, the origin means something: it is the null (zero) vector. Translations do not make sense because they say that a different element is now the null vector. This makes no sense because the null vector is the unique element for which $\mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{v}$ for any $\mathbf{v}$.

Of course, translations *are* a key part of understanding physics. Vector spaces are just the wrong structure to describe them. The vector spaces that we deal with for spacetime are **tangent planes** to spacetime. A particle's velocity $\mathbf{v}$ is a

[11] See also https://threadreaderapp.com/thread/1702863061047263428

vector of a tangent plane at a point p on spacetime. We say that

$$v \in \mathrm{T}_p M \ , \tag{3.9.17}$$

which means the tangent plane at point p on the spacetime (manifold) M. At a different point, p', the particle's velocity is part of a different tangent plane, $\mathrm{T}_{p'}M$. Imagine the spaces of tangent vectors at some point on the equator versus at the north pole. Both of these are two dimensional vector spaces embedded in tree-dimensional ambient space. But these two dimensional planes are quite different. You cannot place a generic vector from one space onto the other space. [**Flip**: Insert picture]

All this is to say that the mathematical structure for including spacetime translations is different from the structure of a vector spaces. In fact, it is the collection of each vector space over each point in the spacetime. This collection is called $\mathrm{T}M$, the **tangent bundle** over the spacetime M.

By way of analogy: the metric is implicitly present in Euclidean space, but it is so trivial that we do not even define it when we first learn about vectors. In the same way, the tangent bundle is a structure that is implicitly present in special relativity but is so trivial—every tangent space looks like every other tangent space—that we never bring it up. However, in general relativity, the tangent bundle comes to life. In the example of the tangent spaces at the equator versus the north pole: because the underlying space has *curvature*, the ways in which tangent spaces are related to one another is no longer trivial. The curvature of spacetime—gravity—is mathematically understood to be a change in the relationships of tangent spaces. Because tangent spaces are where velocities live, this picture gives you an idea of how velocities may change (accelerate) in the presence of gravity.

To push this further: you also appreciate that there are other forces in nature—say, electromagnetism—that can accelerate test particles. Can these also be understood as some kind of tangent bundle? Yes! This is the mathematical structure of *gauge theory* and is the framework for what we call the fundamental forces. In this way, curvature in 'internal spaces' (not spacetime) is the origin of the fundamental forces in particle physics. The mathematical structure is precisely the same as that of gravity. In general, both gravity and the fundamental forces of particle physics are gauge theories; all of these forces are understood as tangent bundles over spacetime.

For a good starting point for all of this, I recommend the lecture notes by Collinucci and Wijns: Andres Collinucci and Alexander Wijns. "Topology of Fibre bundles and Global Aspects of Gauge Theories". In: *2nd Modave Summer School in Theoretical Physics*. Nov. 2006. arXiv: hep-th/0611201.

3.A Relativity by Thought Experiment

You can derive special relativity from the assumption that the speed of light is constant in any reference frame and then doing so called *gedanken-experiments*.[a] We continue our convention of using natural units where $c = 1$, though it should be obvious that using any other units just throws in factors of c all over the place.

[a] 'Thought experiments' in German.

Exercise 3.A.1. Rewrite all the equations in this appendix with the appropriate factors of c. Stop when it becomes obvious how to do this.

I first saw this done in Chapter 15 of *The Feynman Lectures on Physics*[12][b] I refer you to that resource for a systematic derivation from the *gedanken* approach. In this appendix, follow the the general idea and focus on a few subtle points that are not often explained carefully in the standard literature.[c]

[b] *The Feynman Lectures on Physics* are beloved gems of freshman-level physics insight—but the consensus is largely that they are a bit too non-sequitur in style for physics students. In fact, you come to deeply enjoy the lectures only *after* you already understand most of the material—then you can appreciate the little brilliant twists that Feynman makes compared to the usual pedagogy.

[c] I thank Matthew Lugatiman and Adam Green for talking through some of these subtleties with me.

Just as the word *gedankenexperiment* tells you where special relativity was developed, so too does the standard setting of the *gedankenexperiment*: start by imagining a train moving with some velocity $v = \beta$ in the x direction relative to our coordinate system. We are *outside* the train. We are observers, $\mathcal{O}$ and our coordinate system is (t, x). You can also imagine another observer, Oppie,[d] who is *in* the train and whose coordinate system carries primes: (t', x'). We say that Oppie is a *comoving* observer with the train. We align our coordinate systems so that the origins coincide:

[d] This is *not* a reference to Oppenheimer's nickname or a misspelling of Opie, Ron Howard's most famous character. Instead, it's a name that is reasonably close to $\mathcal{O}'$, "oh-prime."

$$(t = 0, x = 0) = (t' = 0, x' = 0) \ . \tag{3.A.1}$$

The above equation should be treated to mean *the point that we call* $(0, 0)$ *coincides with the point that Oppie calls* $(0, 0)$. In this way, a generic point (t, x) is really a *separation* between that point and the origin.[e]

[e] Recall our caveat in Chapter 3.3.1: positions are not vectors, but differences in positions are vectors.

3.A.1 Time dilation

First imagine that Oppie has a little gizmo that first emits a photon towards a mirror, then detects the reflected photon. See Fig. 3.7 The height of the gizmo is $D = \ell$, the height of the train. It is critical that the gizmo is aligned so that the photon moves *perpendicular* to the direction of motion. We write ℓ for *length* and to avoid ambiguities with [covariant] derivatives, D mesons, and so forth.

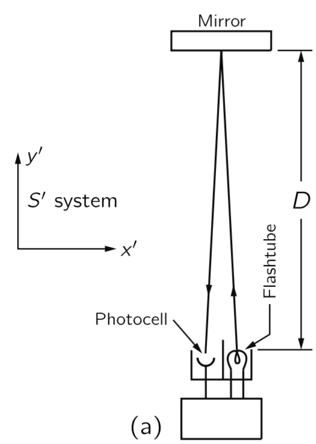


Figure 3.7: From *The Feynman Lectures on Physics*, Chapter 15.

Like us, Oppie understands that the speed of light is $c = 1$ and that this is true in any reference frame. At the origin ($t' = 0$), Oppie turns ont he gizmo and measures the time t' that it takes for the photon to traverse the distance ℓ and come back.

$$t' = 2\ell/c = 2\ell \ . \tag{3.A.2}$$

What do we see? Like Oppie, we see the device turn on at the origin ($t = 0$) and we see that the photon hits the roof of the train car and bounces back down. However, unlike Oppie, we also see the entire system move in direction of the train's motion. The trajectory is shown in Fig. 3.8. Because the motion of the train is perpendicular to the up-and-down direction, we assume that the height of the train is unchanged by relativistic effects: we measure this height to be ℓ just as Oppie does. We then measure that the photon takes times t to hit the top of the train and return. This time is evenly split[f] between the upward-going time and downward-going time.

[f] By symmetry.

[12] https://www.feynmanlectures.caltech.edu/I_15.html

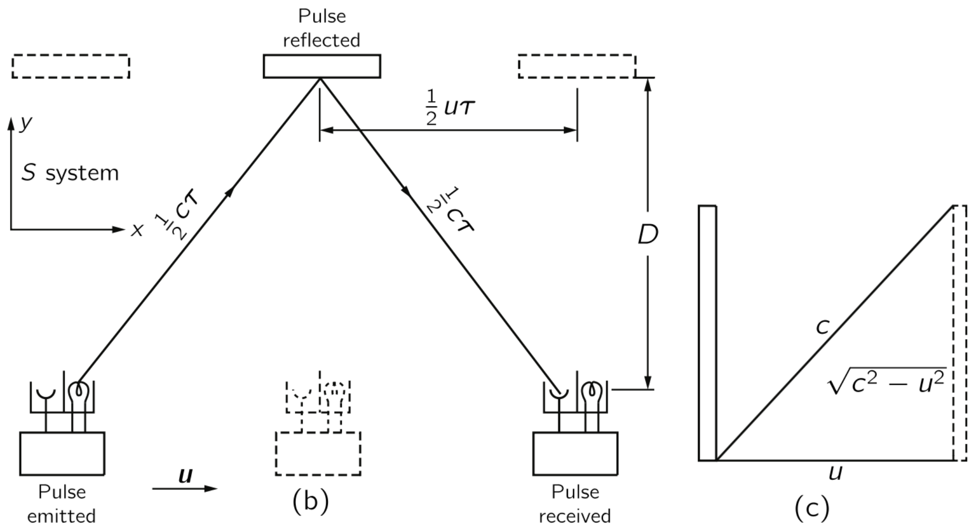


Figure 3.8: Same as Fig. 3.7, but as seen from observers who are not on the train. From *The Feynman Lectures on Physics*, Chapter 15.

During this time, the train has moved along the x direction by an amount βt. From trigonometry, we thus have:

$$\left(\frac{t}{2}\right)^2 = \left(\frac{\beta t}{2}\right)^2 + \ell^2 \tag{3.A.3}$$

Plugging in Oppie's observation that that relates the height of the train to the time Oppie measures, (3.A.2), we have

$$(t')^2 = (1-\beta^2)t^2 = \frac{t}{\gamma} \ . \tag{3.A.4}$$

That is to say, the time that we measure is related to the time that Oppie measures by

$$t = \gamma t' \ . \tag{3.A.5}$$

Because $\gamma > 1$, we the time that we see is *dilated* relative to what Oppie experiences. Because the gizmo is essentially an idealized clock, we say that observers see the events on the moving train moving slower than they would expect by a factor γ.

3.A.2 Length contraction

Having learned something about the passage of time, we can now figure out what motion does to the relative perception of space. To do this, let us do a second experiment where we turn the gizmo in Fig. 3.7 onto its side. Place the gizmo at the back of a train car and arrange it so that the photon first moves *along the direction of the train's motion*, then reflects off of the mirror to go *against the direction of the train for its return motion*. Let us now set the length of the gizmo to be the length of the train car: Ollie observes[g] this to be ℓ_0 while we observe it to be ℓ.

[g] We write ℓ_0 rather than ℓ' because Ollie is in the rest frame of the train. ℓ_0 is called a *proper length*, it is the length of an object to an observer for whom the object is not moving.

In pre-special relativity Galilean physics we would expect $\ell_0 = \ell$. However, the fact that $c = 1$ is constant means that we cannot make this assumption. Points on a spacetime diagram are called *events* because they are a place and a time. Figure 3.9 shows the experiment. Let the origin be the back of the train at time zero—where both we and Oppie synchronize our clocks. A photon is shot towards the end of the train. The two edge of the train correspond to the t' axis and its parallel translation—this is because in Oppie's frame the train is not moving and so it stays at $x' = 0$; similarly, the front of the strain stays at $x' = \ell_0$.

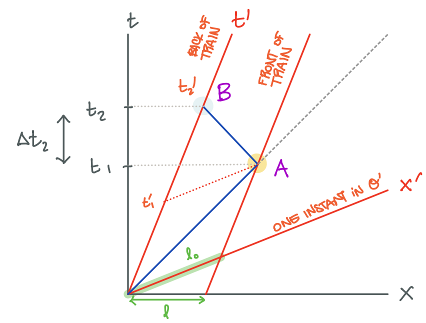


Figure 3.9: The trajectory of the photon (dark blue) from the back of the train at $t = t' = 0$ to the event A then the event B where it returns to the back of the train. Oppie's coordinate system is in red.

The photon travels at the speed of light $c = 1$, which is a 45° diagonal line on the way out, then a 135° line on the way back. The thought experiment is again simpler in Oppie's frame. Oppie measures that the total time for the back-and-forth journey is $t_2' = 2\ell'$. It is clear both from the thought experiment and the diagram that each leg of the journey takes the same amount of time $t_2' = 2t_1'$.

We see something a bit different. On the outbound journey the photon has to *catch up* to the front train. If we could observe it, we would measure that it takes time t_1 for the photon to reach the mirror. After it reflects off the mirror at the front of the train, the photon and the back of the train are moving *toward each other*. We observe that this total journey takes a time t_2. We also define $\Delta t_2 = t_2 - t_1$ to be just the time of the journey from the front of the train to the back again. Observe that $\Delta t_2 \neq t_1$; both logically and from Figure 3.9. This means we have the relations:

$$t_1 = \ell + \beta t_1 \qquad\qquad \Delta t_2 = \ell - \beta \Delta t_2 \ . \tag{3.A.6}$$

Here we recall that β is the speed of the train and the different signs correspond to whether the train is moving with or against the photon. Since t_1 is already measured against the origin, there is no need to write $\Delta t_1 \equiv t_1 - 0 = t_1$. These give us the relations:

$$(1-\beta)t_1 = \ell \qquad\qquad (1+\beta)\Delta t_2 = \ell \ . \tag{3.A.7}$$

The total time is then

$$t_2 = t_1 + \Delta t_2 = \frac{[(1+\beta) + (1-\beta)]\,\ell}{1-\beta^2} = 2\gamma^2 \ell \ . \tag{3.A.8}$$

Exercise 3.A.2. Prove (3.A.8). It is worth doing.

Now we have to invoke time dilation in (3.A.5). The total time for the round trip journey that we measure t_2 is related to the total time that Oppie measures t_2' by

$$t_2 = \gamma t_2' \tag{3.A.9}$$
$$2\gamma^2 \ell = 2\ell' \tag{3.A.10}$$
$$\ell = \frac{\ell'}{\gamma} \ . \tag{3.A.11}$$

Because $\gamma > 1$, we find that what *we* measure to be the length of the train is *contracted* (smaller) relative to what Oppie measures in the comoving frame of the train.

3.A.3 Confusion

The lessons of this *gedanken* appendix are simple. If the frame $\mathcal{O}'$ is moving relative to our frame, $\mathcal{O}$, then

- From (3.A.5): The time that we measure is *dilated* (slower) compared to the time that is measured in $\mathcal{O}'$.
- From (3.A.11): Length along the direction of motion are *contracted* (smaller) compared to the length along the direction of motion measured in $\mathcal{O}'$.

Having established these, I leave it to you to explore some of the apparent paradoxes in special relativity such as the pole-in-barn paradox and the twin paradox. The spacetime diagrams that we draw are useful guides. As a hint, often the resolution to paradoxes is the observation that simultaneity is a frame-dependent notion.[h]

[h] This has a manifestation for us: the locality of the fundamental interactions is a logical consequence of requiring causality in our theory. Interactions must happen at a single *event* rather than over a finite separation in order for there to be a clear *cause* that precedes an *effect* in any valid reference frame.

Thus far, everything we have discussed in this appendix is standard fare in a decent treatment of first-year modern physics. Let us focus on some of the conceptual hiccups that are not always explained carefully.

Could we derive length contraction without using time dilation? You should feel unsatisfied. In relativity, space and time have equal footing. However, *every* thought-experiment based derivation of special relativity starts by deriving time dilation and *then* uses that result to derive length contraction. The reason for this asymmetry is that the time dilation thought experiment did not depend on the directions perpendicular to the train's motion: both the $\mathcal{O}$ and $\mathcal{O}'$ frame measured the height of the train to be ℓ. On the other hand, the length contraction thought experiment had both spatial displacement ℓ *and* a time displacement t_2 that needed to be measured in each reference frame. To say it differently: the time dilation experiment allowed us to ignore part of the spatial configuration. However, in the length contraction experiment you *cannot* ignore the time separation because all experiments evolve in time.[i] Rest assured, the geometry *does* respect the symmetry between space and time.

[i] Philosophically: we perceive the universe as a sequence in time. We can imagine experiences where we are stationary in space, but we have no experiment—*gedanken* or otherwise—where we are stationary in time.

If they are so useful, why did we not draw a spacetime diagram for the time dilation experiment? The time dilation experiment made use of a third dimension of spacetime, the height of the train. It is a bit of a pain to draw, and you still end up drawing the same right triangles in Fig. 3.8

Was it necessary for the photon to complete a full loop? In these thought experiments, the the photon is emitted and detected by the same device: the trajectory is an 'out and back' in runners parlance. For the time dilation experiment, it is obviously sufficient to consider only the trajectory from the origin to the ceiling. What about for the length contraction experiment? This is also possible.

> **Exercise 3.A.3.** Derive time dilation using the same thought experiment, but without the 'return journey' to the back of the train.

That this is possible is obvious for those with experience with relativity. However, it is useful pedagogically to have the photon perform a round trip journey because this way all observations are performed by the same observer. I can say that *Oppie* emits a photon and *Oppie* observes its return. Then we can identify the event where we observe Oppie emitting a photon and the event where we observe Oppie observing the photon.[j]

[j] Given the subtleties of spacetime separated events with regard to simultaneity, I can appreciate this pedagogical choice.

I tried calculating this using Lorentz transformations and I got stuck. Here is a common error. Suppose we know that Oppie measures the train car to have length ℓ_0 and that Oppie measures this at some time slice $t' = 0$. Then we can perform a Lorentz transformation:

$$\begin{pmatrix} t \\ \ell \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta \\ \gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} 0 \\ \ell_0 \end{pmatrix} = \begin{pmatrix} \gamma\beta\ell_0 \gamma\ell_0 \end{pmatrix} . \tag{3.A.12}$$

From this we believe that $\ell = \gamma\ell_0$, i.e. length is *also* dilated—which is *incorrect*! What went wrong here? Recall that the *event* that we are using to measure the length of the train is a photon hitting the back of the train. That is labeled event A in Fig. 3.9. This occurs on the light cone where $t = x$ (and also $t' = x'$). That means that we need to Lorentz transform the separation between the *event* $(t'_1 = \ell_0, x' = \ell_0)$. This gives

$$\begin{pmatrix} t_1 \\ \ell \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta \\ \gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} \ell_0 \\ \ell_0 \end{pmatrix} = \begin{pmatrix} \gamma(1+\beta)\ell_0 \\ \gamma(1+\beta)\ell_0 \end{pmatrix} . \tag{3.A.13}$$

We can use the first component and our time dilation result to find:

$$t_1 \equiv \ell = \gamma(1+\beta)(1-\beta)\ell' \tag{3.A.14}$$

$$= \frac{\ell'}{\gamma} . \tag{3.A.15}$$

We have used the definition that $\gamma = (1-\beta^2)^{-1}$. From this we indeed confirm $\ell = \ell'/\gamma$ and that length is *contracted*.

3.B More examples in relativity

3.B.1 The electromagnetic field strength

Another place where special relativity rears its head is in electrodynamics. Electricity and magnetism are two manifestations of the same electromagnetic phenomenon.[a] This is illustrated in Fig. 3.10. If you did not know

[a] Unification of apparently unrelated forces is a big theme of this course. Electromagnetism is the example that you already know.

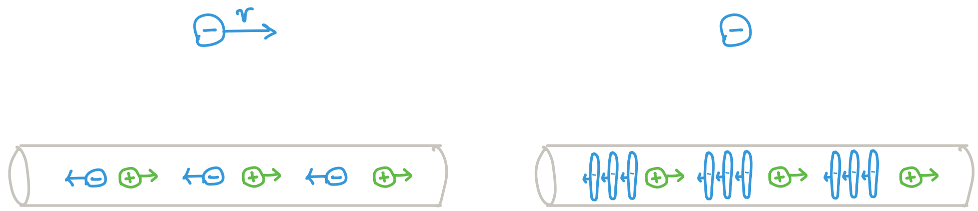


Figure 3.10: LEFT: an electron moves near a current. The current has no net electric charge. In the absence of magnetism, we expect the electron to move in a straight line. RIGHT: if we boost into the electron frame, the particles in the current moving in the opposite direction are length-contracted. This means that the charge density increases. The stationary electron now feels a net electric force. This implies that something is missing in the picture on the left.

about magnetism, you would find a paradox when you consider a charged particle moving along a current. In one frame, the current is an equal number of positive and negative charges moving in opposite directions.[b] If you boost to the external charged particle's rest frame, length contraction forces one species of the current particles to increase their charge density relative to the other species. This creates a net electric field that acts on the stationary external particle. Without magnetism, the first frame is missing this additional force.

[b] I suppose more realistically the negative charges move while the positive charges stay put—that does not change the conclusion.

The electric and magnetic fields are unified in the *electromagnetic field strength*, which is a two-index tensor:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^x & -E^y & -E^z \\ E^x & 0 & -B^z & B^y \\ E^y & B^z & 0 & -B_x \\ E^z & -B^y & B^x & 0 \end{pmatrix} \quad F_{\mu\nu} = \begin{pmatrix} 0 & E^x & E^y & E^z \\ -E^x & 0 & -B^z & B^y \\ -E^y & B^z & 0 & -B_x \\ -E^z & -B^y & B^x & 0 \end{pmatrix} . \tag{3.B.1}$$

We see that electricity and magnetism are unified in that their components mix into one another under a Lorentz transformation.

Exercise 3.B.1. Confirm that $F_{\mu\nu} = g_{\mu\alpha} g_{\nu\beta} F^{\alpha\beta}$ with the (3+1)-dimensional Minkowski metric $g_{\mu\nu}$.

Exercise 3.B.2. Find the components of $F'^{\mu\nu}$ under a boost along the z-direction.

3.B.2 Simultaneity

One of the key ideas in special relativity is that we sacrifice our notion of simultaneity for objects that are not observed at the same spacetime point. We may reuse our diagram in Fig. 3.5. Consider two different points on the x' axis. These are simultaneous with respect to the primed observer: they both occur at $t' = 9$. However, these two points obviously do *not* have the same t coordinate to the unprimed observer. This observation helps clear up several apparent paradoxes that may show up in relativity. More importantly, it completely upends our notion of causality.

One of the unwritten-but-understood laws of physics is that the cause precedes the effect. I have to drop my mug before I hear the sound of the ceramic shattering. This notion is imperiled if in some other reference frame someone else would have heard the shattering before they observe the mug being dropped. One deduction of this is that the laws of physics should be *local* in spacetime. A consequence of this observation is that the laws of physics should be written with derivatives.

3.B.3 Discrete isometries

Isometries are symmetries of the metric. The Minkowski metric has a parity isometry that acts as a discrete symmetry. In matrix form the action of parity is a Lorentz transformation

$$P = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} . \tag{3.B.2}$$

This symmetry is *discrete* in contrast to continuous symmetries. Rotations are a continuous symmetry because you can rotate by any real number amount, θ. They parameterize an infinite number of Lorentz transformations. Discrete symmetries, on the other hand, represent a countable number of Lorentz transformation. In fact, because $P^2 = \mathbb{1}$, there is only one such Lorentz transformation.

There is a second discrete transformation called **time reversal**:

$$T = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} . \tag{3.B.3}$$

This one sounds rather dramatic, doesn't it?[c] It should be clear that T indeed reverses the direction of time. This, however, does not mean that time reversal is something we can do.[d] Our understanding of causality breaks if we allow time reversal. However, mathematically time reversal is a clear isometry of the Minkowski metric.

In fact, there is something 'deep' to say that the classical laws of physics are time reversal invariant. If you run a video backwards, everything that happens obeys the laws of physics. The sign of the gravitational force may swap, but the dynamics of such an "anti-gravity" force law follows Newtonian mechanics. Entropy may decrease rather than increase, but there is no sense in which the microscopic transition from one configuration to the next would violate any laws of microphysics.[e] Though, to be fair, we also do not know how to take a right-handed person and do a parity transformation on them to turn them left-handed.

[c] One of my mentors in graduate school once told the story that he was preparing a lecture on particle physics at the pub. When his friends asked him what the topic of the lecture would be, he said "time reversal." The bar crowd suddenly grew silent until the bartender quietly asked: "... we can do that?"

[d] Time reversal is a major part of *The Legend of Zelda: Tears of the Kingdom*. Before this, there was the ground-breaking independent computer game *Braid* that pioneered this mechanic. The latter has an additional connection to physics in that the story is largely understood to be a parable about the development of atomic weapons.

[e] This is to say that the "arrow of time" from statistical mechanics is not a statement about microscopic interactions nor is it a statement about what is not *possible*, only about what is increasingly *improbable*.

In order to restrict to sensible isometries, we say that valid observers in special relativity are those that are related by isometries that are *connected to the identity*. This means that one may write the isometry with respect to a continuous parameter and that for some value of that parameter, the isometry is simply the identity matrix. In this way, we restrict our physical isometries to those that maintain the direction of time.

There is one more discrete symmetry that is often mentioned along with parity and time reversal: charge inversion. Unlike the other two, charge inversion is *not* a spacetime symmetry since it acts only on particles (fields). Charge inversion takes every particle and flips their charges. For now you may think about flipping the *electric* charge of the particle—but this actually holds for all of the types of charges that we examine in this course.[f] There are two combinations of these discrete symmetries that are notable:

[f] Charges are conserved quantities. Remember that conserved quantities come from symmetries of the action. Unlike the spacetime symmetries of this chapter, those charges are related to *internal* symmetries.

- The combination CP (charge–parity) is the transformation that takes a particle to its anti-particle.
- The combination $CPT \equiv \mathbb{1}$. That is: if we perform all three discrete symmetries, we return to the same state.[g]

[g] I leave this here with no proof. I know such proofs exist, but they are largely in the domain of a construction called axiomatic quantum field theory, which I know nothing about. You can learn more about this in[13]

Because these are all parities—in the sense that $C^2 = P^2 = T^2 = \mathbb{1}$—we see an relation between the antimatter transformation CP and the idea of moving backward in time. You may want to remember then when our Feynman diagram notation makes it *look like* antiparticles are particles moving backward in time.[h]

[h] To be clear, this is *not* what is happening.

3.C Bird's eye view of tensors

The reason why we make a big deal about tensors and the 'indexology' view of particle physics is that this perspective is particularly helpful in physics. Table 3.1, borrowed from my Physics 17 lecture notes, gives a hint of this. Column and row vectors have a few equivalent names:

- Upper index: [column] vector, contravariant vector, ket,
- Lower index: row/dual/covariant vector, covector, bra, one-form .

These are all fancy names for the *same* idea. A row vector[a] $\underaccent{\tilde}{\mathbf{w}}$ is a *linear function* on vectors to numbers. This means that given two vectors $\mathbf{v}$ and $\mathbf{u}$ and two numbers α and β:

[a] Here we stick to the simplest-sounding name.

$$\underaccent{\tilde}{\mathbf{w}}\left(\alpha\mathbf{v} + \beta\mathbf{u}\right) = \alpha\underaccent{\tilde}{\mathbf{w}}(\mathbf{v}) + \beta\underaccent{\tilde}{\mathbf{w}}(\mathbf{u}) \ . \tag{3.C.1}$$

Example 3.C.1. To make it clear that this really is a simple statement, let us

[13] Alexander S. Blum and Andrés Martínez de Velasco. "The genesis of the CPT theorem". In: *Eur. Phys. J. H* 47.1 (2022), p. 5. DOI: 10.1140/epjh/s13129-022-00037-w

try it out for a two-dimensional real vector space in matrix notation:

$$\begin{pmatrix} w_1 & w_2 \end{pmatrix} \left(\alpha \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} + \beta \begin{pmatrix} u^1 \\ u^2 \end{pmatrix} \right) = \alpha(w_1 v^1 + w_2 v^2) + \beta(w_1 u^1 + w_2 u^2) \ . \quad (3.C.2)$$

Example 3.C.2. Here is an even sillier example in a *one*-dimensional real vector space, $\mathbb{R}$. In this case, vectors are just numbers $v^1 = x$. Row vectors are linear functions, f characterized by one component, a. They satisfy:

$$f(\alpha x + \beta y) = \alpha x + \beta y \ . \quad (3.C.3)$$

Even though α, β, x, and y are practically all numbers, we should recognize that x and y are 'vectors' while α and β are 'numbers' in the sense of rescaling.

Exercise 3.C.1. By our definition of linear, show that the general equation for a line in the Cartesian plane, $f(x) = ax + b$ is *not* linear.

Of course, you can equivalently think about $\mathbf{v}$ acting on $\underset{\sim}{\mathbf{w}}$,

$$\mathbf{v}(\underset{\sim}{\mathbf{w}}) \coloneqq v^i w_i = w_i v^i = \underset{\sim}{\mathbf{w}}(\mathbf{v}) \ . \quad (3.C.4)$$

Here we see why it helps to think about indices rather than matrices as arrays of numbers that have some multiplication rule. Given a column and a row vector, there is one obvious way to form a number. If you treat one of the objects as a 'function' and the other object as an 'argument,' then the function is linear in the argument. In this sense, we say that the column and row vectors are *dual* to one another. They encode the same structure relative to each other.[b]

The bras and kets of quantum mechanics are also simply column and row vectors. This is obvious for simple finite-dimensional systems, like the spin states of an electron. After all, did we not write spinors as columns of two numbers?[c] There are also systems with countably-infinite dimensions, like the hydrogen atom. There we wrote our states as eigenvectors of energy and angular momentum, $|E, \ell, m\rangle$, where these *quantum numbers* take discrete-but-unbounded values. You will recall that these states also have bras $\langle E', \ell', m'|$ whose wavefunctions are related by a complex conjuate to the associated ket. There are also uncountably infinite quantum systems[d] like plane waves where a 'quantum'-number like the wave momentum can take on any value, $e^{i\mathbf{p}\cdot\mathbf{x}}$ where the 'infiniteness' of the vector space means that it does not make sense to use a discrete index for the components—instead, the components are part of a continuum. As a taste for how these all come together, present Table 3.1 reproduced from my Physics 17 course.

[b] This notion of duality is completely independent of physical dualities or parity symmetries of physical systems. However, it is true that if a particle has *internal* symmetries that are described by, say, upper indices, then its antiparticle has these indices lowered.

[c] They were kind of weird because they obviously were not vectors in Euclidean two dimensional space—but they *are* vectors in a different space. See Appendix C for a discussion—note: that appendix is significantly more advanced than most of this course.

[d] You may object that plane waves are not normalizable states.

Vector Space	$\mathbb{R}$	$\mathbb{C}$	∞-dimensional
Vector/ket	$\mathbf{v} = \lvert v\rangle$	$\mathbf{v} = \lvert v\rangle$	f
Basis vector	$\hat{\mathbf{e}}_i = \lvert i\rangle$	$\hat{\mathbf{e}}_i = \lvert i\rangle$	$\hat{e}_i(x)$
			$\hat{e}_p(x)$
Components	$v^i \in \mathbb{R}$	$v^i \in \mathbb{C}$	$f^i, \widetilde{f}(p) \in \mathbb{C}$
	$\mathbf{v} = v^i \hat{\mathbf{e}}_i = v^i \lvert i\rangle$	$\mathbf{v} = v^i \hat{\mathbf{e}}_i = v^i \lvert i\rangle$	$f(x) = f^i\, \hat{e}_i(x)$
			$f(x) = \int đp\, \widetilde{f}(p) e_p(x)$
Row vector/bra	$\underset{\sim}{\mathbf{w}} = w_i \underset{\sim}{\mathbf{e}}^i = w_i \langle i\rvert$	$\underset{\sim}{\mathbf{w}} = w_i \underset{\sim}{\mathbf{e}}^i = w_i \langle i\rvert$	distribution, e.g. $\delta(x)$
Matrix	$A = A^i{}_j \lvert i\rangle \langle j\rvert$	$A = A^i{}_j \lvert i\rangle \langle j\rvert$	operator, e.g. $\frac{d^2}{dx^2}$
Inner Product	$\langle v, w\rangle = g_{ij} v^i w^j$	$\langle v, w\rangle$	$\langle f, g\rangle = \int dx\, f^*(x) g(x)$
	$\langle v, w\rangle = \langle w, v\rangle$	$\langle v, w\rangle = \langle w, v\rangle^*$	$\langle f, g\rangle = \langle g, f\rangle^*$
Adjoint	Transpose	Hermitian Conjugate	Integration by parts
	$(A^T)^i{}_j = g_{jk} A^k{}_\ell g^{\ell i}$	$(A^\dagger)^i{}_j = [(A^T)^i{}_j]^*$	e.g. $\left(\frac{d}{dx}\right)^\dagger = -\left(\frac{d}{dx}\right)$
Self-adjoint	Symmetric	Hermitian	Sturm–Liouville
	$A^T = A$	$A^\dagger = A$	$\mathcal{O}^\dagger = \mathcal{O}$
	$\mathbb{R}$ Eigenvalues	$\mathbb{R}$ Eigenvalues	$\mathbb{R}$ Eigenvalues
	$\perp$ Eigenvectors	$\perp$ Eigenvectors	$\perp$ Eigenvectors
Isometry, e.g.	Rotations, Boosts	Unitary Matrices	Change of variable

Table 3.1: Terms and notation in real, complex, and infinite-dimensional vector spaces. From my Physics 17 notes.

Chapter 4

Quantum Mechanics

4.1 Sum Over Histories

[**Flip**: To be filled in. Please see the example in class of the N-slit experiment and the proposal that each path contributes e^{iS}.]

Rule 4.1.1 (Sum over histories)**.** Given some observed initial state $|\text{in}\rangle$ and some observed final state $|\text{out}\rangle$, the amplitude to go from $|\text{in}\rangle$ to $|\text{out}\rangle$ is

$$\mathcal{M}(\text{in} \to \text{out}) = \sum_{\text{path}} e^{iS_{\text{path}}} \ , \tag{4.1.1}$$

where the sum is over all 'paths' (histories) that permit the initial state to evolve into the final state. S_{path} is the action for a given path—the same action that you are used to from classical mechanics, the time integral over the Lagrangian, $S = \int dt\ L$.

4.2 Path Integral

Suppose a quantum state has initial and final positions[a]

$$|\text{in}\rangle = |q_0\rangle \qquad\qquad |\text{out}\rangle = |q_T\rangle \ . \tag{4.2.1}$$

These are labeled so that we observe $|q_0\rangle$ at time $t = 0$ and we observe $|q_T\rangle$ at time $t = T$. The amplitude to go from $|q_0\rangle \to |q_T\rangle$ over the time interval T is

$$\langle q_T| \hat{U}(T) |q_0\rangle = \langle q_T| e^{-iT\hat{H}} |q_0\rangle \ . \tag{4.2.2}$$

Here $\hat{U}(t)$ is the operator that evolves states over time t. As you know from quantum mechanics, the Hamiltonian $\hat{H}$, is the operator that enacts infinitesimal evolution in time. This means that we enact finite time translations by exponentiating the Hamiltonian.[b]

There are two tricks to deploy at this step:

[a] We deliberate use the notation $\boldsymbol{q}$ for position instead of $\boldsymbol{x}$. This is to aid our transition to quantum *field theory*. Old textbooks sometimes refer to *second quantization*; it is precisely to distinguish the role of $\boldsymbol{q}$, a field displacement, and $\boldsymbol{x}$ a continuous field index indicating spacetime position.

[b] Please make sure you are familiar with this idea: Hermitian operators *generate* infinitesimal changes to a system. The exponential of these Hermitian operators are unitary operators that enact finite transformations. The idea shows up over and over again in this class.

1. Break the finite time evolution operator into a product of $(N+1)$ time evolutions:

$$\hat{U}(T) = \hat{U}(\Delta t)\hat{U}(\Delta t)\cdots\hat{U}(\Delta t) \ . \tag{4.2.3}$$

It is *obvious* that $\hat{U}(t)\hat{U}(s) = \hat{U}(t+s)$.

2. Insert the identity as a sum over a complete set of intermediate states:

$$\mathbb{1} = \int \mathrm{d}q_i \ |q_i\rangle\langle q_i| \ . \tag{4.2.4}$$

We insert a copy of the identity in-between each of the $U(\Delta t)$s. Thus the label i runs from 1 to N.

Exercise 4.2.1. Confirm that $\hat{U}(t)\hat{U}(s) = \hat{U}(t+s)$. Hint: this is trivial, but convince yourself that it is trivial.

Applying these two tricks to our amplitude[c]

[c] $\mathrm{d}^N\mathbf{q} := \mathrm{d}q_1 \cdots \mathrm{d}q_N$

$$\langle q_T|\,\hat{U}(T)\,|q_0\rangle = \int \mathrm{d}^N\mathbf{q} \ \langle q_T|\,\hat{U}(\Delta T)\,|q_N\rangle\,\langle q_N|\,\hat{U}(\Delta T)\,|q_{N-1}\rangle \cdots \tag{4.2.5}$$

$$= \int \mathrm{d}^N\mathbf{q} \ \prod_i^{N+1} K_i \ , \tag{4.2.6}$$

where the factor K_i is a complex number:

$$K_i := \langle q_i|\,\hat{U}(\Delta T)\,|q_{i-1}\rangle \ , \tag{4.2.7}$$

where we define $\langle q_{N+1}| = \langle q_T|$. We have reduced the amplitude to go from our initial state to our final state to a product of K_i. The form of K_i is simply that of the amplitude to go from $|q_{i-1}\rangle$ to $|q_i\rangle$ over a small time Δt. The integral over intermediate states $\mathrm{d}^N\mathbf{q}$ tells us that we are summing over all possible configurations of those unobserved intermediate states.

In order to evaluate K_i, let us remind ourselves of the form of the Hamiltonian in $U(\Delta T) = \exp(i\Delta T\hat{H})$,

$$\hat{H} = \frac{\hat{P}^2}{2m} + V(\hat{Q}) \ , \tag{4.2.8}$$

where we use the non-relativistic kinetic term $\frac{1}{2}m\dot{q}^2 = p^2/2m$ and an arbitrary potential $V(q)$ that only depends on the q displacements, but not the momenta p. Because $\hat{H}$ is an operator, we write $\hat{P}$ for the momentum operator and $\hat{Q}$ for the position[d] operator. All Hamiltonians in ordinary field theory have this type of separation: there is a *kinetic term* that depends on momentum and a *potential term* that depends on displacements. When we pass to relativistic theories, the kinetic term converts into a form that includes spatial derivatives in a way that it is Lorentz invariant.

[d] This 'position' is not necessarily spatial position. In general it is a displacement of a field value.

Once again there is another silly trick.

3. Insert the momentum-space representation of the identity,

$$\mathbb{1} = \int đp\ |p\rangle\langle p| \qquad\qquad đp := \frac{\mathrm{d}p}{2\pi}\ . \tag{4.2.9}$$

Observe that this has a relative factor of 2π compared to the resolution of the 'position space' identity. This is the famous—and seemingly inscrutable—factor that shows up in the Fourier transform, see Appendix A.[e][1]

[e] If you spend time thinking about this, you may wonder why we do not split the 2π evenly between positions and momenta. This would certainly be symmetric, but this choice ends up being fairly convenient in physics. It means, for example, that the conjugate variable for time is essentially the angular momentum, $E = \omega$.

Because $\hat{H}$ contains a term that evaluates easily in momentum space and a term that evaluates easily in position space, we insert the complete set of momentum states into K_i to allow the $\hat{P}$-dependent terms to hit a $\langle p|$ while the Q-dependent terms hit a $|q_i\rangle$:

$$K_i = \int đp\, \langle q_i|p\rangle\, \langle p|\, \mathrm{e}^{-i\Delta t\frac{\hat{P}^2}{2m}}\, \mathrm{e}^{-i\Delta t V(\hat{Q})}\, |q_{i-1}\rangle \tag{4.2.10}$$

$$= \int đp\ \mathrm{e}^{-i\Delta t\frac{p^2}{2m}}\, \mathrm{e}^{-i\Delta t V(q_{i-1})}\, \langle q_i|p\rangle\langle p|q_{i-1}\rangle\ . \tag{4.2.11}$$

The factors on the right are simply the Fourier components between the position and momentum eigenstates:

$$\langle q_i|p\rangle = \mathrm{e}^{-ipq_i} \qquad\qquad \langle p|q_{i-1}\rangle = \mathrm{e}^{+ipq_{i-1}}\ . \tag{4.2.12}$$

Plugging this into K_i gives

$$K_i = \mathrm{e}^{-i\Delta t V(q_{i-1})} \int đp\ \mathrm{e}^{-i\Delta t\frac{p^2}{2m} - ip(q_i - q_{i-1})}\ . \tag{4.2.13}$$

The integral is simply a Gaussian with a source term. If you do not worry[f] about the factors of i, then this integral is precisely in the form of (A.1.4), which stated:

[f] Will not worry. On the other hand, you may want to worry about whether the assumptions we made to solve these Gaussian integrals is still valid when there are complex coefficients. This scratches the surface of why some mathematicians are skeptical about whether anything we do in quantum field theory is well posed.

$$G_{a,J} \equiv \int \mathrm{d}x\ e^{-\frac{a}{2}x^2 + Jx} = \sqrt{\frac{2\pi}{a}}\, e^{\frac{J^2}{2a}}\ . \tag{A.1.4}$$

We identify

$$a = \frac{i\Delta t}{m} \qquad\qquad J = -i(q_i - q_{i-1})\ . \tag{4.2.14}$$

Plugging in this result gives

$$K_i = \mathrm{e}^{-i\Delta t V(q_{i-1})}\, \frac{1}{2\pi}\sqrt{\frac{2\pi m}{i\Delta t}}\, \mathrm{e}^{\frac{-(q_i - q_{i-1})^2}{2i\Delta t/m}} \tag{4.2.15}$$

$$= \sqrt{\frac{m}{i2\pi\Delta t}}\, \mathrm{e}^{\frac{i}{2}m\frac{(q_i - q_{i-1})^2}{\Delta t^2}\Delta t}\, \mathrm{e}^{-iV(q_{i-1})\Delta t} \tag{4.2.16}$$

$$= \mathcal{N}\, \mathrm{e}^{i\Delta t\, L[q_i]}\ . \tag{4.2.17}$$

[1] See e.g. https://physics.stackexchange.com/a/141737/166736

In the last line we have identified $(q_i - q_{i-1})/\Delta t = \dot{q}_i$ and observed that the argument of the combined exponential is simply $i\Delta t$ times the *Lagrangian* at point i. The ugly square root is some complex constant factor that *does not matter*. We simply write it as a normalization $\mathcal{N}$.

Now we insert this into the amplitude to go from our initial state to our final state, (4.2.6):

$$\langle q_T | \, \hat{U}(T) \, |q_0\rangle = \mathcal{N}^{N+1} \int \mathrm{d}^N\mathbf{q} \;\; \mathrm{e}^{i\sum_i \Delta t \, L[q_i]} \; . \tag{4.2.18}$$

Recalling that i encodes 'time slices,' we recognize that the argument of the exponential is simply an *integral* of the Lagrangian,

$$i\sum_i \Delta t \, L[q_i] \to i \int \mathrm{d}t \;\, L[q(t)] \equiv iS[q(t)] \; , \tag{4.2.19}$$

where we have recovered the action. In the continuum limit, it is conventional to replace the notation $\mathrm{d}^N\mathbf{q}$ with $\mathcal{D}q(t)$ so that our final expression for the $\Delta t \to 0$ limit is

$$\langle q_T | \, \hat{U}(T) \, |q_0\rangle = \mathcal{N}' \int \mathcal{D}q(t) \;\; \mathrm{e}^{iS[q(t)]} \; . \tag{4.2.20}$$

Here $\mathcal{N}'$ is a re-packaging of the powers of $\mathcal{N}$. The integration over all trajectories $\mathcal{D}q(t)$ is called a **path integral**.

> **Exercise 4.2.2.** Why was it necessary for us to split $U(T)$ into several small steps? HINT: did we make any assumptions to go from (4.2.10) to (4.2.11)?

4.3 Recovering Lagrange

[**Flip**: To be filled in. Discussion of phasor diagrams, steepest descent, and why the classical path dominates. Motivates the principle of least action as the $\hbar \to 0$ limit.]

4.4 Recovering Schrödinger

This section is optional and is outside the primary scope of the class. However, I encourage you to at least skim it to see an alternative explanation of where the Schrödinger equations comes from. It should convince you that the sum over histories (path integral) construction gives you the same dynamics as before.

We show that the Schrödinger equation appears as the single-particle, non-relativistic limit over the path integral formalism.[a] The power of

[a] This section draws from the excellent pedagogical article by David Derbes in the *American Journal of Physics* (one of my favorite publications) in 1996. This artcile was selected as one of the Fermat's Library journal club articles.[2]

[2] David Derbes. "Feynman's derivation of the Schrödinger equation". In: *American Journal of Physics* 64.7 (July 1996), pp. 881–884. ISSN: 0002-9505. DOI: 10.1119/1.18114. eprint: https://pubs.aip.org/aapt/ajp/article-pdf/64/7/881/11886238/881_1_online.pdf. URL: https://doi.org/10.1119/1.18114; https://fermatslibrary.com/s/feynmans-derivation-of-the-schrodinger-equation

the path integral formalism is that it generalizes beyond non-relativistic quantum mechanics. Like with all new theoretical formalisms, it behooves one to confirm that the new formalism is mathematically equivalent to the established formalism in the appropriate limit.

Suppose we interpret $q(t)$ as the position of a non-relativistic particle at time t. Then in this basis, we can think of the amplitude $\langle q_t|\hat{U}(t)|q_0\rangle$ as the wavefunction $\psi(q,t)$ for the particle to be in position q at time t given some initial state $|q_0\rangle$ at $t=0$. We can rewrite (4.2.6) as follows:

$$\psi(q,t) = \langle q_t|\hat{U}(t)|q_0\rangle \tag{4.4.1}$$

$$= \int \mathrm{d}q_N \; K_N \int \mathrm{d}^{N-1}\mathbf{q} \; \prod_i^N K_i \tag{4.4.2}$$

$$= \int \mathrm{d}q' \; K_N \, \psi(q', t-\Delta t) \; . \tag{4.4.3}$$

We have relabeled $q_N \to q'$. It is the last displacement variable that we integrate over. In fact, for convenience,[b] let us shift all of our time coordinates by Δt:

[b] You do not have to do this step.

$$\psi(q, t+\Delta t) = \int \mathrm{d}q' \; K_N \, \psi(q', t) \; . \tag{4.4.4}$$

This equation says that to go from the $(N-1)^{\text{th}}$ step to the N^{th} step, you integrate the wavefunction at last position over all possible intermediate steps, weighted by the *kernel*,[c] K_N. We can now use the form of the kernel, (4.2.17):

[c] Kernels are just fancy names for distributions, which are in turn fancy names for functions that are integrated over.

$$K_N = \mathcal{N}\, \mathrm{e}^{i\Delta t\, L[q_i]} \qquad\qquad \mathcal{N} = \sqrt{\frac{m}{2\pi i \Delta t}} \; . \tag{4.4.5}$$

Let us now be careful with the scaling with the small time slice Δt:

$$i\Delta t L[q_i] = \frac{im}{2}\frac{(q_i - q_{i-1})^2}{\Delta t} - i\Delta t V(\bar{q}) \; . \tag{4.4.6}$$

Note that the kinetic term goes like Δt^{-1}, which is *large*. The potential term scales linearly with Δt, so it is small. We write $V(\bar{q})$ to mean the average value of V in the interval (q_{i-1}, q_i). For our purposes, $\bar{q} = q$ assuming a sufficiently smooth potential. This means we can Taylor expand the potential term, but not the kinetic term:

$$\psi(q, t+\Delta t) = \mathcal{N} \int \mathrm{d}q' \;\; \mathrm{e}^{\frac{im}{2}\frac{(q-q')^2}{\Delta t}} \left[1 - i\Delta t\, V(\bar{q})\right] \, \psi(q', t) \; . \tag{4.4.7}$$

We drop higher order terms in Δt. Now we assume that $q = q' + \xi$ and make the plausible case that ξ is small.[d] This means we can also replace $\bar{q} \to q$ to leading order. We can then expand $\psi(q', t)$, where we shall suppress the t argument so that spatial derivatives are clear:

[d] To be clear, we integrate over even large values of ξ. The sin that I am sweeping under the rug is the idea that the rapidly varying oscillation in the $\mathrm{e}^{\text{kinetic}}$ term will cause integrants with large values of ξ to cancel. This is the steepest descent method in the method of phasors.

$$\psi(q') = \psi(q) - \xi\psi'(q) + \frac{1}{2}\xi^2\psi''(q) + \cdots \; . \tag{4.4.8}$$

To be clear: we now write $\psi'(q,t)$ to mean $\frac{\mathrm{d}}{\mathrm{d}q}\psi(q,t)$. Plugging this in:

$$\psi(q,t+\Delta t) = \mathcal{N}\int \mathrm{d}\xi \;\; \mathrm{e}^{\frac{im}{2}\frac{\xi^2}{\Delta t}}\left[1 - i\Delta t\, V(q)\right]\left[\psi(q,t) - \xi\psi'(q,t) + \frac{1}{2}\xi^2\psi''(q,t)\right] \tag{4.4.9}$$

$$= \mathcal{N}\int \mathrm{d}\xi \;\; \mathrm{e}^{\frac{im}{2}\frac{\xi^2}{\Delta t}}\left[\psi(q,t) - i\Delta t\, V(q)\,\psi(q,t) + \frac{1}{2}\xi^2\psi''(q,t)\right] . \tag{4.4.10}$$

In the last line we removed the term linear in ξ because it cancels against the rest of the integral, which is even in ξ. We also remove terms of order $\mathcal{O}(\Delta t\, \xi^2)$ since these are small. Observe that the ξ-independent terms in the bracket are simple Gaussian integrals:

$$\int \mathrm{d}\xi \;\; \mathrm{e}^{\frac{im}{2}\frac{\xi^2}{\Delta t}} = \sqrt{\frac{2\pi i\Delta t}{m}} = \mathcal{N}^{-1} . \tag{4.4.11}$$

Observe that this factor exactly cancels the annoying normalization in (4.4.5)—but only for the ξ-independent terms in the bracket. The remaining term is a moment of the Gaussian integral. There happens to be a convenient trick for this by taking a derivative of (A.1.4):

$$G_{a,J} \equiv \int \mathrm{d}x \;\; e^{-\frac{a}{2}x^2+Jx} = \sqrt{\frac{2\pi}{a}}e^{\frac{J^2}{2a}} , \tag{A.1.4}$$

where we see that

$$-2\frac{\mathrm{d}G_{a,0}}{\mathrm{d}a} = \int \mathrm{d}x \;\; x^2 e^{-\frac{a}{2}x^2} = \frac{1}{a}\sqrt{\frac{2\pi}{a}} . \tag{4.4.12}$$

Exercise 4.4.1. While the method above reconstructs even moments of the Gaussian integral, one can reconstruct all moments by taking derivatives with respect to the source, J. Show that you can reconstruct the second moment of the Gaussian integral, (4.4.12), by taking two derivatives with respect to J.

This gives us

$$\int \mathrm{d}\xi \;\; \xi^2\, \mathrm{e}^{\frac{im}{2}\frac{\xi^2}{\Delta t}} = \frac{1}{-im/\Delta t}\sqrt{\frac{2\pi i\Delta t}{m}} = \frac{i\Delta t}{m}\mathcal{N}^{-1} . \tag{4.4.13}$$

We can insert (4.4.11) and (4.4.13) into the expression for $\psi(q,t+\Delta t)$, (4.4.10). The ugly factor of $\mathcal{N}$ conveniently cancels. We are left with:

$$\psi(q,t+\Delta t) = \psi(q,t) - i\Delta t V(q)\,\psi(q,t) + \frac{i\Delta t}{2m}\psi''(q,t) \tag{4.4.14}$$

We can rearrange terms and recognize the time derivative:

$$\frac{\mathrm{d}\psi(q,t)}{\mathrm{d}t} = i\left[\frac{1}{2m}\left(\frac{\mathrm{d}}{\mathrm{d}q}\right)^2 - V(q)\right]\psi(q,t) , \tag{4.4.15}$$

which we recognize as the Schrödinger equation. You can debate whether or not this counts as a 'derivation' of the Schrödinger equation since we implicitly assumed the Schrödinger equation when writing the time evolution operator. However, if you instead took as a starting point the observation that each path is weighted by $\mathrm{e}^{iS[q(t)]}$, then one would recover the Schrödinger equation as a consequence.

Exercise 4.4.2. Derive the same results by taking a time derivative of the continuous expression (4.2.20). The steps are analogous, but this will test if you understand what the continuous expression means.

Chapter 5

Closing Thoughts

Acknowledgments

PT thanks all the people who taught him quantum field theory and particle physics over the years. In particular, courses from Scott Thomas, Pat Burchat, Savas Dimopoulos, Aaron Roodman, Michael Peskin, Shamit Kachru, David Tong, Maciej Dunajski, Ben Allanach, Hugh Osborn, Fernando Quevedo, Silvia Pascoli, Csaba Csáki, Maxim Perelstein, and Yuval Grossman. I am further indebted to those who were (and are still) on this journey to figure this all out—those are too many to list explicitly, but I especially thank my postdoctoral mentors Tim Tait and Jonathan Feng, and everyone who has ever shared an office or done problem sets with me. My approach to writing and pedagogy is inspired by the writing of Sidney Coleman, Anthony Zee, David Tong, and Matthew Strassler—the physicists who you read if they have ever written anything vaguely related to whatever it is you are trying to learn. PT is supported by a NSF CAREER award (#2045333).

Appendix A

Gauss, Dirac, Fourier

We review some key mathematical tools: Gaussian integrals, the Dirac δ-function, and Fourier transform conventions.

A.1 Gaussian Integrals

A.1.1 One dimension

The most basic Gaussian integral is

$$G \equiv \int_{-\infty}^{\infty} \mathrm{d}x \ \ \mathrm{e}^{-\frac{1}{2}x^2} = \sqrt{2\pi} \ . \tag{A.1.1}$$

In what follows I will always assume that the integration limits are $(-\infty, \infty)$ unless otherwise explicitly stated. There is a trick: evaluate G^2 in polar coordinates:

$$G^2 = \int dx\, dy \ \ \mathrm{e}^{-\frac{x^2+y^2}{2}} = \int_0^{\infty} r \ \mathrm{d}r \ \mathrm{d}\theta \ \ \mathrm{e}^{-\frac{1}{2}r^2} \ . \tag{A.1.2}$$

You can then use a substitution $w = r^2/2$ and the fact that $\int_0^\infty \mathrm{d}w \ \ e^{-w} = 1$.

A.1.2 Rescaling

A trivial extension is to rescale the quadratic term in the exponential.

$$G_a \equiv \int \mathrm{d}x \ \ \mathrm{e}^{-\frac{a}{2}x^2} = \sqrt{\frac{2\pi}{a}} \ . \tag{A.1.3}$$

There are two ways of seeing this:

1. One can perform a change of variables by defining $y \equiv \sqrt{a}x$, which case $\mathrm{d}x \ = 1/\sqrt{a} \ \mathrm{d}y$. Fromt his one finds $G_a = G/\sqrt{a}$.
2. A slicker way of doing this is by dimensional analysis. One recognizes that by assigning arbitrary dimensions to $[x] = 1$, the dimension of the coefficient is $[a] = -1/2$ in order for the argument of the exponential to be dimensionless. Then we see that $[G_a] = 1$ and so the right-hand side must scale like $a^{-1/2}$. The any additional coefficient must 1 because $G_{a=1} = G$.

A.1.3 Adding a source

For reasons that are clear in a physical context, we can shift the argument of the exponential by adding a linear term

$$G_{a,J} \equiv \int \mathrm{d}x \; e^{-\frac{a}{2}x^2 + Jx} = \sqrt{\frac{2\pi}{a}} e^{\frac{J^2}{2a}} \; . \tag{A.1.4}$$

The the constant J is called a **source**. The trick we use is completing the square:

$$(a+b)^2 = a^2 + 2ab + b^2 \Rightarrow a^2 + 2ab = (a+b)^2 - b^2 \; . \tag{A.1.5}$$

Then by shifting your integration variable, this reduces to the result for G_a times an overall prefactor.

A.1.4 Higher dimensions

Let's ramp it up! Suppose you have N variables, q^1, q^2, $\cdots$, q^N. These are not powers, they are upper indices. We have changed variables from x to q for reasons that will be clear when we pass to field theory. Show that

$$\mathbf{G}_A \equiv \int \mathrm{d}q^{\,1} \cdots \mathrm{d}q^{\,N} \; e^{-\frac{1}{2}q^i A_{ij} q^j} = \sqrt{\frac{(2\pi)^N}{\det A}} \; . \tag{A.1.6}$$

We can think of each of the N integration variables as being components of an N component vector $\mathbf{q}$ so that the integration is actually over N-dimensional real space, $\mathbb{R}^N$. The argument of the exponential is

$$q^i A_{ij} q^j = \mathbf{q}^{\mathrm{T}} A \mathbf{q} \; , \tag{A.1.7}$$

where A is an $N \times N$ matrix with components A_{ij}.[a] Assume that A is a symmetric matrix, $A_{ij} = A_{ji}$ that is *non-degenerate* (nonzero eigenvalues).

[a] If you are concerned about index heights, *good*! For now just use the summation convention. Alternatively, you may remember that 'real matrices' have an index structure $A^i{}_j$ and 'row vectors' have an index structure $\mathbf{q}^{\mathrm{T}} = q_i$ so that we can write $\mathbf{q}^{\mathrm{T}} A \mathbf{q} = q_i A^i{}_j q^j$. In our Euclidean space we have a trivial metric, so the upper/lower indices are just a formality.

Exercise A.1.1. Prove (A.1.6). It may help to remember that you can diagonalize a symmetric matrix with a rotation, R. Because $R^{\mathrm{T}}R = I$, the identity, we may "multiply by one" to show:

$$\mathbf{q}^{\mathrm{T}} A \mathbf{q} \to \mathbf{q}^{\mathrm{T}} R^{\mathrm{T}} R A R^{\mathrm{T}} R \mathbf{q} = (R\mathbf{q})^{\mathrm{T}} \left(R A R^{\mathrm{T}} \right) R\mathbf{q} = (R\mathbf{q})^{\mathrm{T}} \hat{A} R \mathbf{q} \tag{A.1.8}$$

where we define the *diagonal* matrix

$$\hat{A} \equiv R A R^{\mathrm{T}} \; . \tag{A.1.9}$$

Remember that $\hat{A} = \mathrm{diag}(\lambda_1, \cdots, \lambda_N)$: the diagonalized matrix contains the eigenvalues along the diagonal. Remember that the determinant of such a matrix is really easy. Finally, remember that rotating coordinate systems does not change your integration measure:

$$\mathrm{d}^N \mathbf{q} \equiv \mathrm{d}q^1 \cdots \mathrm{d}q^N = \mathrm{d}^n \mathbf{s} \tag{A.1.10}$$

where $\mathbf{s} \equiv R\mathbf{q}$. This should be enough for the evaluation of $\mathbf{G}_A$ to reduce to N products of G_a.

A.2 Integral representation of $\delta(x)$

As a starting point, we start with the Fourier representation of the Dirac δ-function.[a] Recall that the δ-function is defined by

[a] The $\boldsymbol{\delta}$-function is not formally a function, it is a *distribution*. This means it only makes sense when it is being integrated over.

$$\int_{-\infty}^{\infty} \mathrm{d}x\ \delta(x)\, f(x) = \lim_{\varepsilon\to 0}\int_{-\varepsilon}^{\varepsilon} \mathrm{d}x\ \delta(x)\, f(x) = f(0) \tag{A.2.1}$$

$$\delta(x) = 0 \text{ if } x \neq 0\ . \tag{A.2.2}$$

$\delta(0)$ is not defined—and you had better make sure you never see it show up in the prediction of any physical quantity. In this way, δ "cancels" an integral and picks out a value of the integrand,

$$\int_{-\infty}^{\infty} \mathrm{d}x\ \delta(x - x_0)\, f(x) = f(x_0)\ . \tag{A.2.3}$$

There is a useful integral representation of the δ-function:

$$\delta(x) = \int_{\infty}^{\infty} \mathrm{d}\xi\ e^{2\pi i x\xi}\ . \tag{A.2.4}$$

[**Flip**: To fill in. I was going to derive this expression.]

A.3 A general Fourier transform

There are two choices one can make when defining a Fourier transform convention[a]; we parameterize these choices by real numbers a and b. The Fourier transform $\widetilde{f}(\omega)$ of a function $f(t)$ is

[a] This appendix draws from an excellent discussion on Physics Stack Exchange.[1]

$$\widetilde{f}(\omega) = \sqrt{\frac{|b|}{(2\pi)^{1-a}}}\int_{-\infty}^{\infty} \mathrm{d}t\ e^{ib\omega t} f(t)\ . \tag{A.3.1}$$

We see that a tells us about the (2π) factors and b tells us about the argument of the basis function $e^{ib\omega t}$. With this basis, the inverse Fourier transform is

$$f(t) = \sqrt{\frac{|b|}{(2\pi)^{1+a}}}\int_{-\infty}^{\infty} \mathrm{d}\omega\ e^{-ib\omega t} f(\omega)\ . \tag{A.3.2}$$

One may check that the inverse Fourier transform of a Fourier transform gives the original function:

$$\widetilde{\widetilde{f}} = \frac{|b|}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\ e^{-ib\omega t}\int_{-\infty}^{\infty} ds\, e^{ib\omega s} f(s) \tag{A.3.3}$$

$$= \frac{|b|}{2\pi}\int ds\, f(z)\int \mathrm{d}\omega\ e^{ib\omega(s-t)} \tag{A.3.4}$$

$$= \int \mathrm{d}s\ \delta(s-t) f(s)\ , \tag{A.3.5}$$

where we have used $\int \mathrm{d}\xi\ \exp(2\pi i x\xi) = \delta(x)$.

[1] https://physics.stackexchange.com/a/308248

A.4 Our Conventions

The convention that we will choose for the *time–frequency* [inverse] Fourier transform is

$$f(t) = \int_{-\infty}^{\infty} đ\omega \, e^{-i\omega t} \widetilde{f}(\omega) \qquad\qquad đ\omega \equiv \frac{d\omega}{2\pi} \ . \tag{A.4.1}$$

This corresponds to $a = b = 1$. The corresponding transform for the frequency-domain function is

$$\widetilde{f}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}t \ e^{i\omega t} f(t) \ . \tag{A.4.2}$$

A.5 Higher Dimensions

All of this generalizes to higher dimensions: you simply Fourier transform each dimension. In fact, one is free to use a different Fourier transform convention for each direction. We can use this freedom to pick a convention that 'automatically' fits our conventions for spacetime. In particular, given a four-vector $x = (t, \mathbf{x})$ and its conjugate four-momentum $p = (\omega, \mathbf{k})$, one may choose to Fourier transform as follows:

$$f(x) = \int đ\omega \ đ^3\mathbf{k} \, e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} \widetilde{f}(p) \ . \tag{A.5.1}$$

With this convention, the basis function is simply

$$e^{-i(\omega t - \mathbf{k}\cdot\mathbf{x})} = e^{-ip\cdot x} \ , \tag{A.5.2}$$

where $p \cdot x$ is the usual Minkowski dot product, $p_\mu x^\mu$. This makes it clear that the basis function is Lorentz invariant. The Fourier transform would still respect the spacetime symmetries even if we had not chosen a convenient notation—it just wouldn't be as simple to see.

Example A.5.1. Statistical Mechanics. One motivation for our Fourier convention is statistical mechanics. One formulation of classical statistical mechanics is to assume that phase space is discrete: a particle has momentum $\mathbf{p}$ whose components take integer multiples of some unit momentum, h. Assuming that the particle has g internal degrees of freedom (e.g. $g = 2$ for a particle that can be spin-up or spin-down), then the density of states is g/h^3. Quite remarkably in the history of physics, the value of h can be identified with Planck's constant in quantum mechanics. In natural units we take $\hbar = h/(2\pi) \equiv 1$, so the phase space density is $g/(2\pi)^3$. For a particle with a phase space distribution function $f(\mathbf{x}, \mathbf{p})$, this means that the number density of particles is

$$n = g \int đ^3p \, f(p) \ . \tag{A.5.3}$$

We see that it is convenient to take a convention where every dp comes with a $(2\pi)^{-1}$.

Exercise A.5.1. Lorentz-Invariant Phase Space. In relativistic systems, the energy and the momenta are related by $E^2 = \mathbf{p}^2 + m^2$. We are, of course, using natural units where $c = 1$. The phase space integral over $đ^3p$ is thus also an integral over the energy. In order to enforce the relativistic relation, the full phase space density is usually written as $đ^4p\,(2\pi)\delta(E^2 - p^2 - m^2)$. Show that integrating over the δ-function gives

$$\int đ^4p\,\delta(E^2 - p^2 - m^2) = \int \frac{đ^3\mathbf{p}}{2E(p)} \qquad E(p) \equiv \sqrt{p^2 + m^2}\,. \tag{A.5.4}$$

Appendix B

Notation and Conventions

B.1 Spacetime Conventions

4D Minkowski indices are written with lower-case Greek letters from the middle of the alphabet, $\mu, \nu, \cdots$. 5D indices are written in capital Roman letters from the middle of the alphabet, $M, N, \cdots$. Tangent space indices are written in lower-case Roman letters from the beginning of the alphabet, $a, b, \cdots$. Flavor indices are written in lower-case Roman letters near the beginning of the alphabet, $i, j, \cdots$.

Dirac spinors Ψ are related to left- and right-chiral Weyl spinors ($\chi, \bar{\psi}$ respectively) via

$$\Psi = \begin{pmatrix} \chi \\ \bar{\psi} \end{pmatrix}. \tag{B.1.1}$$

Note that sometimes we will write $\Psi = (\psi, \bar{\chi})^T$. The point is that un-barred Weyl spinors are—by convention—left-handed while barred spinors are right-handed. Our convention for σ^0 and the three Pauli matrices σ is

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{B.1.2}$$

with the flat-space γ matrices given by

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \qquad \gamma^5 = \begin{pmatrix} i\mathbb{1} & 0 \\ 0 & -i\mathbb{1} \end{pmatrix}, \tag{B.1.3}$$

where $\bar{\sigma}^\mu = (\sigma^0, -\sigma)$. This convention for γ^5 gives us the correct Clifford Algebra. (Note that this differs from the definition of γ^5 in Peskin & Schroeder.)

Appendix C

Representations of the Poincare Group

This appendix is slightly more advanced than the rest of these notes.

We briefly review the Poincaré group and its spinor representations. Readers with a strong background in field theory will be familiar with these topics and can skip this subsection. Other readers with a weaker background in the representations of the Poincaré group are encouraged to do peruse more thorough literature on this topic. Excellent references include section 2 of Weinberg, Vol. I,[1] Section 1 of Buchbinder and Kuzenko,[2] Section 5 of Gutowski,[3] Section 4 of Osborn,[4] and Section 10 of Jones.[5]

C.1 The Poincaré group and its properties

The **Poincaré group** describes the symmetries of Minkowski space and is composed of transformations of the form

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \tag{C.1.1}$$

where a^μ parameterizes translations and $\Lambda^\mu{}_\nu$ parameterizes transformations of the Lorentz group containing rotations and boosts. We can write elements of the Poincaré group as $\{(\Lambda, a)\}$. A pure Lorentz transformation is thus $(\Lambda, 0)$ while a pure translation is $(\mathbb{1}, a)$. Elements are multiplied according to the rule

$$(\Lambda_2, a_2) \cdot (\Lambda_1, a_1) = (\Lambda_2\Lambda_1,\, \Lambda_2 a_1 + a_2.). \tag{C.1.2}$$

[1] Steven Weinberg. *The Quantum Theory of Fields, Vol I-III*. Cambridge University Press, 2005.

[2] I. L. Buchbinder and S. M. Kuzenko. *Ideas and methods of supersymmetry and supergravity: Or a walk through superspace*. Bristol, UK: IOP (1998) 656 p.

[3] Jan Gutowski. *Symmetries and Particle Physics*. Michaelmas term Part III course, Department of Applied Mathematics and Theoretical Physics, Cambridge University. 2007.

[4] Hugh Osborn. *Symmetries and Particle Physics*. Lecture Notes. 2010. URL: http://www.damtp.cam.ac.uk/user/ho/GNotes.pdf.

[5] H. F. Jones. *Groups, representations and physics, 2nd Ed.* Institute of Physics, 1998.

Note that these transformations *do not commute*,

$$(\Lambda, 0) \cdot (\mathbb{1}, a) = (\Lambda, \Lambda a) \tag{C.1.3}$$
$$(\mathbb{1}, a) \cdot (\Lambda, 0) = (\Lambda, a). \tag{C.1.4}$$

Thus the Poincaré group is *not* a direct product of the Lorentz group and the group of 4-translations. The technical term for the relation between these groups is that the Poincaré group is a **semi-direct product** of the Lorentz and 4-translation groups.

Locally the Poincaré group is represented by the algebra

$$[M^{\mu\nu}, M^{\rho\sigma}] = i(M^{\mu\sigma}\eta^{\nu\rho} + M^{\nu\rho}\eta^{\mu\sigma} - M^{\mu\rho}\eta^{\nu\sigma} - M^{\nu\sigma}\eta^{\mu\rho}) \tag{C.1.5}$$
$$[P^\mu, P^\nu] = 0 \tag{C.1.6}$$
$$[M^{\mu\nu}, P^\sigma] = i(P^\mu\eta^{\nu\sigma} - P^\nu\eta^{\mu\sigma}). \tag{C.1.7}$$

The $\mathbf{M}$ are the antisymmetric generators of the Lorentz group,

$$(M^{\mu\nu})_{\rho\sigma} = i(\delta^\mu_\rho\delta^\nu_\sigma - \delta^\mu_\sigma\delta^\nu_\rho), \tag{C.1.8}$$

and the $\mathbf{P}$ are the generators of translations. We will derive the form of the Lorentz generators below. As a 'sanity check,' one should be able to recognize in (C.1.5) the usual $O(3)$ Euclidean symmetry by taking $\mu, \nu, \rho, \sigma \in \{1, 2, 3\}$ and noting that at most only one term on the right-hand side survives. One may check that this coincides with the algebra for angular momenta, $\mathbf{J}$. Equation (C.1.6) says that translations commute, while (C.1.7) says that the generators of translations transform as vectors under the Lorentz group. This is, of course, expected since the generators of translations are precisely the four-momenta. The factors of i should also be clear since we're taking the generators $\mathbf{P}$ and $\mathbf{M}$ to be Hermitian.

One can represent this algebra in matrix form as

$$\left(\begin{array}{cc:c} \multicolumn{2}{c:}{M} & P \\ \hdashline 0 & 0 & 1 \end{array}\right). \tag{C.1.9}$$

One can check explicitly this reproduces the algebra in (C.1.5–C.1.7) and the multiplication law (C.1.2). The 'translation' part of the Poincaré algebra is boring and requires no further elucidation. It is the Lorentz algebra that yields the interesting features of our fields under Poincaré transformations.

C.2 The Lorentz Group

Let us now explore the **Lorentz group**, which is sometimes called the **homogeneous Lorentz group** to disambiguate it from the Poincaré group which is sometimes called the **inhomogeneous Lorentz group**.

The Lorentz group is composed of the transformations that preserve the inner product on Minkowski space, $\langle x^\mu, x^\nu\rangle = x^\mu\eta_{\mu\nu}x^\nu = x^\mu x_\mu$. In particular, for $x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu$, we have

$$\left(\Lambda^\mu{}_\rho x^\rho\right)\eta_{\mu\nu}\left(\Lambda^\nu{}_\sigma x^\sigma\right) = x^\rho\eta_{\rho\sigma}x^\sigma. \tag{C.2.1}$$

From this we may deduce that the fundamental transformations of the Lorentz group satisfy the relation

$$\Lambda^{\mu}{}_{\rho}\,\eta_{\mu\nu}\,\Lambda^{\nu}{}_{\sigma} = \eta_{\rho\sigma}, \tag{C.2.2}$$

or in matrix notation,

$$\mathbf{\Lambda}^{\mathbf{T}}\eta\mathbf{\Lambda} = \eta, \tag{C.2.3}$$

where $\eta = \mathrm{diag}(+,-,-,-)$ is the usual Minkowski metric used by particle physicists[a].

[a] If you are a particle physicist born after 1980 and you write papers using a different metric, then I am mad at you.

C.2.1 Generators of the Lorentz Group

Let's spell out the procedure for determining the generators of the Lorentz group. We will later follow an analogous procedure to determine the generators of supersymmetry. We start by writing out any [finite] Lorentz transformation as the exponentiation

$$\mathbf{\Lambda} = e^{it\mathbf{W}}, \tag{C.2.4}$$

where t is a transformation parameter and $\mathbf{W}$ is the generator we'd like to determine. We stick with the convention that generators of unitary representations should be Hermitian. Clever readers will question whether it is true that *all* Lorentz transformations can be written as the exponentiation of a generator at the identity. This is true for the cases of physical interest, where we only deal with the part of the subgroup which is connected to the identity. We will discuss the disconnected parts shortly.

Plugging (C.2.4) into (C.2.3) and setting $t = 0$, we obtain the relation

$$\eta\mathbf{W} + \mathbf{W}^T\eta = 0, \tag{C.2.5}$$

or with explicit indices,

$$\eta_{\mu\rho}W^{\rho}{}_{\nu} + W^{\rho}{}_{\mu}\eta_{\rho\nu} = 0 \tag{C.2.6}$$

$$W_{\mu\nu} + W_{\nu\mu} = 0. \tag{C.2.7}$$

Thus the generators $\mathbf{W}$ are 4×4 antisymmetric matrices characterized by six real transformation parameters so that there are six generators. Let us thus write the exponent of the finite transformation (C.2.4) as

$$it\,W^{\lambda\sigma} = it\omega^{\mu\nu}\left(M_{\mu\nu}\right)^{\lambda\sigma}, \tag{C.2.8}$$

where $\omega^{\mu\nu}$ is an antisymmetric 4×4 matrix parameterizing the linear combination of the independent generators and $\left(M_{\mu\nu}\right)^{\lambda\sigma}$ are the Hermitian generators of the Lorentz group. The μ,ν indices label the six generators, while the λ,σ indices label the matrix structure of each generator.

We may thus verify that (C.1.8) indeed furnishes a basis for the generators of the Lorentz group connected to the identity

$$(M^{\mu\nu})_{\rho\sigma} = i(\delta^{\mu}_{\rho}\delta^{\nu}_{\sigma} - \delta^{\mu}_{\sigma}\delta^{\nu}_{\rho}). \tag{C.2.9}$$

The qualification "connected to the identity" turns out to be rather important, as we shall see when we consider representations of the Lorentz algebra.

C.2.2 Components of the Lorentz Group

Recall that the Lorentz group has four disconnected parts. The defining (C.2.3) implies that

$$(\det \mathbf{\Lambda})^2 = 1, \tag{C.2.10}$$

$$({\Lambda^0}_0)^2 - \sum_i ({\Lambda_0}^i)^2 = 1, \tag{C.2.11}$$

where the first equation comes from taking a determinant and the second equation comes from taking $\rho = \sigma = 0$ in (C.2.2). From these equations we see that

$$\det \mathbf{\Lambda} = \pm 1 \tag{C.2.12}$$

$${\Lambda^0}_0 = \pm \sqrt{1 + \sum_i ({\Lambda_0}^i)^2}. \tag{C.2.13}$$

The choice of the two signs on the right-hand sides of these equations labels the four components of the Lorentz group. One cannot form a smooth path in the space of Lorentz transformations starting in one component of and ending in another (i.e. they are disconnected).

The component of the Lorentz group with $\det \mathbf{\Lambda} = +1$ contains the identity element and is a subgroup that preserves parity. In order to preserve the direction of time, one ought to further choose ${\Lambda^0}_0 \geq 1$. We shall specialize to this subgroup, which is called the **orthochronous Lorentz group**, $SO(3,1)^{\uparrow}_{+}$ which further satisfies

$$\det \mathbf{\Lambda} = +1 \tag{C.2.14}$$

$$\Lambda^0_0 \geq 1. \tag{C.2.15}$$

Other parts of the Lorentz group can be obtained from $SO(3,1)^{\uparrow}_{+}$ by applying the transformations

$$\Lambda_P = \text{diag}(+,-,-,-) \tag{C.2.16}$$

$$\Lambda_T = \text{diag}(-,+,+,+). \tag{C.2.17}$$

Here Λ_P and Λ_T respectively refer to parity and time-reversal transformations. One may thus write the Lorentz group in terms of 'components' (not necessarily 'subgroups'),

$$SO(3,1) = SO(3,1)^{\uparrow}_{+} \oplus SO(3,1)^{\uparrow}_{-} \oplus SO(3,1)^{\downarrow}_{+} \oplus SO(3,1)^{\downarrow}_{-}, \tag{C.2.18}$$

where the up/down arrow refers to ${\Lambda^0}_0$ greater/less than ± 1, while the $\pm$ refers to the sign of $\det \mathbf{\Lambda}$. Again, only $SO(3,1)^{\uparrow}_{\pm}$ form subgroups. In these notes we will almost exclusively work with the orthochronous Lorentz group so that we will drop the $\pm$ and write this as $SO(3,1)^{\uparrow}$.

It is worth noting that the fact that the Lorentz group is not simply connected is related to the existence of a 'physical' spinor representation, as we will mention below.

C.2.3 The Lorentz Group is related to $SU(2)\times SU(2)$

Locally the Lorentz group is related to the group $SU(2)\times SU(2)$, i.e. one might suggestively write

$$SO(3,1) \approx SU(2)\times SU(2). \tag{C.2.19}$$

Let's flesh this out a bit. One can explicitly separate the Lorentz generators $M^{\mu\nu}$ into the generators of rotations, J_i, and boosts, K_i:

$$J_i = \frac{1}{2}\,\epsilon_{ijk}\,M_{jk} \tag{C.2.20}$$

$$K_i = M_{0i}, \tag{C.2.21}$$

where ϵ_{ijk} is the usual antisymmetric Levi-Civita tensor. $\mathbf{J}$ and $\mathbf{K}$ satisfy the algebra

$$[J_i, J_j] = \quad i\epsilon_{ijk}J_k \tag{C.2.22}$$

$$[K_i, K_j] = -i\epsilon_{ijk}J_k \tag{C.2.23}$$

$$[J_i, K_j] = \quad i\epsilon_{ijk}K_k. \tag{C.2.24}$$

We can now define 'nice' combinations of these two sets of generators,

$$A_i = \frac{1}{2}(J_i + iK_i) \tag{C.2.25}$$

$$B_i = \frac{1}{2}(J_i - iK_i). \tag{C.2.26}$$

This may seem like a very arbitrary thing to do, and indeed it's *a priori* unmotivated. However, we now see that the algebra of these generators decouple into two $SU(2)$ algebras,

$$[A_i, A_j] = i\,\epsilon_{ijk}\,A_k \tag{C.2.27}$$

$$[B_i, B_j] = i\,\epsilon_{ijk}\,B_k \tag{C.2.28}$$

$$[A_i, B_j] = 0. \tag{C.2.29}$$

Magic! Note, however, that from (C.2.25) and (C.2.26) that these generators are *not* Hermitian (gasp!). Recall that a Lie group is generated by Hermitian operators.[b] Thus we were careful above *not* to say that $SU(3,1)$ *equals* $SU(2)\times SU(2)$, where 'equals' usually means either isomorphic or homomorphic. As a further sanity check, $SU(2)\times SU(2)$ is manifestly compact while the Lorentz group cannot be since the elements corresponding to boosts can be arbitrarily far from the origin. This is all traced back to the sign difference in the time-like component of the metric, i.e. the difference between $SO(4)$ and $SO(3,1)$. While rotations are Hermitian and generate unitary matrices, boosts are anti-Hermitian and generate anti-unitary matrices. At this level, then, our representations are non-unitary.

Anyway, we needn't worry about the precise sense in which $SO(3,1)$ and $SU(2)\times SU(2)$ are related, the point is that we may label representations of $SO(3,1)$ by the quantum numbers of $SU(2)\times SU(2)$, (A,B). For example, a Dirac spinor is in the $(\frac{1}{2},\frac{1}{2}) = (\frac{1}{2},0)\oplus(0,\frac{1}{2})$ representation, i.e. the direct sum of two Weyl reps. To connect back to reality[c], the physical meaning of all this is that we may write the spin of a representation as $J = A + B$.

[b] Mathematicians often use anti-Hermitian generators.

[c] Ope! There goes gravity. Apologies to Eminem, "Lose Yourself."

C.2.4 An aside: complexified algebras

So how are $SO(3,1)$ and $SU(2) \times SU(2)$ *actually* related?

We've been deliberately vague about the exact relationship between the Lorentz group and $SU(2) \times SU(2)$. The precise relationship between the two groups are that the *complex* linear combinations of the generators of the Lorentz algebra are isomorphic to the *complex* linear combinations of the Lie *algebra* of $SU(2) \times SU(2)$.

$$\mathcal{L}_{\mathbb{C}}(SO(3,1)) \cong \mathcal{L}_{\mathbb{C}}(SU(2) \times SU(2)) \tag{C.2.30}$$

Be careful not to say that the Lie algebras of the two groups are identical, it is important to emphasize that only the *complexified* algebras are identifiable. The complexification of $SU(2) \times SU(2)$ is the special linear group, $SL(2,\mathbb{C})$. In the next subsection we will identify $SL(2,\mathbb{C})$ as the **universal cover** of the Lorentz group. First, however, we shall show that the Lorentz group is isomorphic to $SL(2,\mathbb{C})/\mathbb{Z}_2$.

C.2.5 The Lorentz group is isomorphic to $SL(2,\mathbb{C})/\mathbb{Z}_2$

While the Lorentz group and $SU(2) \times SU(2)$ were neither related by isomorphism nor homomorphism, we *can* concretely relate the Lorentz group to $SL(2,\mathbb{C})$. More precisely, the Lorentz group is isomorphic to the coset space $SL(2,\mathbb{C})/\mathbb{Z}_2$

$$SO(3,1) \cong SL(2,\mathbb{C})/\mathbb{Z}_2 \tag{C.2.31}$$

Recall that we may represent four-vectors in Minkowski space as complex Hermitian 2×2 matrices via $V^\mu \to V_\mu \sigma^\mu$, where the σ^μ are the usual Pauli matrices,

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{C.2.32}$$

$SL(2,\mathbb{C})$ is the group of complex 2×2 matrices with unit determinant. It is spanned precisely by these Pauli matrices. Purists will admonish us for not explicitly distinguishing between a Lie group ($SL(2,\mathbb{C})$) and its algebra ($sl(2,\mathbb{C})$ or $\mathcal{L}[SL(2,\mathbb{C})]$ or $\mathfrak{sl}(2,\mathbb{C})$). The distinction is not worth the extra notational baggage since the meaning is clear in context

To be explicit, we may associate a vector $\mathbf{x}$ with either a vector in Minkowski space $\mathbb{M}^4$ spanned by the unit vectors e^μ,

$$\mathbf{x} = x^\mu e_\mu = \left(x^0, x^1, x^2, x^3\right), \tag{C.2.33}$$

or with a matrix in $SL(2,\mathbb{C})$,

$$\mathbf{x} = x_\mu \sigma^\mu = \begin{pmatrix} x_0 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{pmatrix}. \tag{C.2.34}$$

Note the lowered indices on the components of x_μ, i.e. $(x^0, x^1, x^2, x^3) = (x_0, -x_1, -x_2, -x_3)$. The four-vector components are recovered from the $SL(2,\mathbb{C})$ matrices via

$$x_0 = \frac{1}{2}\mathrm{Tr}(\mathbf{x}), \qquad\qquad x_i = \frac{1}{2}\mathrm{Tr}(\mathbf{x}\sigma^i). \tag{C.2.35}$$

The latter of these is easy to show by expanding $\mathbf{x} = x_0\mathbb{1}^0 + x_i\sigma^i$ and then noting that $\mathbb{1}\sigma^i \propto \sigma^i$, $\sigma^j\sigma^i|_{j\neq i} \propto \sigma^k_{k\neq i}$, and $\sigma^i\sigma^i|_{\text{no sum}} \propto \mathbb{1}$. Thus only the $\sigma^i\sigma^i$ term of $\mathbf{x}\sigma^i$ has a trace, so that taking the trace projects out the other components.

For the Minkowski four-vectors, we already understand how a Lorentz transformation $\mathbf{\Lambda}$ acts on a [covariant] vector x^μ while preserving the vector norm,

$$|\mathbf{x}|^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2. \tag{C.2.36}$$

This is just the content of (C.2.3), which defines the Lorentz group.

For Hermitian matrices, there is an analogous transformation by the action of the group of invertible complex matrices of unitary determinant, $SL(2,\mathbb{C})$. For $\mathbf{N} \in SL(2,\mathbb{C})$, $\mathbf{N}^\dagger\mathbf{x}\mathbf{N}$ is also in the space of Hermitian 2×2 matrices. Such transformations preserve the determinant of $\mathbf{x}$,

$$\det\mathbf{x} = x_0^2 - x_1^2 - x_2^2 - x_3^2. \tag{C.2.37}$$

The equivalence of the right-hand sides of s (C.2.36) and (C.2.37) are very suggestive of an identification between the Lorentz group $SO(3,1)$ and $SL(2,\mathbb{C})$. Indeed, (C.2.37) implies that for each $SL(2,\mathbb{C})$ matrix $\mathbf{N}$, there exists a Lorentz transformation Λ such that

$$\mathbf{N}^\dagger x^\mu\sigma_\mu\mathbf{N} = (\Lambda x)^\mu\sigma_\mu. \tag{C.2.38}$$

A very important feature should already be apparent: the map from $SL(2,\mathbb{C}) \to SO(3,1)$ is two-to-one. This is clear since the matrices $\mathbf{N}$ and $-\mathbf{N}$ yield the *same* Lorentz transformation, $\Lambda^\mu{}_\nu$. Hence it is not $SO(3,1)$ and $SL(2,\mathbb{C})$ that are isomorphic, but rather $SO(3,1)$ and $SL(2,\mathbb{C})/\mathbb{Z}_2$.

The point is that one will miss something if one only looks at representations of $SO(3,1)$ and not the representations of $SL(2,\mathbb{C})$. This 'something' is the spinor representation. How should we have known that $SL(2,\mathbb{C})$ is the important group? One way of seeing this is noting that $SL(2,\mathbb{C})$ is **simply connected** as a group manifold.

By the polar decomposition for matrices, any $g \in SL(2,\mathbb{C})$ can be written as the product of a unitary matrix U times the exponentiation of a traceless Hermitian matrix h,

$$g = Ue^h. \tag{C.2.39}$$

We may write these matrices explicitly in terms of real parameters $a, \cdots, g$;

$$h = \begin{pmatrix} c & a - ib \\ a + ib & -c \end{pmatrix} \tag{C.2.40}$$

$$U = \begin{pmatrix} d + ie & f + ig \\ -f + ig & d - ie \end{pmatrix}. \tag{C.2.41}$$

Here a, b, c are unconstrained while $d, \cdots, g$ must satisfy

$$d^2 + e^2 + f^2 + g^2 = 1. \tag{C.2.42}$$

Thus the space of 2×2 traceless Hermitian matrices $\{h\}$ is topologically identical to $\mathbb{R}^3$ while the space of unit determinant 2×2 unitary matrices $\{U\}$ is topologically identical to the three-sphere, S_3. Thus we have

$$SL(2, \mathbb{C}) = \mathbb{R}^3 \times S_3. \tag{C.2.43}$$

As both of the spaces on the right-hand side are simply connected, their product, $SL(2, \mathbb{C})$, is also simply connected. This is a 'nice' property because we can write down any element of the group by exponentiating its generators at the identity. But even further, since $SL(2, \mathbb{C})$ is simply connected, its quotient space $SL(2, \mathbb{C})/\mathbb{Z}_2 = SO(3, 1)^\uparrow$ is *not* simply connected.

C.2.6 Universal cover of the Lorentz group

The fact that $SO(3, 1)^\uparrow$ is not simply connected should bother you. In the back of your mind, your physical intuition should be unsatisfied with non-simply connected groups. This is because simply-connected groups have the very handy property of having a one-to-one correspondence between representations of the group and representations of its algebra; i.e. we can write any element of the group as the exponentiation of an element of the algebra about the origin.

What's so great about this property? In quantum field theory fields transform according to representations of a symmetry's algebra, not representations of the group. Since $SO(3, 1)^\uparrow$ is not simply connected, the elements of the algebra at the identity that we used do *not* tell the whole story. They were fine for constructing finite elements of the Lorentz group *that were connected to the identity*, but they don't capture the *entire* algebra of $SO(3, 1)^\uparrow$.[d]

[d] To spoil the surprise, the point is this: the elements of the algebra of $SO(3, 1)^\uparrow$ at the identity capture the vector (i.e. fundamental) representation of the Lorentz group, but it misses the spinor representation.

Now we're in a pickle. Given a group, we know how to construct representations of an algebra near the identity based on group elements connected to the identity. But this only characterizes the entire algebra if the group is simply connected. $SO(3, 1)^\uparrow$ is *not* simply connected. Fortunately, there's a trick. It turns out that for any connected Lie group, there exists a unique 'minimal' simply connected group that is homeomorphic to it called the **universal covering group**.

Stated slightly more formally, for any connected Lie group G, there exists a simply connected universal cover $\widetilde{G}$ such that there exists a homomorphism $\pi : \widetilde{G} \to G$ where $G \cong \widetilde{G}/\ker \pi$ and $\ker \pi$ is a discrete subgroup of the center of $\widetilde{G}$. Phew, that was a mouthful. For the Lorentz group this statement is $SO(3, 1)^\uparrow \cong SL(2, \mathbb{C})/\mathbb{Z}_2$. Thus the key statement is:

The Lorentz group is covered by $SL(2, \mathbb{C})$.

The point is that the homomorphism π is locally one-to-one and thus G and $\widetilde{G}$ have the same Lie algebras. Thus we can determine the Lie algebra of G away from the identity by considering the Lie algebra for $\widetilde{G}$ at the identity.

This universal covering group of $SO(3,1)^{\uparrow}$ is often referred to as $Spin(3,1)$. The name is no coincidence, it has everything to do with the spinor representation.

C.2.7 Projective representations

For the uninitiated, it may not be clear why the above rigamarole is necessary or interesting. Here we would like to approach the topic from a different direction to explain why the spinor representation is necessarily the 'most basic' representation of four-dimensional spacetime symmetry.

A typical "representation theory for physicists" course goes into detail about constructing the usual tensor representations of groups but only mentions the spinor representation of the Lorentz group in passing. Students 'inoculated' with a quantum field theory course will not bat an eyelid at this, since they're already used to the technical manipulation of spinors. But where does the spinor representation come from if all of the 'usual' representations we're used to are tensors?

The answer lies in quantum mechanics. Recall that when we write representations U of a group G, we have $U(g_1)U(g_2) = U(g_1 g_2)$ for $g_1, g_2 \in G$. In quantum physics, however, physical states are invariant under phases, so we have the freedom to be more general with our multiplication rule for representations: $U(g_1)U(g_2) = U(g_1 g_2)\exp(i\phi(g_1, g_2))$. Such 'representations' are called **projective representations**. In other words, quantum mechanics allows us to use projective representations rather than ordinary representations.

It turns out that not every group admits 'inherently' projective representations. In cases where such reps are not allowed, a representation that one *tries* to construct to be projective can have its generators redefined to reveal that it is actually an ordinary non-projective representation. The relevant mathematical result for our purposes is that groups which are *not* simply connected—such as the Lorentz group—*do* admit inherently projective representations.

The Lorentz group is *doubly* connected, i.e. going over any loop *twice* will allow it to be contracted to a point. This means that the phase in the projective representation must be ± 1. One can consider taking a loop in the Lorentz group that corresponds to rotating by 2π along the $\hat{z}$-axis. Representations with a projective phase $+1$ will return to their original state after a single rotation, these are the particles with integer spin. Representations with a projective phase -1 will return to their original state only after *two* rotations, and these correspond to fractional-spin particles, or spinors.

An excellent discussion of projective representations can be found in Weinberg, Volume I.[6] More on the representation theory of the Poincaré

[6]Steven Weinberg. "The Quantum theory of fields. Vol. 1: Foundations". In: ().

group and its SUSY extension can be found in Buchbinder and Kuzenko.[7] Further pedagogical discussion of spinors can be found in.[8] Some discussion of the topology of the Lorentz group can be found in Frankel section 19.3a.[9]

C.2.8 Lorentz representations are non-unitary and non-compact

If you weren't vigilant you might have missed a potential deal-breaker. We mentioned briefly that the representations of the Lorentz group are not unitary. The generators of boosts are imaginary. This makes them *anti-unitary* rather than unitary. From the point of view of quantum mechanics this is the kiss of death since we know that only *unitary* representations preserve probability. Things aren't looking so rosy anymore, are they?

Where does this non-unitarity come from?[e] It all comes from the factor of i associated with boosts. Just look at (C.2.25) and (C.2.26). This factor of i is crucial since is is related to the non-compactness of the Lorentz group. This is a very intuitive statement: rotations are compact since the rotation parameter lives on a circle ($\theta = 0$ and $\theta = 2\pi$ are identified) while boosts are non-compact since the rapidity can take on any value along the real line. The dreaded factor of i, then, is deeply connected to the structure of the group. In fact, it's precisely the difference between $SO(3,1)$ and $SO(4)$, i.e. the difference between space and time: a minus sign in the metric. To make the situation look even more grim, even if we were able to finagle a way out of the non-unitarity issue (and we can't), there is a theorem that unitary representations of non-compact groups are infinite-dimensional. There is nothing infinite-dimensional about the particles we hope to describe with the Lorentz group. This is looking like quite a pickle!

[e] Maybe we can fix it?

Great—what do we do now? Up until now we'd been thinking about representations of the Lorentz group as if they would properly describe particles. Have we been wasting our time looking at representations of the Lorentz group? No; fortunately not—but we'll have to wait until we examine the representations of this group before we properly resolve this apparent problem. The key, however, is that one must look at *full* Poincaré group (incorporating translations as well as Lorentz transformations) to develop a consistent picture. Adding the translation generator P to the algebra surprisingly turns out to cure the ills of non-unitarity (i.e. of non-compactness). The cost for these features, as mentioned above, is that the representations will become infinite dimensional, but this infinite dimensionality is well-understood physically: we can boost into any of a continuum of frames where the particle has arbitrarily boosted four-

Cambridge, UK: Univ. Pr. (1995) 609 p.

[7] Buchbinder and Kuzenko, *Ideas and methods of supersymmetry and supergravity: Or a walk through superspace*.

[8] Ethan D. Bolker. "The Spinor Spanner". In: *The American Mathematical Monthly* 80.9 (1973), pp. 977–984. ISSN: 00029890. URL: http://www.jstor.org/stable/2318771.

[9] T. Frankel. *The geometry of physics: An introduction, 2nd Ed.* Cambridge University Press, 2004.

momentum.

This is why most treatments of this subject don't make a big deal about the Lorentz group not being satisfactory for particle representations. They actually end up being rather useful for describing *fields*, where the non-unitarity of the Lorentz representations isn't a problem because the actual states in the quantum Hilbert space are the *particles* which are representations of the full Poincaré group.

For now our strategy will be to sweep any concerns about unitarity under the rug and continue to learn more about representations of the Lorentz group. We will then get back to the matter at hand and settle the present issues by working with representations of the Poincaré group. The take-home message, however, is that it is not sufficient to just have the Lorentz group; nature really needs the full Poincaré group to make sense of itself. This principle of extending symmetries is a motivation for studying supersymmetry and extra dimensions, and indeed has been a guiding principle for particle physics over the past three decades.

C.2.9 Spinors: the fundamental representation of $SL(2, \mathbb{C})$

The representations of the universal cover of the Lorentz group, $SL(2, \mathbb{C})$, are spinors. Most standard quantum field theory texts do calculations in terms of four-component Dirac spinors. This has the benefit of representing all the degrees of freedom of a typical Standard Model massive fermion into a single object.[f] Remember that the Standard Model is a *chiral* theory which already is most naturally written in terms of Weyl spinors—that's why the Standard Model Feynman rules in Dirac notation always end up with ugly factors of $\frac{1}{2}(\mathbb{1} \pm \gamma^5)$. A comprehensive guide review of two-component spinors can be found in Dreiner et al..[10]

Let us start by defining the **fundamental** and **conjugate** (or **antifundamental**) representations of $SL(2, \mathbb{C})$. These are just the matrices $N_\alpha{}^\beta$ and $(N^*)_{\dot{\alpha}}{}^{\dot{\beta}}$. Don't be startled by the dots on the indices, they're just a book-keeping device to keep the fundamental and the conjugate indices from getting mixed up. One cannot contract a dotted with an undotted $SL(2, \mathbb{C})$ index; this would be like trying to contract spinor indices (α or $\dot{\alpha}$) with vector indices (μ): they index two totally different representations.[g]

We are particularly interested in the objects that these matrices act on. Let us thus define **left-handed Weyl spinors**, ψ, as those acted upon by the fundamental rep and **right-handed Weyl spinors**, $\overline{\chi}$, as those that are acted upon by the conjugate rep. Again, do not be startled with the extra jewelry that our spinors display. The bar on the right-handed spinor just serves to distinguish it from the left-handed spinor. To be clear, they're both spinors, but they're different types of spinors that have different types of indices and that transform under different representations of $SL(2, \mathbb{C})$.

[f] In supersymmetry, on the other hand, it will turn out to be natural to work with two-component spinors. For example, a complex scalar field has two real degrees of freedom. In order to have a supersymmetry between complex scalars and fermions, we require that the number of degrees of freedom match for both types of object. A Dirac spinor, however, has four real degrees of freedom ($2\times$ (4 complex degrees of freedom) $-$ 4 from the Dirac equation). Weyl spinors, of course, have just the right number of degrees of freedom.

[g] This doesn't mean that we can't swap between different types of indices. In fact, this is exactly what we did in (C.2.33) and (C.2.34). We will get to the role of the σ matrices very shortly.

[10] Herbi K. Dreiner, Howard E. Haber, and Stephen P. Martin. "Two-component spinor techniques and Feynman rules for quantum field theory and supersymmetry". In: (2008). arXiv: 0812.1594 [hep-ph].

Explicitly,

$$\psi'_\alpha = N_\alpha{}^\beta \psi_\beta \tag{C.2.44}$$

$$\overline{\chi}'_{\dot\alpha} = (N^*)_{\dot\alpha}{}^{\dot\beta}\, \overline{\chi}_{\dot\beta}. \tag{C.2.45}$$

C.3 Invariant Tensors

We know that $\eta_{\mu\nu}$ is invariant under $SO(3,1)$ and can be used (along with the inverse metric) to raise and lower $SO(3,1)$ indices. For $SL(2,\mathbb{C})$, we can build an analogous tensor, the unimodular antisymmetric tensor

$$\epsilon^{\alpha\beta} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{C.3.1}$$

Unimodularity (unit determinant) and antisymmetry uniquely define the above form up to an overall sign. The choice of sign ($\epsilon^{12} = 1$) is a convention. As a mnemonic, $\epsilon^{\alpha\beta} = i\,(\sigma^2)_{\alpha\beta}$, but note that this is not a formal equality since we will see below that the index structure on the σ^2 is incorrect. This tensor is invariant under $SL(2,\mathbb{C})$ since

$$\epsilon'^{\alpha\beta} = \epsilon^{\rho\sigma} N_\rho{}^\alpha N_\sigma{}^\beta = \epsilon^{\alpha\beta} \det N = \epsilon^{\alpha\beta}. \tag{C.3.2}$$

We can now use this tensor to raise undotted $SL(2\mathbb{C})$ indices:

$$\psi^\alpha \equiv \epsilon^{\alpha\beta}\psi_\beta. \tag{C.3.3}$$

To lower indices we can use an analogous unimodular antisymmetric tensor with two lower indices. For consistency, the overall sign of the lowered-indices tensor must be defined as

$$\epsilon_{\alpha\beta} = -\epsilon^{\alpha\beta}, \tag{C.3.4}$$

so that raising and then lowering returns us to our original spinor:

$$\epsilon_{\alpha\beta}\epsilon^{\beta\gamma} = \delta^\gamma_\alpha. \tag{C.3.5}$$

This is to ensure that the upper- and lower-index tensors are inverses, i.e. so that the combined operation of raising then lowering an index does not introduce a sign. Dotted indices indicate the complex conjugate representation, $\epsilon^*_{\alpha\beta} = \epsilon_{\dot\alpha\dot\beta}$. Since ϵ is real we thus use the same sign convention for dotted indices as undotted indices,

$$\epsilon^{\dot 1\dot 2} = \epsilon^{12} = -\epsilon_{\dot 1\dot 2} = -\epsilon_{12}. \tag{C.3.6}$$

We may raise dotted indices in exactly the same way:

$$\overline{\chi}^{\dot\alpha} \equiv \epsilon^{\dot\alpha\dot\beta}\overline{\chi}_{\dot\beta}. \tag{C.3.7}$$

In order to avoid sign errors, it is a useful mnemonic to always put the ϵ tensor directly to the left of the spinor whose indices it is manipulating, this way the index closest to the spinor contracts with the spinor index. In other words, one needs to be careful since $\epsilon^{\alpha\beta}\psi_\beta \neq \psi_\beta\epsilon^{\beta\alpha}$.

In summary:

$$\psi^\alpha = \epsilon^{\alpha\beta}\psi_\beta \qquad \psi_\alpha = \epsilon_{\alpha\beta}\psi^\beta \qquad \overline{\chi}^\alpha = \epsilon^{\dot\alpha\dot\beta}\psi_{\dot\beta} \qquad \overline{\chi}_{\dot\alpha} = \epsilon_{\dot\alpha\dot\beta}\overline{\chi}^{\dot\beta}. \tag{C.3.8}$$

C.4 Contravariant representations

Now that we're familiar with the ϵ tensor, we should tie up a loose end from Section C.2.9. There we introduced the fundamental and conjugate representations of $SL(2,\mathbb{C})$. What happened to the **contravariant** representations that transform under the inverse matrices N^{-1} and N^{*-1}? For a general group, e.g. $GL(N,\mathbb{C})$, these are unique representations so that we have a total of four different reps for a given group.

It turns out that for $SL(2,\mathbb{C})$ these contravariant representations are equivalent (in the group theoretical sense) to the fundamental and conjugate representations presented above. Using the antisymmetric tensor $\epsilon_{\alpha\beta}$ ($\epsilon^{12} = 1$) and the unimodularity of $N \in SL(2,\mathbb{C})$,

$$\epsilon_{\alpha\beta}N^{\alpha}{}_{\gamma}N^{\beta}{}_{\delta} = \epsilon_{\gamma\delta}\det N \tag{C.4.1}$$

$$\epsilon_{\alpha\beta}N^{\alpha}{}_{\gamma}N^{\beta}{}_{\delta} = \epsilon_{\gamma\delta} \tag{C.4.2}$$

$$\left(N^T\right)_{\gamma}{}^{\alpha}\epsilon_{\alpha\beta}N^{\beta}{}_{\delta} = \epsilon_{\gamma\delta} \tag{C.4.3}$$

$$\epsilon_{\alpha\beta}N^{\beta}{}_{\delta} = \left[\left(N^T\right)^{-1}\right]_{\alpha}{}^{\gamma}\epsilon_{\gamma\delta} \tag{C.4.4}$$

And hence by Schur's Lemma N and $(N^T)^{-1}$ are equivalent. Similarly, N^* and $(N^\dagger)^{-1}$ are equivalent. This is not surprising, of course, since we already knew that the antisymmetric tensor, ϵ, is used to raise and lower indices in $SL(2,\mathbb{C})$. Thus the equivalence of these representations is no more 'surprising' than the fact that Lorentz vectors with upper indices are equivalent to Lorentz vectors with lower indices. Explicitly, then, the contravariant representations transform as

$$\psi'^{\alpha} = \psi^{\beta}(N^{-1})_{\beta}{}^{\alpha} \tag{C.4.5}$$

$$\overline{\chi}'^{\dot\alpha} = \overline{\chi}^{\dot\beta}(N^{*-1})_{\dot\beta}{}^{\dot\alpha}. \tag{C.4.6}$$

Let us summarize the different representations for $SL(2,\mathbb{C})$,

Representation	**Index Structure**	**Transformation**
Fundamental	Lower	$\psi'_{\alpha} = N_{\alpha}{}^{\beta}\psi_{\beta}$
Conjugate	Lower dotted	$\overline{\chi}'_{\dot\alpha} = (N^*)_{\dot\alpha}{}^{\dot\beta}\overline{\chi}_{\dot\beta}$
Contravariant	Upper	$\psi'^{\alpha} = \psi^{\beta}(N^{-1})_{\beta}{}^{\alpha}$
Contra-conjugate	Upper dotted	$\overline{\chi}'^{\dot\alpha} = \overline{\chi}^{\dot\beta}(N^{*-1})_{\dot\beta}{}^{\dot\alpha}$

Occasionally one will see the conjugate and contravariant-conjugate transformations written in terms of Hermitian conjugates,

$$\overline{\chi}'_{\dot\alpha} = \overline{\chi}_{\dot\beta}(N^\dagger)^{\dot\beta}{}_{\dot\alpha} \tag{C.4.7}$$

$$\overline{\chi}'^{\dot\alpha} = (N^{\dagger-1})^{\dot\alpha}{}_{\dot\beta}\overline{\chi}^{\dot\beta}. \tag{C.4.8}$$

We will not advocate this notation, however, since Hermitian conjugates are a bit delicate notationally in quantum field theories. For more details about the representations of $SL(2,\mathbb{C})$, see the appendix of Wess and Bagger.[11]

Example C.4.1 (Equivalent representations)**.** We stated before that the fundamental, conjugate, contravariant, and contravariant-conjugate representations of $GL(N,C)$ are generally not equivalent. We've seen that for $SL(2,\mathbb{C})$ this is not true, and there are only two unique representations. Another example is $U(N)$, for which $U^\dagger = U^{-1}$ and $U^{\dagger -1} = U$. Thus, unlike $SL(2,\mathbb{C})$ where the upper- and lower-index representations are equivalent, for $U(N)$ the dotted- and undotted-index representations are equivalent.

Note that there are different conventions for the index height, but the point is that the upper and lower index objects are equivalent due to the ϵ tensor. We can see this in a different way by looking at tensor structures.

C.5 Stars and Daggers

Let us take a pause from our main narrative to clarify some notation. When dealing with *classical* fields, the complex conjugate representation is the usual complex conjugate of the field; i.e. $\psi \to \psi^*$. When dealing with *quantum* fields, on the other hand, it is conventional to write a Hermitian conjugate; i.e. $\psi \to \psi^\dagger$. This is because the quantum field contains creation and annihilation *operators*. The Hermitian conjugate here is the quantum version of complex conjugation. (We'll explain this statement below.)

This is the same reason why Lagrangians are often written $\mathcal{L} =$ term + h.c. In classical physics, the Lagrangian is a scalar quantity so one would expect that one could have just written 'c.c.' (complex conjugate) rather than 'h.c.' (Hermitian conjugate). In QFT, however, the fields in the Lagrangian are operators that must be Hermitian conjugated. When taking a first general relativity course, some students develop a very bad habit: they think of lower-index objects as row vectors and upper-index objects as column vectors, so that

$$V_\mu W^\mu = \begin{pmatrix} V_0 & V_1 & V_2 & V_3 \end{pmatrix} \cdot \begin{pmatrix} W^0 \\ W^1 \\ W^2 \\ W^3 \end{pmatrix}. \tag{C.5.1}$$

Thus they think of the covariant vector as somehow a 'transpose' of contravariant vectors. This is is a bad, bad, *bad* habit and those students must pay their penance when they work with spinors. In addition to confusion generated from the antisymmetry of the metric and the anticommutation relations of the spinors, such students become confused when reading an

[11] J. Wess and J. Bagger. *Supersymmetry and supergravity.* Princeton, USA: Univ. Pr. (1992) 259 p.

expression like $\psi^\dagger_\alpha$ because they interpret this as

$$\psi^\dagger_\alpha \stackrel{?}{=} (\psi^*_\alpha)^T = (\overline{\psi}_{\dot\alpha})^T \stackrel{?}{=} \overline{\psi}^{\dot\alpha} \tag{C.5.2}$$

Wrong! Fail! Go directly to jail, do not pass go! The dagger (†) on the ψ acts *only* on the quantum operators in the field ψ, it doesn't know and doesn't care about the Lorentz index. Said once again, with emphasis: *There is no transpose in the quantum Hermitian conjugate!*

To be safe, one could always write the Hermitian conjugate since this is 'technically' always correct. The meaning, however, is not always clear. Hermitian conjugation is always defined with respect to an inner product. Anyone who shows you a Hermitian conjugate without an accompanying inner product might as well be selling you a used car with no engine.

In matrix quantum theory the inner product comes with the appropriate Hilbert space. This is what is usually assumed when you see a dagger in QFT. In quantum wave mechanics, on the other hand, the Hermitian conjugate is defined with respect to the L^2 inner product,

$$\langle f, g\rangle = \int dx\, f^*(x)\, g(x), \tag{C.5.3}$$

so that its action on fields is just complex conjugation. The structure of the inner product still manifests itself, though. Due to integration by parts, the Hermitian conjugate of the derivative is non-trivial,

$$\left(\frac{\partial}{\partial x}\right)^\dagger = -\frac{\partial}{\partial x}. \tag{C.5.4}$$

As you know very well we're really just looking at different aspects of the same physics.

> **Example C.5.1** (Not-so-obvious inner products)**.** It may seem like we're beating a dead undergrad with a stick with all this talk of the canonical inner product. However, in spaces with boundaries—such as the Randall-Sundrum model of a warped extra dimension—the choice of an inner product can be non-trivial. In order for certain operators to be Hermitian (e.g. so that physical masses are non-negative), functional inner products have to be modified. This changes the expansion of a function (e.g. a wavefunction) in terms of an 'orthonormal basis.'

The relevant inner product depends on the particular representation one is dealing with. In one-particle representations of symmetry, the inner product is the usual bracket for linear matrix[a] operators on the finite dimensional space of states in a supersymmetirc multiplet and we can think of daggers as usual for quantum mechanics operators (e.g. turning raising and lowering operators into each other). That gives some insight about the particle content in supersymmetry, but one doesn't really see the field theory come through until one works with 'superspace' upon wich 'superfields' propagate. In this case we'll use the appropriate generalization of the L^2 inner product for the infinite dimensional space of functions on superspace.

[a] We say 'matrix' because the one-particle state space is finite-dimensional so we could actually write our operators explicitly as matrices. This doesn't actually provide much insight and we won't do this, but the terminology is helpful to distinguish between the superspace picture where the operators are differential operators.

The operators in the superspace picture will no longer be matrix operators but differential operators which one must be careful keeping track of minus signs from integration-by-parts upon Hermitian conjugation.

It is worth making one further note about notation. Sometimes authors will write

$$\overline{\psi}_{\dot{\alpha}} = \psi^{\dagger}_{\alpha}. \tag{C.5.5}$$

This is technically correct, but it can be a bit misleading since one shouldn't get into the habit of thinking of the bar as some kind of operator. The bar and its dotted index are notation to distinguish the right-handed representation from the left-handed representation. The content of the above equation is the statement that the conjugate of a left-handed spinor transforms as a right-handed spinor.

If none of that made any sense, then you should take a deep breath, pretend you didn't read any of this, and continue to the next subsection. It be clear when we start making use of these ideas in proper context.

C.6 Tensor representations

Now that we've said a few things about raised/lowered and dotted/undotted indices, it's worth repeating the mantra of tensor representations of Lie groups: *symmetrize, antisymmetrize, and trace* (see, for example, Section 4.3 of Cheng and Li Ta-Pei Cheng and Ling-Fong Li. *Gauge Theory of Elementary Particle Physics*. Oxford University Press, 1988). Let's recall the familiar $SU(N)$ case. We can write down tensor representations by just writing out the appropriate indices, e.g. if $\psi^a \to U^a{}_b\psi^b$ and $\psi_a \to U^{\dagger}{}_a{}^b\psi_b$, then we can write an (n,m)-tensor Ψ and its transformation as

$$\Psi^{i_1\cdots i_m}{}_{j_1\cdots j_m} \to U^{i_1}{}_{i_1'}\cdots U^{i_n}{}_{i_n'}U^{\dagger}{}_{j_1}{}^{j_1'}\cdots U^{\dagger}{}_{j_m}{}^{j_m'}\Psi^{i_1'\cdots i_m'}{}_{j_1'\cdots j_m'}. \tag{C.6.1}$$

This, however, is not generally an irreducible representation. In order to find the irreps, we can make use of the fact that tensors of symmetrized/antisymmetrized indices don't mix under the matrix symmetry group. For $U(N)$,

$$\Psi^{ij} \to \Psi'^{ij} = U^i{}_kU^j{}_\ell\Psi^{k\ell} \tag{C.6.2}$$

$$\Psi^{ji} \to \Psi'^{ji} = U^j{}_\ell U^i{}_k\Psi^{\ell k} \tag{C.6.3}$$

$$= U^i{}_kU^j{}_\ell\Psi^{\ell k}, \tag{C.6.4}$$

thus if $P(i,j)$ is the operator that swaps the indices $i \leftrightarrow j$, then Ψ^{ij} and $\Psi^{ji} = P(i,j)\Psi^{ij}$ transform in the same way. In other words, $P(i,j)$ commutes with the matrices of $U(N)$, and hence we may construct simultaneous eigenstates of each. This means that the eigenstates of $P(i,j)$, i.e. symmetric and antisymmetric tensors, form invariant subspaces under $U(N)$. This argument is straightforwardly generalized to any matrix group and arbitrarily complicated index structures. Thus we may commit to memory an important lesson: we ought to symmetrize and antisymmetrize our tensor representations.

As with any good informercial, one can expect a "but wait, there's more!" deal to spice things up a little bit. Indeed, it turns our there are two more tricks we can invoke to further reduce our tensor reps. The first is taking the trace. For $U(N)$ this is somewhat obvious once it's suggested: we know from basic linear algebra that the trace is invariant under unitary rotations; it is properly a scalar quantity. What this amounts to for a general tensor is taking the contraction of an upper index i and lower index j with the Kronecker delta, δ_i^j. This is guaranteed to commute with the symmetry group because δ_i^j is invariant under $U(N)$. This is analogous to $\epsilon_{\alpha\beta}$ being an invariant tensor of $SL(2,\mathbb{C})$.

There's one more trick for $SU(N)$ (but not $U(N)$) which comes from another invariant tensor, $\epsilon_{i_1\cdots i_N}$. This is invariant under $SU(N)$ since

$$U_{i_1}^{i'_1}\cdots U_{i_N}^{i'_N}\epsilon_{i'_1\cdots i'_N} = \det U \epsilon_{i_1\cdots i_N} = \epsilon_{i_1\cdots i_N}. \tag{C.6.5}$$

Thus any time one has N antisymmetric indices of an $U(N)$ tensor, one can go ahead and drop them. Just like that. Note that this is *totally different* from the $\epsilon_{\alpha\beta}$ of $SL(2,\mathbb{C})$.

By now, the slightly more group theoretically savvy will have recognized the basic rules for $U(N)$ Young tableaux. This formalism further allows one to formulaically determine the dimension of a tensor representation and its decomposition into irreducible representations. Details for the $SU(N)$ Young tableaux are worked out in Cheng and Li Cheng and Li, *Gauge Theory of Elementary Particle Physics*, with a more general discussion in any self-respecting representation theory text.

The points that one should take away from this, however, isn't the formalism of Young tableaux, but rather the 'big picture' intuition of decomposing into irreducible tensors. In particular for $SL(2,\mathbb{C})$ the irreducible two-index ϵ tensor tells us that we can always reduce any tensorial representation into direct sums irreducible tensors which are symmetric in their dotted and (separately) undotted indices,

$$\Psi_{\alpha_1\cdots\alpha_{2n}\dot{\alpha}_1\cdots\dot{\alpha}_{2m}} = \Psi_{(\alpha_1\cdots\alpha_{2n})(\dot{\alpha}_1\cdots\dot{\alpha}_{2m})}. \tag{C.6.6}$$

We label such an irreducible tensor-of-spinor-indices with the $SO(3,1)$ notation (n,m). In this notation the fundamental left- (ψ) and right-handed ($\bar{\chi}$) spinors transform as ($\frac{1}{2}$,0) and (0,$\frac{1}{2}$) respectively. One might now want to consider how to reduce Poincaré tensor products following the analogous procedure that textbooks present for $SU(2)$. Recall that $SO(3) \cong SU(2)/\mathbb{Z}_2$ so that such an analogy amounts to 'promoting' $SO(3)$ to $SO(3,1)$. For further details see, e.g., Osborn's lecture notes.[12]

[12]Osborn, *Symmetries and Particle Physics*.

C.7 Lorentz-Invariant Spinor Products

Armed with a metric to raise and lower indices, we can also define the inner product of spinors as the contraction of upper and lower indices. Note that in order for inner products to be Lorentz-invariant, one cannot contract dotted and undotted indices.

There is a very nice short-hand that is commonly used in supersymmetry that allows us to drop contracted indices. Since it's important to distinguish between left- and right-handed Weyl spinors, we have to be careful that dropping indices doesn't introduce an ambiguity. This is why right-handed spinors are barred in addition to having dotted indices. Let us now define the contractions

$$\psi\chi \equiv \psi^\alpha \chi_\alpha \tag{C.7.1}$$
$$\overline{\psi}\overline{\chi} \equiv \overline{\psi}_{\dot{\alpha}}\overline{\chi}^{\dot{\alpha}}. \tag{C.7.2}$$

Note that the contractions are different for the left- and right-handed spinors. A useful mnemonic is to imagine swordsman who carries his/her sword along his waist. A right-handed swordsman would have his sword along his left leg so that he could easily unsheathe it by pulling up and to the right—just like the way the right-handed dotted Weyl spinor indices contract. Similarly, a left-handed swordsman would have his sword along his right leg so that he would unsheathe by pulling up and to the left.

This is a choice of convention such that

$$(\psi\chi)^\dagger \equiv (\psi^\alpha\chi_\alpha)^\dagger = \overline{\chi}_{\dot{\alpha}}\overline{\psi}^{\dot{\alpha}} \equiv \overline{\chi}\overline{\psi} = \overline{\psi}\overline{\chi}. \tag{C.7.3}$$

The second equality is worth explaining. Why is it that $(\psi^\alpha\chi_\alpha)^\dagger = \overline{\chi}_{\dot{\alpha}}\overline{\psi}^{\dot{\alpha}}$? Recall from that the Hermitian conjugation acts on the creation and annihilation operators in the quantum fields ψ and χ. The Hermitian conjugate of the product of two Hermitian operators AB is given by $B^\dagger A^\dagger$. The coefficients of these operators in the quantum fields are just c-numbers ('commuting' numbers), so the conjugate of $\psi^\alpha\chi_\alpha$ is $\left(\chi^\dagger\right)_{\dot{\alpha}}\left(\psi^\dagger\right)^{\dot{\alpha}}$.

Now let's get back to our contraction convention. Recall that quantum spinor fields are Grassmann, i.e. they anticommute. Thus we show that with our contraction convention, the order of the contracted fields don't matter:

$$\psi\chi = \psi^\alpha\chi_\alpha = -\psi_\alpha\chi^\alpha = \chi^\alpha\psi_\alpha = \chi\psi \tag{C.7.4}$$
$$\overline{\psi}\overline{\chi} = \overline{\psi}_{\dot{\alpha}}\overline{\chi}^{\dot{\alpha}} = -\overline{\psi}^{\dot{\alpha}}\overline{\chi}_{\dot{\alpha}} = \overline{\chi}_{\dot{\alpha}}\overline{\psi}^{\dot{\alpha}} = \overline{\chi}\overline{\psi}. \tag{C.7.5}$$

It is actually rather important that quantum spinors anticommute. If the ψ were *commuting* objects, then

$$\psi^2 = \psi\psi = \epsilon^{\alpha\beta}\psi_\beta\psi_\alpha = \psi_2\psi_1 - \psi_1\psi_2 = 0. \tag{C.7.6}$$

Thus we must have ψ such that

$$\psi_1\psi_2 = -\psi_2\psi_1, \tag{C.7.7}$$

i.e. the components of the Weyl spinor must be Grassmann. So one way of understanding why spinors are anticommuting is that metric that raises and lowers the indices is antisymmetric. (We know, of course, that from another perspective this anticommutativity comes from the quantum creation and annihilation operators.)

C.8 Vector-like Spinor Products

Notice that the Pauli matrices give a natural way to go between $SO(3,1)$ and $SL(2,\mathbb{C})$ indices. Using (C.2.38),

$$(x_\mu \sigma^\mu)_{\alpha\dot{\alpha}} \to N_\alpha{}^\beta (x_\nu \sigma^\nu)_{\beta\dot{\gamma}} N^*{}_{\dot{\alpha}}{}^{\dot{\gamma}} \tag{C.8.1}$$

$$= (\Lambda_\mu{}^\nu x_\nu) \sigma^\mu{}_{\alpha\dot{\alpha}}. \tag{C.8.2}$$

Then we have

$$(\sigma^\mu)_{\alpha\dot{\alpha}} = N_\alpha{}^\beta (\sigma^\nu)_{\beta\dot{\gamma}} (\Lambda^{-1})^\mu{}_\nu N^*{}_{\dot{\alpha}}{}^{\dot{\gamma}}. \tag{C.8.3}$$

One could, for example, swap between the vector and spinor indices by writing

$$V_\mu \to V_{\alpha\dot{\beta}} \equiv V_\mu (\sigma^\mu)_{\alpha\dot{\beta}}. \tag{C.8.4}$$

We can define a 'raised index' σ matrix,

$$(\overline{\sigma}^\mu)^{\dot{\alpha}\alpha} \equiv \epsilon^{\alpha\beta} \epsilon^{\dot{\alpha}\dot{\beta}} (\sigma^\mu)_{\beta\dot{\beta}} \tag{C.8.5}$$

$$= (\mathbb{1}, -\sigma). \tag{C.8.6}$$

Note the bar and the reversed order of the dotted and undotted indices. The bar is just notation to indicate the index structure, similarly to the bars on the right-handed spinors.

Now an important question: How do we understand the indices? Why do we know that the un-barred Pauli matrices have lower indices $\alpha\dot{\alpha}$ while the barred Pauli matrices have upper indices $\dot{\alpha}\alpha$? Clearly this allows us to maintain our convention about how indices contract, but some further checks might help clarify the matter. Let us go back to the matrix form of the Pauli matrices (C.2.32) and the upper-indices epsilon tensor (C.3.1). One may use $\epsilon = i\sigma^2$ and to directly verify that

$$\epsilon \overline{\sigma}_\mu = \sigma_\mu^T \epsilon, \tag{C.8.7}$$

and hence

$$\overline{\sigma}_\mu = \epsilon \sigma_\mu^T \epsilon^T. \tag{C.8.8}$$

Now let us write these in terms of dot-less indices—i.e. write all indices without dots, whether or not they ought to have dots—then we can restore the indices later to see how they turn out. To avoid confusion we'll write dot-less indices with lowercase Roman letters $\alpha, \beta, \gamma, \delta \to a, b, c, d$.

$$(\overline{\sigma}^\mu)^{ad} = \epsilon^{ab}(\sigma^{\mu T})_{bc}(\epsilon^T)^{cd} \tag{C.8.9}$$
$$= \epsilon^{ab}(\sigma^\mu)_{cb}\epsilon^{dc} \tag{C.8.10}$$
$$= \epsilon^{ab}\epsilon^{dc}(\sigma^\mu)_{cb}. \tag{C.8.11}$$

We already know what the dot structure of the σ^μ is, so we may go ahead and convert to the dotted/undotted lowercase Greek indices. Thus $c, b \to \gamma, \dot{\beta}$. Further, we know that the ϵs carry only one type of index, so that $a, d \to \dot{\alpha}, \delta$. Thus we see that the $\overline{\sigma}^\mu$ have a dotted-then-undotted index structure. A further consistency check comes from looking at the structure of the γ matrices as applied to the Dirac spinors formed using Weyl spinors with our index convention.

C.9 Generators of $SL(2,\mathbb{C})$

How do Lorentz transformations act on Weyl spinors? We should already have a hint from the generators of Lorentz transformations on Dirac spinors. (Go ahead and review this subsection of your favorite QFT textbook.) The objects that obey the Lorentz algebra, (C.1.5), and generate the desired transformations are given by the matrices,

$$(\sigma^{\mu\nu})_\alpha{}^\beta = \frac{i}{4}(\sigma^\mu\overline{\sigma}^\nu - \sigma^\nu\overline{\sigma}^\mu)_\alpha{}^\beta \tag{C.9.1}$$
$$(\overline{\sigma}^{\mu\nu})^{\dot{\alpha}}{}_{\dot{\beta}} = \frac{i}{4}(\overline{\sigma}^\mu\sigma^\nu - \overline{\sigma}^\nu\sigma^\mu)^{\dot{\alpha}}{}_{\dot{\beta}}. \tag{C.9.2}$$

It is important to note that these matrices are Hermitian. (Some sources will define these to be anti-Hermitian, see below.) The assignment of dotted and undotted indices are deliberate; they tell us which generator corresponds to the fundamental versus the conjugate representation. The choice of *which* one is fundamental versus conjugate, of course, is arbitrary. Thus the left and right-handed Weyl spinors transform as

$$\psi_\alpha \to \left(e^{-\frac{i}{2}\omega_{\mu\nu}\sigma^{\mu\nu}}\right)_\alpha{}^\beta \psi_\beta \tag{C.9.3}$$
$$\overline{\chi}^{\dot{\alpha}} \to \left(e^{-\frac{i}{2}\omega_{\mu\nu}\overline{\sigma}^{\mu\nu}}\right)^{\dot{\alpha}}{}_{\dot{\beta}} \overline{\chi}^{\dot{\beta}}. \tag{C.9.4}$$

Example C.9.1 (Factors of i)**.** Oh boy.

> "In the kingdom of the blind, the one-i'd man is king," D. Erasmus.

Some standard references (such as Bailin & Love[13]) define $\sigma^{\mu\nu}$ and $\overline{\sigma}^{\mu\nu}$ without the leading factor of i. Their spinor transformations thus also lack the $(-i)$ in the exponential so that the actual quantity being exponentiated is the same in either convention. As much as it pains me—and it will certainly pain *you* even more—to

> differ from 'standard notation', we will be adamant about our choice since only then do our Lorentz transformations take the usual format for exponentiating Lie algebras:
>
> $$e^{i\cdot(\mathbb{R}\text{ Transformation Parameter})\cdot(\text{Hermitian Generator})}. \tag{C.9.5}$$
>
> Thus propriety demands that we define *Hermitian* $\sigma^{\mu\nu}$ as in (C.9.2), even if this means having a slightly different algebra from Bailin & Love. This will be very important when using the Fierz identities

We can invoke the $SU(2)\times SU(2)$ representation (and we use that word *very* loosely) of the Lorentz group from (C.2.20) and (C.2.21) to write the left-handed $\sigma^{\mu\nu}$ generators as

$$J_i = \frac{1}{2}\epsilon_{ijk}\sigma_{jk} = \frac{1}{2}\sigma_i \tag{C.9.6}$$

$$K_i = \sigma_{0i} = -\frac{i}{2}\sigma_i. \tag{C.9.7}$$

One then finds

$$A_i = \frac{1}{2}(J_i + iK_i) = \frac{1}{2}\sigma_i \tag{C.9.8}$$

$$B_i = \frac{1}{2}(J_i - iK_i) = 0. \tag{C.9.9}$$

Thus the left-handed Weyl spinors ψ_α are $(\frac{1}{2},0)$ spinor representations Similarly, one finds that the right-handed Weyl spinors $\overline{\chi}^{\dot\alpha}$ are $(0,\frac{1}{2})$ spinor representations.

Alternately, we could have *guessed* the generators of the Lorentz group acting on Weyl spinors from the algebra of rotations and boosts in (C.2.22) – (C.2.24). With a modicum of cleverness one could have made the ansatz that the $\mathbf{J}$ and $\mathbf{K}$ are represented on Weyl spinors via (C.9.6) and (C.9.7). From this one could exponentiate to derive a finite Lorentz transformation,

$$\exp\left(\frac{i}{2}\sigma\cdot\theta \pm \frac{1}{2}\sigma\cdot\phi\right) = \exp\left(i\frac{1}{2}\sigma\cdot(\theta\mp i\phi)\right), \tag{C.9.10}$$

where the upper sign refers to left-handed spinors while the lower sign refers to right-handed spinors. θ and ϕ are the parameters of rotations and boosts, respectively. One can then calculate the values of σ^{0i} and σ^{ij} to confirm that they indeed match the above expression.

Alternately, we could have *guessed* the generators of the Lorentz group acting on Weyl spinors from the algebra of rotations and boosts in (C.2.22) – (C.2.24). With a modicum of cleverness one could have made the ansatz that the $\mathbf{J}$ and $\mathbf{K}$ are represented on Weyl spinors via (C.9.6) and (C.9.7). From this one could exponentiate to derive a finite Lorentz transformation,

$$\exp\left(\frac{i}{2}\sigma\cdot\theta \pm \frac{1}{2}\sigma\cdot\phi\right) = \exp\left(i\frac{1}{2}\sigma\cdot(\theta\mp i\phi)\right), \tag{C.9.11}$$

[13] D. Bailin and A. Love. "Supersymmetric gauge field theory and string theory". In: (). Bristol, UK: IOP (1994) 322 p. (Graduate student series in physics)

where the upper sign refers to left-handed spinors while the lower sign refers to right-handed spinors. θ and ϕ are the parameters of rotations and boosts, respectively. One can then calculate the values of σ^{0i} and σ^{ij} to confirm that they indeed match the above expression.

> **Example C.9.2. Representations of the Lorentz group.** It is worth making an aside about the representations of the Lorentz group since there is a nice point that is often glossed over during one's first QFT course; see, for example, subsection 3.2 of Peskin.[a] Dirac's trick for finding n-dimensional representations of the Lorentz generators was to postulate the existence of $n \times n$ matrices γ^μ that satisfy the Clifford algebra,
>
> $$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \cdot \mathbb{1}_{n\times n}. \tag{C.9.12}$$
>
> One can then use this to explicitly show that such matrices satisfy the algebra of the Lorentz group, (C.1.5). You already know that for $n = 4$ in the Weyl representation, this construction is explicitly reducible into the left- and right-handed Weyl representations and we see that we get the appropriate generators are indeed what we defined to be $\sigma^{\mu\nu}$ and $\overline{\sigma}^{\mu\nu}$. We could have constructed the $n = 2$ generators directly, but this would then lead to a factor of i in the definition of σ.
>
> [a] Michael Edward Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley (1995) 842 p.

C.10 Chirality

Now let's get back to a point of nomenclature. Why do we call them left- and right-handed spinors? The Dirac equation tells us[a]

$$p_\mu \sigma^\mu \psi = m\psi \tag{C.10.1}$$
$$p_\mu \overline{\sigma}^\mu \overline{\chi} = m\overline{\chi}. \tag{C.10.2}$$

[a] To be clear, there's some arbitrariness here. How do we know which 'Dirac equation' (i.e. with $\boldsymbol{\sigma}$ or $\overline{\boldsymbol{\sigma}}$) to apply to ψ (the fundamental rep) versus $\overline{\chi}$ (the conjugate rep)? This is convention, 'by the interchangeability of the fundamental and conjugate reps' and 'the interchangeability of $\boldsymbol{\sigma}$ and $\overline{\boldsymbol{\sigma}}$' if you wish. Once we have chosen the convention of (C.10.1), then (C.10.2) follows from Hermitian conjugation. In other words, once we've chosen that the fundamental representation goes with the '$\boldsymbol{\sigma}$' Dirac equation (C.10.1), we know that the conjugate representation goes with the '$\overline{\boldsymbol{\sigma}}$' Dirac equation (C.10.2). If you ever get confused, check the index structure of $\boldsymbol{\sigma}$ and $\overline{\boldsymbol{\sigma}}$ and make sure they are contracting honestly.

Equation (C.10.2) follows from (C.10.1) via Hermitian conjugation, as appropriate for the conjugate representation.

In the massless limit, then, $p^0 \to |\mathbf{p}|$ and hence

$$\left(\frac{\sigma\cdot\mathbf{p}}{|\mathbf{p}|}\psi\right) = \quad \psi \tag{C.10.3}$$
$$\left(\frac{\sigma\cdot\mathbf{p}}{|\mathbf{p}|}\overline{\chi}\right) = -\overline{\chi}. \tag{C.10.4}$$

We recognize the quantity in parenthesis as the helicity operator, so that ψ has helicity $+1$ (left-handed) and $\overline{\chi}$ has helicity -1 (right-handed). Non-zero masses complicate things, of course. In fact, they complicate things differently depending on whether the masses are Dirac or Majorana. We'll get to this in due course, but the point is that even though ψ and $\overline{\chi}$ are no longer helicity eigenstates, they are *chirality* eigenstates:

$$\gamma_5 \begin{pmatrix} \psi \\ 0 \end{pmatrix} = - \begin{pmatrix} \psi \\ 0 \end{pmatrix} \tag{C.10.5}$$

$$\gamma_5 \begin{pmatrix} 0 \\ \overline{\chi} \end{pmatrix} = \begin{pmatrix} 0 \\ \overline{\chi} \end{pmatrix}, \tag{C.10.6}$$

where we've put the Weyl spinors into four-component Dirac spinors in the usual way so that we may apply the chirality operator, γ_5.

Example C.10.1. Keeping the broad program in mind, let us take a moment to note that chirality will play an important role in whatever new physics we might find at the Terascale. The Standard Model is a chiral theory (e.g. q_L and q_R are in different gauge representations), so whatever Terascale completion supersedes it must also be chiral. This is no problem in SUSY where we may place chiral fields into different supermultiplets ('superfields'). In extra dimensions, however, we run into the problem that there is no chirality operator in five dimensions. This leads to a lot of subtlety in model-building. (Not to mention issues with anomaly cancellation.) By the way, it prudence requires that a diligent student should thoroughly understand the difference between chirality and helicity.

C.11 Fierz Rearrangement

Fierz identities are useful for rewriting spinor operators by swapping the way indices are contracted. For example,

$$(\chi\psi)(\chi\psi) = -\frac{1}{2}(\psi\psi)(\chi\chi). \tag{C.11.1}$$

One can understand these Fierz identities as an expression of the decomposition of tensor products in group theory. For example, we could consider the decomposition $(\frac{1}{2}, 0) \otimes (0, \frac{1}{2}) = (\frac{1}{2}, \frac{1}{2})$:

$$\psi_\alpha \overline{\chi}_{\dot{\alpha}} = \frac{1}{2}(\psi\sigma_\mu\overline{\chi})\sigma^\mu{}_{\alpha\dot{\alpha}}, \tag{C.11.2}$$

where, on the right-hand side, the object in the parenthesis is a vector in the same sense as (C.8.4). The factor of $\frac{1}{2}$ is, if you want, a Clebsch-Gordan coefficient.

Another example is the decomposition for $(\frac{1}{2}, 0) \otimes (\frac{1}{2}, 0) = (0, 0) + (1, 0)$:

$$\psi_\alpha \chi_\beta = \frac{1}{2}\epsilon_{\alpha\beta}(\psi\chi) + \frac{1}{2}(\sigma^{\mu\nu}\epsilon^T)_{\alpha\beta}(\psi\sigma_{\mu\nu}\chi). \tag{C.11.3}$$

Note that the $(1, 0)$ rep is the antisymmetric tensor representation. All higher dimensional representations can be obtained from products of spinors. Explicit calculations can be found in the lecture notes by Müller-Kirsten and Wiedemann.[14]

[14] H. J. W. Muller-Kirsten and A. Wiedemann. "SUPERSYMMETRY: AN INTRODUCTION WITH CONCEPTUAL AND CALCULATIONAL DETAILS". in: (). Print-86-0955 (KAISERSLAUTERN).

These identities are derived from completeness relations for the basis $\{\sigma^\mu\}$ of 2×2 complex matrices. Using (C.2.35), we may write a generic $SL(2, \mathbb{C})$ matrix $\mathbf{x}$ as

$$\mathbf{x} = x_\mu \sigma^\mu \tag{C.11.4}$$

$$= \frac{1}{2}\mathrm{Tr}(\mathbf{x})\sigma^0 + \frac{1}{2}\mathrm{Tr}(\mathbf{x}\sigma^i)\sigma^i \tag{C.11.5}$$

$$(\mathbf{x})_{ac} = \frac{1}{2}(\mathbf{x})_{zz}\delta_{ac} + \frac{1}{2}(\mathbf{x})_{yz}(\sigma^i)_{zy}(\sigma^i)_{ac}, \tag{C.11.6}$$

where we sum over repeated indices and we are using lower-case Roman indices to emphasize that this is a *matrix* relation. The above statement knows nothing about metrics or raised/lowered indices or representations of $SL(2, \mathbb{C})$; it's just a fact regarding the multiplication these particular 2×2 matrices together. We already know the canonical index structure of σ matrices so we can later make this statement '$SL(2, \mathbb{C})$-covariant.' First, though, we take $\mathbf{x}$ to be one of the canonical basis elements of the space of 2×2 matrices, $\mathbf{x} = \mathbf{e}_{bd}$ with index structure $(\mathbf{e}_{bd})_{ac} = \delta_{ab}\delta_{cd}$. Thus we may write (C.11.6) as

$$\delta_{ab}\delta_{cd} = \frac{1}{2}\delta_{zb}\delta_{zd}\delta_{ac} - \frac{1}{2}\delta_{yb}\delta_{zd}(\sigma^i)_{zy}(\sigma^i)_{ac} \tag{C.11.7}$$

$$= \frac{1}{2}\delta_{ac}\delta_{db} + \frac{1}{2}(\sigma^i)_{ac}(\sigma^i)_{db}. \tag{C.11.8}$$

Now remembering the *matrix* definition of $\overline{\sigma}$, i.e. (C.8.6) rather than (C.8.5), we may write

$$\delta_{ab}\delta_{cd} = \frac{1}{2}\delta_{ac}\delta_{db} - \frac{1}{2}(\sigma^i)_{ac}(\overline{\sigma}^i)_{db} \tag{C.11.9}$$

$$= \frac{1}{2}\delta_{ac}\delta_{db} + \frac{1}{2}(\sigma^i)_{ac}(\overline{\sigma}_i)_{db} \tag{C.11.10}$$

$$= \frac{1}{2}(\sigma^\mu)_{ac}(\overline{\sigma}_\mu)_{db}, \tag{C.11.11}$$

where we have now used the Minkowski space metric to lower the contracted index and combine both terms. Let's now promote this from a matrix equation to an $SL(2, \mathbb{C})$ covariant equation by raising and dotting indices according to the conventions for σ^μ and $\overline{\sigma}^\mu$. We shall also now use our conventional $SL(2, \mathbb{C})$ lower-case Greek indices. We find

$$\delta_\alpha{}^\beta \delta_{\dot\gamma}{}^{\dot\delta} = \frac{1}{2}(\sigma^\mu)_{\alpha\dot\gamma}(\overline{\sigma}_\mu)^{\dot\delta\dot\beta}. \tag{C.11.12}$$

Upon contracting indices with the appropriate spinors, this is precisely (C.11.2). This relation can now be used to generate further Fierz identities, as the eager student may check independently. It is useful to remember the general structure of this 'mother' Fierz identity. On the left-hand side we have two Kronecker δs, one with undotted indices and the other with dotted indices. As usual both δs have one upper and one lower index. On the right-hand side we have two Pauli 'four-matrices,' one barred and the other unbarred with their vector index contracted. The indices on the right-hand side is fixed by the index structure on the left-hand side, so this is actually rather trivial to write down. Just don't forget the factor of one-half. (As with many of the metric identities, one can easily obtain the overall normalization by tracing both sides.)

Example C.11.1 (How to properly write this Fierz Identity)**.** Let's make a rather pedantic point about notation. The form of the 'mother' Fierz identity in (C.11.12) is what one should use to derive further Fierz identities since all indices are explicit. However, if one wanted to write down a list of useful identities—for example as Bailin & Love[a] do in their appendix—one would instead write out the explicit spinors, i.e.

$$\psi_\alpha \overline{\chi}^{\dot{\gamma}} = \frac{1}{2}(\sigma^\mu \overline{\chi})_\alpha (\overline{\sigma}_\mu \psi)^{\dot{\gamma}}. \tag{C.11.13}$$

This is more useful than (C.11.12) since that equation still had a lot of freedom to re-order indices. For example, one could have written the left-hand side of (C.11.8) as any of

$$\delta_{\alpha\delta}\delta_{\beta\gamma} = \delta_{\delta\alpha}\delta_{\beta\gamma} = \delta_{\alpha\delta}\delta_{\gamma\beta} = \delta_{\delta\alpha}\delta_{\gamma\beta} = \delta_{\beta\gamma}\delta_{\alpha\delta} = \cdots \tag{C.11.14}$$

[a]Bailin and Love, "Supersymmetric gauge field theory and string theory".

A comprehensive list of Fierz identities can be found in Appendix A of Bailin & Love,[15] note the different convention for $\sigma^{\mu\nu}$. A very pedagogical exposition on deriving Fierz identities for Dirac spinors is presented in Nishi.[16]

C.12 Connection to Dirac Spinors

We would now like to explicitly connect the machinery of two-component Weyl spinors to the four-component Dirac spinors that we (unfortunately) teach our children.

Let us define

$$\gamma^\mu \equiv \begin{pmatrix} 0 & \sigma^\mu \\ \overline{\sigma}^\mu & 0 \end{pmatrix}. \tag{C.12.1}$$

This, one can check, gives us the Clifford algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \cdot \mathbb{1}. \tag{C.12.2}$$

We can further define the fifth γ-matrix, the four-dimensional chirality

[15]Bailin and Love, "Supersymmetric gauge field theory and string theory".

[16]C. C. Nishi. "Simple derivation of general Fierz-type identities". In: *American Journal of Physics* 73.12 (2005), pp. 1160–1163. DOI: 10.1119/1.2074087. URL: http://link.aip.org/link/?AJP/73/1160/1.

operator,

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \tag{C.12.3}$$

A **Dirac spinor** is defined, as mentioned above, as the direct sum of left- and right-handed Weyl spinors, $\Psi_D = \psi \oplus \overline{\chi}$, or

$$\Psi_D = \begin{pmatrix} \psi_\alpha \\ \overline{\chi}^{\dot{\alpha}} \end{pmatrix}. \tag{C.12.4}$$

The choice of having a lower undotted index and an upper dotted index is convention and comes from how we defined our spinor contractions. The generator of Lorentz transformations takes the form

$$\Sigma^{\mu\nu} = \begin{pmatrix} \sigma^{\mu\nu} & 0 \\ 0 & \overline{\sigma}^{\mu\nu} \end{pmatrix}, \tag{C.12.5}$$

with spinors transforming as

$$\Psi_D \to e^{-\frac{i}{2}\omega_{\mu\nu}\Sigma^{\mu\nu}}\Psi_D. \tag{C.12.6}$$

In our representation the action of the chirality operator is given by γ_5,

$$\gamma^5\Psi_D = \begin{pmatrix} -\psi_\alpha \\ \overline{\chi}^{\dot{\alpha}} \end{pmatrix}. \tag{C.12.7}$$

We can then define left- and right-handed projection operators,

$$P_{L,R} = \frac{1}{2}\left(\mathbb{1} \mp \gamma^5\right). \tag{C.12.8}$$

Using the standard notation, we shall define a barred *Dirac* spinor as $\overline{\Psi}_D \equiv \Psi^\dagger\gamma^0$. Note that this bar has nothing to do with the bar on a Weyl spinor. We can then define a charge conjugation matrix C via $C^{-1}\gamma^\mu C = -(\gamma^\mu)^T$ and the Dirac conjugate spinor ${\Psi_D}^c = C{\overline{\Psi}_D}^T$, or explicitly in our representation,

$${\Psi_D}^c = \begin{pmatrix} \chi_\alpha \\ \overline{\psi}^{\dot{\alpha}} \end{pmatrix}. \tag{C.12.9}$$

A **Majorana spinor** is defined to be a Dirac spinor that is its own conjugate, $\Psi_M = \Psi_M^c$. We can thus write a Majorana spinor in terms of a Weyl spinor,

$$\Psi_M = \begin{pmatrix} \psi_\alpha \\ \overline{\psi}^{\dot{\alpha}} \end{pmatrix}. \tag{C.12.10}$$

Here our notation is that $\overline{\psi} = \psi^\dagger$, i.e. we treat the bar as an operation acting on the Weyl spinor (a terrible idea, but we'll do it just for now). One can choose a basis of the γ matrices such that the Majorana spinors are manifestly real. Thus this is sometimes called the 'real representation' of a Weyl spinor.Note that a Majorana spinor contains exactly the same amount of information as a Weyl spinor. Some textbooks thus package the Weyl SUSY generators into Majorana Dirac spinors, eschewing the dotted and undotted indices.

It is worth noting that in four dimensions there are no Majorana-Weyl spinors. This, however, is a dimension-dependent statement A good treatment of this can be found in the appendix of Polchinksi volume II.[17]

C.13 Dots and Bars

It's worth emphasizing once more that the dots and bars are just book-keeping tools. Essentially they are a result of not having enough alphabets available to write different kinds of objects. The bars on Weyl spinors can be especially confusing for beginning supersymmetry students since one is tempted to associate them with the barred Dirac spinors, $\overline{\Psi} = \Psi^\dagger \gamma_0$. *Do not make this mistake.* Weyl and Dirac spinors are different objects. The bar on a Weyl spinor has *nothing* to do with the bar on a Dirac spinor. We see this explicitly when we construct Dirac spinors out of Weyl spinors (namely $\Psi = \psi \oplus \overline{\chi}$), but it's worth remembering because the notation can be misleading.

In principle ψ and $\overline{\psi}$ are totally different spinors in the same way that α and $\dot{\alpha}$ are totally different indices. Sometimes—as we have done above—we may also use the bar as an operation that converts an unbarred Weyl spinor into a barred Weyl spinor. That is to say that for a left-handed spinor ψ, we may define $\overline{\psi} = \psi^\dagger$. To avoid ambiguity it is customary—as we have also done—to write ψ for left-handed Weyl spinors, $\overline{\chi}$ for right-handed Weyl spinors, and $\overline{\psi}$ to for the right-handed Weyl spinor formed by taking the Hermitian conjugate of the left-handed spinor ψ.

To make things even trickier, much of the literature on extra dimensions use the convention that ψ and χ (unbarred) refer to left- and right-'chiral' *Dirac* spinors. Here 'chiral' means that they permit chiral zero modes, a non-trivial subtlety of extra dimensional models that we will get to in due course. For now we'll use the supersymmetric convention that ψ and $\overline{\chi}$ are left- and right-handed Weyl spinors.

[17] J. Polchinski. "String theory. Vol. 2: Superstring theory and beyond". In: (). Cambridge, UK: Univ. Pr. (1998) 531 p.

Index