DRAFT - WORK IN PROGRESS

# Quantum Physics from Symmetry and Effective Theory

Trevor Scheopner
trevor@scheopner.com

Please send any errors you notice or feedback for improvement to the author.

©2026 Trevor Scheopner.
All Rights Reserved.

If you got something out of this book and want to support me:




https://ko-fi.com/trevorscheopner

November 2, 2026

# Contents


**Why do we need it?** v

**1 Single Particle Quantum Mechanics** 1
- 1.1 Linear Algebra of Fields . . . 4
- 1.2 Schrödinger Equation . . . 7
  - 1.2.1 Hermitian vs. Self-Adjoint . . . 10
  - 1.2.2 Spectral Theorem . . . 11
  - 1.2.3 Adiabatic Invariants in Quantum Mechanics . . . 16
- 1.3 Wigner's Theorem . . . 19
- 1.4 Continuous Transformations in Quantum Mechanics . . . 24
- 1.5 Summary and Spread Coordinates . . . 29
  - 1.5.1 Operator Spreads . . . 35
  - 1.5.2 Classical Limit . . . 38
  - 1.5.3 Meaning of $\hbar$ . . . 40
  - 1.5.4 Rigged Hilbert Space . . . 42
  - 1.5.5 Phenomenological Born Rule . . . 43
- 1.6 Scalar Theory - Galilean Schrödinger Equation . . . 52
  - 1.6.1 Interaction with Electromagnetism . . . 55
  - 1.6.2 Gauge Transformation Geometry . . . 57
  - 1.6.3 Charge Conservation . . . 61
- 1.7 Angular Momentum . . . 73
  - 1.7.1 Representations of SU(2) . . . 74
  - 1.7.2 Spherical Harmonics . . . 79
- 1.8 Spinor Theory - Pauli Equation . . . 92
- 1.9 Scattering Theory . . . 102
  - 1.9.1 Interaction Picture . . . 103
  - 1.9.2 Dyson Series . . . 104
  - 1.9.3 Lippmann-Schwinger Equation . . . 105
  - 1.9.4 Transition Matrix . . . 107
  - 1.9.5 Transition Rates and Fermi's Master Equation . . . 109
  - 1.9.6 Analytic Properties of the S Matrix . . . 111
- 1.10 Another Big Bold Leap . . . 120
  - 1.10.1 Trouble with Multi-particle Statistics . . . 120
  - 1.10.2 Quantum Field Theory to the Rescue . . . 121
- 1.11 Summary and Additional Reading . . . 134

**2 Quantum Mechanics in General** 137
- 2.1 Why Quantum Mechanics? . . . 137
  - 2.1.1 Quantum Ex Nihilo . . . 137
  - 2.1.2 Why Density Matrices? . . . 143
  - 2.1.3 Entanglement . . . 147
  - 2.1.4 Canonical Structure of Density Matrices . . . 149
- 2.2 Classical Limit . . . 153
  - 2.2.1 Weyl Transform . . . 153

2.2.2 Classicality Conditions . . . 157
2.2.3 Groenewold's theorem . . . 160
2.2.4 Quantum Mechanics in Classical Phase Space . . . 163
2.3 Lindblad Equation . . . 168
2.4 Measurement Problem . . . 178
2.5 Other Views of Quantum Mechanics . . . 186
2.5.1 Copenhagen Interpretation . . . 187
2.5.2 Post-Schrödinger Dynamics . . . 190
2.5.3 Many Worlds Interpretation . . . 191
2.5.4 The Damned Cat . . . 193
2.5.5 Bell-type Theorems and Their Consequences . . . 195
2.6 Quantum Statistical Mechanics . . . 197
2.6.1 First Law of Thermodynamics . . . 198
2.6.2 Thermodynamic Equilibrium . . . 199
2.6.3 Entropy and the Second Law of Thermodynamics . . . 200
2.6.4 Planck's Law . . . 204
2.6.5 Bose-Einstein and Fermi-Dirac Statistics . . . 205
2.7 Summary and Additional Reading . . . 209

**3 Quantum Field Theory 211**
3.1 Representations of the Lorentz Group . . . 212
3.2 Spin-Statistics Theorem . . . 216
3.3 Path Integral . . . 216
3.3.1 Path Integral for Time Evolution . . . 216
3.3.2 Important Subtleties . . . 218
3.3.3 Path Integral for Vacuum Expectation Values . . . 220
3.3.4 Path Integral for the Quantum Harmonic Oscillator . . . 224
3.3.5 Fermionic Path Integral . . . 228
3.4 Scalar Theory - Klein-Gordon Equation . . . 230
3.5 Spinor Theory - Dirac Equation . . . 237
3.6 Scattering Theory . . . 259
3.6.1 Schwinger Dyson Equation . . . 259
3.6.2 LSZ Reduction Formula . . . 259
3.6.3 Feynman Diagrams . . . 259
3.6.4 Cross Sections and Decay Rates . . . 259
3.7 Regularization . . . 259
3.8 Renormalization . . . 259
3.9 Symmetries in Quantum Field Theory . . . 259
3.10 Summary and Additional Reading . . . 259

**4 Gauge Theory 260**
4.1 Quantum Electrodynamics in General . . . 260
4.1.1 Classical Limit . . . 260
4.1.2 Gauge Invariance . . . 260
4.1.3 Path Integral and LSZ Reduction for Photons . . . 260
4.1.4 Ward Identities . . . 260
4.2 Scalar Quantum Electrodynamics . . . 260
4.3 Spinor Quantum Electrodynamics . . . 260
4.4 Geometry of Nonabelian Gauge Theory . . . 260
4.4.1 Geometry of Gauge Transformations . . . 260
4.4.2 Curvature of the Gauge Connection . . . 271
4.4.3 Possible Lagrangians . . . 277
4.4.4 The Simplest Gauge Theories . . . 279
4.4.5 Quantum Chromodynamics . . . 286
4.4.6 The Weak Interaction . . . 287
4.5 Dynamics of Nonabelian Gauge Theory . . . 289

4.5.1 Path Integral and LSZ Reduction . . . 289
4.5.2 Feynman Rules and Renormalization . . . 289
4.5.3 Ward Identities . . . 289
4.5.4 Becchi-Rouet-Stora-Tyutin Formalism . . . 289
4.5.5 Wilson Loops and Confinement . . . 289
4.6 Anomalies . . . 289
4.6.1 Adler-Bell-Jackiw Anomaly . . . 289
4.6.2 Chiral Symmetry Breaking . . . 289
4.6.3 Gauge Anomalies . . . 289
4.7 Summary and Additional Reading . . . 289

**5 The Standard Model and Beyond 290**
5.1 Electroweak Symmetry Breaking and the Higgs Mechanism . . . 290
5.2 Generations of Matter and Mass Mixing Matrices . . . 303
5.3 Standard Model in full glory . . . 313
5.4 Standard Model: Hard and Soft Problems . . . 316
5.5 Standard Model Effective Field Theory . . . 317
5.6 Pati-Salam Model . . . 317
5.7 Dark Matter . . . 317
5.8 Supersymmetry . . . 317
5.9 Summary and Additional Reading . . . 317

**6 Quantum Gravity 318**
6.1 Quantum Field Theory in Curved Spacetime . . . 319
6.2 Naive Quantization . . . 319
6.3 Quadratic Gravity, Renormalizability, and Asymptotic Safety . . . 319
6.4 Loop Quantum Gravity . . . 319
6.5 String Theory . . . 319

**References 319**

# Why do we need it?

**The Failure of Classical Physics**

The mechanics of classical particles and fields cannot explain the experimental facts of our world. In our companion text, "Where Classical Physics Comes From" (henceforth "the Original" or "Orig."), we found that classical point particles of finite mass and charge do not have well-defined interactions with electromagnetism or gravity and that extended rigid bodies, which can have well-defined such interactions, are inconsistent with relativistic causality. Thus, we need a new theory of matter which is more dynamic, not squeezed into a point state and not of fixed shape. Such matter can be described by a classical field, and that will be how we get started with quantum mechanics. We also found in the Orig. that classical field theory and statistical mechanics are incompatible, causing the Ultraviolet Catastrophe. Classical field theories do not reach thermodynamic equilibrium at nonzero temperature. This and various related statistical problems will lead us from single particle quantum mechanics inexorably into our main topic: quantum field theory.

**Quantum Confusion**

For historical reasons, quantum mechanics is typically described in far more mysterious terms than other topics in physics without a fundamentally good reason. High quality, sane, and nonmystical explanations of scattered pieces of quantum mechanics exist in literature (and have existed for decades). This work hopes to collect those pieces in one place to build a coherent story of how the quantum world works. I hope to convince you that all of the important structures of quantum mechanics can be derived to the same degree that the corresponding structures of classical mechanics can. Once you understand a slight shift in thinking, it will just make so much sense. To summarize my perspective, I think quantum mechanics is just deterministic unitary time evolution all the way down, and that the apparent randomness of the Born rule and collapse are macroscopically emergent from statistical mechanics. I will base my case for this perspective on [1,7,36,38,54] and the theory of decoherence. I will explain how this perspective is fully consistent with the various quantum experiments (physical or thought), including Bell-type theorems. In this perspective, the Born rule and collapse can be derived properly for some specific quantum systems (see the references, and we will explore examples), but are very difficult to derive on fully general grounds. However, as we will show, they are essentially the only possibilities consistent with ideal (maximally correlated, fast, macroscopic) quantum measurement. In my view, this is analogous to how it is nigh impossible in classical (statistical) mechanics to properly derive from the Liouville equation n that a generic interacting macroscopic system evolves in time average toward the canonical ensemble, but because that ensemble is in many cases the only possible equilibrium allowed by symmetry and factorization properties, it must be the answer.

**Audience & Style**

We will maintain all of the conventions and notations of the Orig. Each section has accompanying homework questions. **It is essential for understanding this work that you at the very least solve the homework questions marked with ⋆ .** In undergrad, Kyoungchul Kong told me to really claim to have read a physics book you have to solve at least half of its homework questions; at the time I thought that was too strict, now I see it is quite wise.

This book is physics graduate level in general and mainly oriented toward theory. The required math background will mostly be vector calculus, linear algebra, differential equations, partial differential equations, and complex analysis. When more advanced mathematics is necessary, references for learning will be given. For students needing to refresh their knowledge of vector calculus, I suggest [13,14,58]; for linear algebra, [20,41]; for differential equations, [8,60]; for partial differential equations, [40]; and for complex analysis, [23] and [11].

This is a physics book. Mathematical deductions and derivations which culminate in "theorems" will not

be proven or stated as rigorously as they would be by a mathematician. Some derivations and conclusions may need to be rigorized to be satisfying to a mathematician. For example of the imperfections we will endure, we will typically work as if all unspecified functions we are working with are analytic or at least smooth. We will work with the Dirac delta distribution without a discussion of measure theory or a treatment of test functions. We will commute limits without proving if it is safe to do so (although we will comment on some places where it is not). We will only mention briefly the difference between Hermitian and self-adjoint and then neglect the distinction. We may refer to the "position basis" and "momentum basis" even though neither the position nor momentum "eigenstates" are in the Hilbert space of square integrable functions. I place a high value on rigorous mathematics and so all "theorems" here are intended to be understood as "proven at the standards of a physicist" by proofs which are "correct in spirit". That is, when the proofs given lack full rigor, there typically exist proofs which have a similar line of thinking starting from appropriately formalized hypotheses but with full rigor which prove the appropriately rigorized conclusion of the theorem. That rigorous conclusion will have the same meaning for all intents and purposes of this work. I have attempted to the best of my ability to give results essentially provable by rigorous mathematics formulated in a way which is more palatable to a typical physicist but which has not lost the heart of the underlying mathematics.

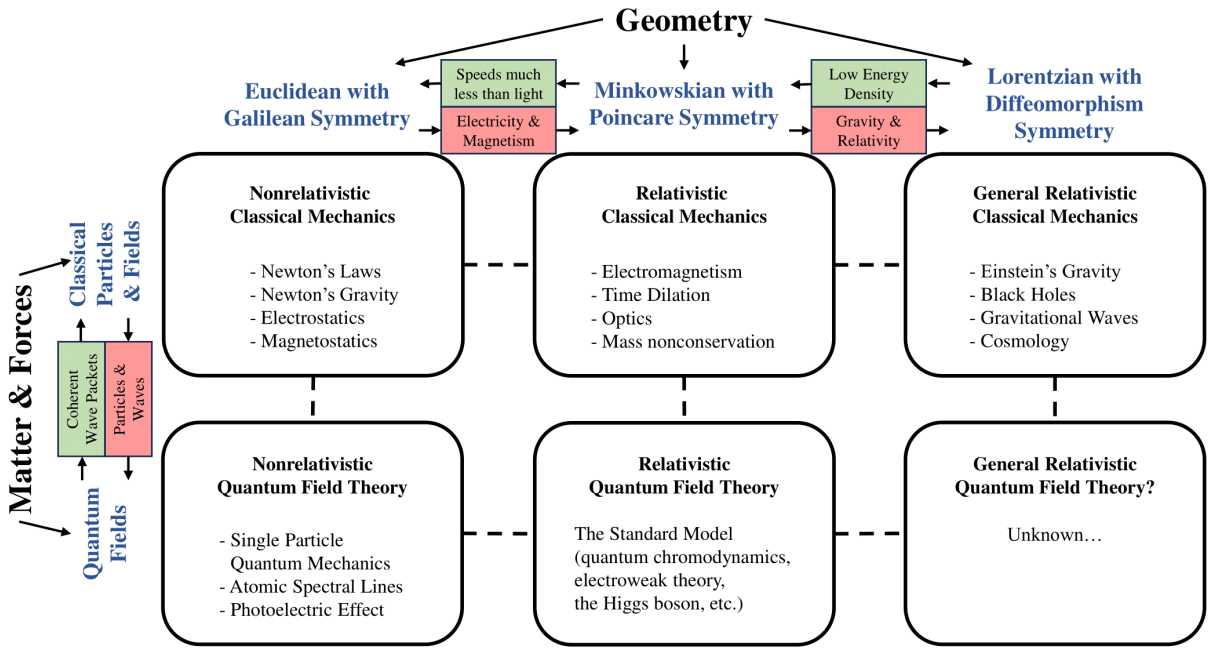



**Figure 1:** The map of the regimes of physical theories together with just some of the successes of each theory.

## Outline

This text serves to finish the explaining Figure 1 as began in the Orig. There we developed three different theories of the geometry of the universe: flat space with invariant time (Euclidean with Galilean Symmetry), flat spacetime with invariant speed of light (Minkowskian with Poincaré Symmetry, and curved spacetime with invariant speed of light (Lorentzian with Diffeomorphism Symmetry). We did so purely with classical degrees of freedom (interacting particles and waves). Here we will develop the second row to fix the problems of classical physics. In the resulting quantum theory, matter and all forces are different interacting modes of one big wave on the space of classical configurations. These two independent selections of a theory of geometry and theory of matter and forces lead to a grid of six different regimes of physics. Each regime but the bottom right corner has known issues with theoretical consistency and/or matching experiment which led to the discover of a new regime. For each regime I have selected a pair of concepts in red which are experimentally necessary but theoretically at odds within that regime; reconciling these pairs requires the discovery of the new regime it points to. Then, the new regime reduces approximately to the old regime when the circumstance shown in green is valid. The bottom right corner, which by the logic of the diagram would be our ultimate theory, is at this time not well understood theoretically and even once understood may very well have new problems. It is also presently inaccessible experimentally. The overwhelming majority of

ongoing physics research is done in the 5 "non-final" regimes, which each still contain exciting questions and applications, not in the final one, which presently is too far out experimentally to warrant saying much about with any confidence.

In Chapter 1 we will develop single particle quantum mechanics, in which a single matter particle is described by a wave-packet, not a moving point. This chapter derives the fundamental results of single particle quantum mechanics (and its interactions with electromagnetism) from basic definitions and symmetries. For further exposition see [2,38,51]. As described, single particle quantum mechanics is just a classical field theory. Quantum mechanics becomes deeper and more interesting when we start to think about the properties of a multiparticle system. In multiparticle quantum mechanics the system is described by a single wave on the space of all possible classical configurations of the particles, not by an independent wave in physical space for each particle. This causes the statistical properties of quantum mechanics to be significantly different from classical field theory. In order to resolve that statistical difference, we will develop quantum field theory and we will find that that theory finally resolves the tension between statistical mechanics and classical fields. In quantum field theory the entire universe is a single large quantum system.

Chapter 2 will derive the general structure of quantum mechanics and its common tools, such as density matrices, the von Neumann equation, the difference between classical and quantum symmetries, the classical limit, the theory of measurement, and the basic structure of quantum statistical mechanics. This chapter establishes all of the most important formalism for going further with quantum mechanics in general.

In Chapter 3 we will take all of our general quantum mechanical tools to properly understand the basics of quantum field theory. This chapter will develop the field representations of the Lorentz group, spin-statistics theorem, the path integral, free field theory, and a few basic interacting field theories. These interacting theories will serve as guiding examples to teach us how to do quantum field theory calculations more practically, introducing renormalization and how to extract scattering observables from the path integral. From this point [42,59] are very helpful supplements.

Chapter 4 will develop gauge theory, which is a general formulation of all three forces that are relevant in particle physics, electromagnetism, the nuclear weak force, and the nuclear strong force. This will bring us to a wonderful geometric view of all interactions while also giving us enormous power to do real calculations of real observable processes in particle colliders.

Chapter 5 will combine quantum electrodynamics, quantum chromodynamics, and further fields (neutrinos, the Higgs, etc.) to form the Standard Model of particle physics, the most successful theory of physics of all time. As a quantum field theory, the Standard Model is able to explain almost every phenomenon we have observed in nature. We will also discuss briefly some proposed physics beyond the Standard Model and discuss what the future of particle physics might hold.

Unfortunately, one of the most glaring failures of the Standard Model is its lack of a theory of gravity. Chapter 6 will discuss why it is so difficult to adapt the Standard Model to include a theory of gravity and discuss what are considered future prospects for how such a thing might eventually be done.

**Acknowledgments**

I began this book in the first year of my PhD at UCLA, because I was confused about physics. I was uncomfortable with the basics and it took me years to disentangle my many misconceptions. Over the time I have been writing I have been supported by many friends, colleagues, students, and mentors without whom this book would have fizzled well before completion. I would like to thank my friends and family for their persistent support of me in developing this book over the several years that it took, in particular my parents Travis and Tameka. I thank my friends and colleagues with whom I have had great discussions that have aided in my writing and whose ideas have become inextricably entangled with my own. Specifically, Juan Pablo Gatica, Daniels Krimans, Richard Myers, Thomas Dellaert, Hassan Farhat, Gabrielle Guttormsen, Justin Vines, Sven Harder, Will Liu, Drew Parks, Anna Wolz, Felix Lichtner, and Francisco Blanco. I would like to thank the students of the 2019-2020 UCLA Honors Physics series, who greatly shaped my growth as an educator, and in the teaching of whom the writing of this book began. I further thank the UCLA graduate student cohorts of 2020 and 2021 who were the first proper test subjects of this book's prototype, and whose feedback greatly guided its development. This book would not have been possible without the tolerance for it and encouragement of its importance that I was given by my supervisors Jan Steinhoff and Zvi Bern. My physics education was fundamentally transformed by John P. Ralston, and it took years for the full depth of his lessons to set it. The underlying philosophy and content of this work are strongly inspired by conversations with him. No generative AI was used in the creation of this book. Use of any part of this work for training large language models or any generative AI tool is expressly prohibited.

*Quantum Mechanics can be understood, and it can be understood by you.*

# Chapter 1

# Single Particle Quantum Mechanics

*If you feel a function $\psi(t, \vec{x})$ is abstract, imagine a whole system full of jello - beautiful green jello of the best transparency - and $\psi(t, \vec{x})$ representing some disturbance of that quantum jello from equilibrium. No position $\vec{x}$ moves. It is the jello that jiggles, shakes, and twists, always and forever. It's beautiful!*

– John P. Ralston [47]

*A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die.*

– Max Planck [43]

*If it's not in your equations, it's not in your theory.*

Let's fix the point particle problem. We found in the Orig. that interacting point particles tend to have infinite energy, do not interact consistently with electromagnetism, and experimentally do not produce the observed macroscopic properties of matter. The most obvious solution is to try particles with a non-zero size. However, if the shape and size of a body are rigid then when one part of it moves the rest must move correspondingly. Hence, a force applied at one end of a rigid body causes the other end to move instantaneously, communicating faster than light. As such, exactly rigid bodies are inconsistent with causality in relativistic theories. Therefore neither point particles nor rigid bodies are ultimately viable for fundamental physics. The size and shape of matter must themselves be dynamical. We need a description about how much matter is at each point in space which is allowed to change in time. That is, there will be some field(s) $Q(t, \vec{x})$ which ultimately give the amount and configuration of matter at the location $\vec{x}$ and time $t$, a classical field theory of matter. Of course, given the success of the point theory of matter under certain conditions, we also need our field theory of matter to reduce to the point theory in those conditions. While when we modeled a unit of matter as being a point particle we ended up with classical mechanics, when we model a unit of matter as being a wave the theory we end up with is called **single particle Quantum Mechanics**. That name is a probably a bit confusing, but over the course of this and the next chapter it will come to make sense.

Before continuing onto the details, we must take a moment to dispel some common misconceptions. In this chapter we cover single particle quantum mechanics, meaning we will describe a single blob of matter by a classical field (and will have it interact with stuff we already know about, like electromagnetism). Quantum mechanics is more general and powerful than covered this chapter; it is the focus of this entire volume. Some readers with prior studies are likely to object to thinking about quantum mechanics like a classical field theory at all, but when considering a single particle it literally is so. They are different when considering multiple interacting particles; we will get there next chapter. Less experienced readers are likely to be very confused that linguistically *quantum particle* is corresponding to a *classical matter wave*. The world particle is used because when our classical matter wave is only in a small region of space compared to all experimentally

relevant length scales and all parts of the wave are moving at about the same velocity, the center of mass and total momentum of that classical matter wave approximately satisfy the classical particle equations of motion (we saw the first sign of this fact in the Original in problem 4.48 and will prove it more generally in section 1.5). Under those circumstances that wave is said to be "in the classical limit" and the relevant physical consequences are generally accurately described by classical point-particle theories. Remember that particle in general does **NOT** mean point-particle. It means the constituents/ingredients of our physical description, not a moving point in space. A point-particle is a specific model of matter particles which has for more than a century now been known to be **WRONG** at the level of fundamental physics.

What does the name quantum theory mean? The full definition will come next chapter, but for now we can take it to mean anything described by the Schrödinger equation (or the related von Neumann equation) introduced in section 1.2. That equation is *enormously* generic. Quantum theory translated into less magical sounding words would just be called "generic Hamiltonian linear wave theory". Readers are likely to encounter some of the following claims about quantum mechanics:

1. Electrons are point particles which instantly jump between discrete atomic orbits (the Bohr model).

2. Everything in quantum mechanics is discrete. The action of any quantum system can only ever be a half-integer multiple of Planck's constant $\hbar$ (the Bohr-Sommerfeld theory).

3. Due to the uncertainty principle[1] it is impossible to know both the position and momentum of a particle at once. Knowing one makes knowing the other impossible.

4. An electron is both a wave and a point-particle at the same time except when you don't look at it because then it is neither and except when you do look at it because then it chooses to be one or the other (wave-particle duality as it might be described in a popular science magazine).

5. When an observable is measured, the system instantaneously, fundamentally randomly, and in a way which cannot be explained collapses to an eigenstate of the operator representing that observable and the experiment records the corresponding exact eigenvalue of that eigenstate as the measured value (naive collapse postulate).

6. *Nobody really understands quantum mechanics.* – Richard Feynman [18].

These are a few among many similar common paradoxical statements about quantum mechanics. Each of these statements is somewhere between misleading and completely wrong. They are based on an incomplete understanding of both the math and physics of quantum mechanics from a century ago. These have long since been explained by much clearer and more reasonable thinking. That thinking has so far failed to propagate into lower level courses, unfortunately, or into the minds of people who just "shut up and calculate". The era from 1900-1925 was bizarre. Problems with classical physics were known but the actual equation defining quantum mechanics (the Schrödinger equation) was not discovered until 1925. For 25 years many physicists said a great deal of wrong, or at best deeply confused, things now called "old quantum theory". **Old quantum theory is NOT the basis of modern quantum mechanics** at all and learning it is mostly actively detrimental to your physics education. While it may be interesting from a historical perspective, it is not helpful for understanding quantum mechanics. Because science is iterative and gradual and it takes time to iron out inconsistencies and change minds, many scientists continued to think about and teach quantum mechanics using old quantum theory terms while presenting actual quantum mechanical equations even well after 1925. Actual quantum mechanics contradicts old quantum theory in countless fundamental ways and old quantum theory was never a coherent theory to begin with, just a loose collection of claims. The tension between those claims has caused enormous confusion for many years among students learning quantum mechanics. This confusion is rooted in history, not physics, and is not necessary for you to experience. I will briefly[2] respond to the above misconceptions, with the fullness of detail to come from

[1] Readers familiar with quantum mechanics are likely to have been taught to use the name "Heisenberg uncertainty principle", named after Werner Heisenberg, who was the scientific head of the Nazi nuclear program. While that program failed, his willing participation and persistent nationalism during the Nazi era make him, in the author's view, ineligible for veneration, rehabilitation, or apologism.

[2] Here I am a victim of Brandolini's law: the amount of energy needed to refute bullshit is an order of magnitude bigger than

reading the rest of this book.

1. In 1925 the Schrödinger equation completely replaced the "old quantum theory" Bohr model of the atom. While the Schrödinger equation does have resonances (specific frequencies it responds very strongly to), it contains no instant jumping nor a discrete response to external sources. The Schrödinger equation is fully consistent with experiment and produces atomic solutions that are smooth in space and time. The resonances of the Schrödinger equation lead to macroscopic behavior which is jump-like on macroscopic time scales but fully continuous on microscopic time scales.

2. Space, time, the action, energy, the fields, the observables, nearly everything in quantum mechanics is continuous. Quantum mechanics is actually *more* continuous than classical mechanics, because it is about waves, not discrete point particles. Some systems have discretely separated eigenvalues, but that does not correspond to having discrete physical properties. For some simple integrable systems those eigenvalues can be calculated approximately using a mathematically souped-up Bohr-Sommerfeld method that explains why that *approximation* ever gave any useful results and why usually it doesn't. You should completely drop the old quantum theory belief that everything is discrete.

3. The uncertainty principle is not about measurement. The unqualified "position" of a body of finite size is not uniquely defined. Nobody believes it is a deep principle that you cannot know the exact definite position of a coffee cup. It is exactly as true as saying you cannot know the exact position of an electron. There are infinitely many points on a coffee cup I could call the position; because it is spread out in space it has no definite position. However, the center of mass position of the coffee cup is perfectly well defined and can be known arbitrarily well. Similarly, while there are infinitely many points you can call the position of a wave, the center of mass position of that wave (such as an electron) is perfectly and exactly well defined and can be known arbitrarily well. There is no fundamental tension between defining the center of mass position of a wave to any given precision and defining the total linear momentum carried by that wave to any given precision.

   When stated properly, the uncertainty principle is actually just a theorem of mathematics (proven in section 1.5). A wave cannot be both a Dirac delta $\delta(x)$ (meaning all located at one place) and a plane wave (meaning consisting entirely of one wavenumber) $e^{ikx}$ at the same time. This is as obvious as it sounds. One can do a bit better and show that the closer to being a Dirac delta a wave is, the further it is from being a plane wave, and vice versa; there is a precise bound on how close you can be to one for a given closeness to the other. There is a purely mathematical a limit to how scrunched up a wave can be in space for a given range of wavenumbers it is allowed to be a linear combination of.

   The uncertainty theorem only connects with physics in the classical limit (the limit that a quantum system approximately behaves like a simpler classical system). Under circumstances where a wave packet behaves approximately like a localized particle, how spread out a wave is in space (call the characteristic spatial width $\sigma_x$) has to be much smaller than our experimental length resolution $w_x$. (Otherwise you would see the wave having obvious finite size, so not approximately point-like.) How spread out a wave is in Fourier space (call the characteristic Fourier width $\sigma_k$) tells you about relative velocities between different parts of the wave, because as we found in Orig. problem 4.43 the wavenumber can be used to compute a wave's group velocity. When a wave contains many wavenumbers, its different parts will move at different speeds, dispersing it in space. If the classical limit is to be a valid approximation for a given experiment, the wave packet needs to remain approximately spatially localized and so we need it to also be approximately localized in Fourier space to limit its dispersion. Therefore, the classical limit requires that both $\sigma_x$ and $\sigma_k$ are small with respect to our experimental resolutions. The uncertainty theorem gives a limit on how small $\sigma_x$ and $\sigma_k$ can simultaneously be, and thus sets a limit to how good of an approximation the classical theory can be to the full quantum theory.

4. The above explanation hopefully helps you develop the intuition that an electron is always described by a wave equation and that only under the circumstances of the classical limit can we approximate that wave by a localized classical particle. We still tend to call the electron a particle, in the sense that it is a basic constituent of the theory, but it is never a point-particle. This real and self-consistent statement of the wave-particle duality is sometimes called the "opticomechanical analogy" and it is just

that needed to produce it.

the observation we made in Orig. problem 4.48 that under suitable approximations the center of a wave packet moves around like a classical particle.

5. Effectively random instantaneous collapse is a macroscopic approximate phenomenological description of the results of some measurement devices. In a real measurement process the system of interest interacts with a measurement device. Because quantum systems are very small compared to real experiments, which are always done in a thermal background of many degrees of freedom in which statistical mechanics is relevant, measurement of individual events in quantum mechanics has to be treated with statistical mechanics. The full system can be treated with continuous deterministic quantum time evolution, will the statistical properties Born rule can emerge dynamically for the measured system from this description. Collapse to an eigenstate is a situationally appropriate approximate model of some measurements after a finite amount of time. This will be explained in sections 1.5.5 and 2.4, based on [1, 7, 36, 38, 54].

6. Feynman's quote is from 1964. Maybe when he said it almost everyone was confused. Sixty years of science have been done since then and his view, given when my grandparents were children, does not represent the present. Many essential technologies of the modern world work only because of how incredibly well we understand quantum mechanics. Do not fall victim to magical thinking. The day you do, you have turned from a scientist into a priest. *This is the* $21^{st}$ *century. This is the century when everybody must understand quantum mechanics.* – John P. Ralston [46].

## 1.1 Linear Algebra of Fields

Call the field we will use to describe matter $Q(t, \vec{x})$ and its conjugate momentum field $P(t, \vec{x})$. We don't really know at this point whether $Q$ should be single number valued (like the scalar potential), spatial vector valued (like the vector potential), or something else. So, for now we will proceed allowing $Q$ to have an arbitrary unknown number of components. Still then, $Q(t, \vec{x})$ at a fixed point in time is a vector valued function of position (in the math sense, meaning an element of some vector space, not in the sense of a vector in space or spacetime). Let us call the vector space that $Q$ at a particular $(t, \vec{x})$ lives in $\mathbb{H}_{\text{int}}(t, \vec{x})$ so that $Q(t, \vec{x}) : \mathbb{R} \times \mathbb{R}^d \to \mathbb{H}_{\text{int}}$ (in $d$ spatial dimensions). This $\mathbb{H}_{\text{int}}$ is called the "internal space" or space of internal degrees of freedom. We notate $\mathbb{H}_{\text{int}}(t, \vec{x})$ with the $(t, \vec{x})$ simply to indicate that because $Q$ is defined at each $(t, \vec{x})$, there is a different "copy" of the vector space $\mathbb{H}_{\text{int}}$ at each $(t, \vec{x})$. The set of vector valued functions of $\vec{x}$ with addition and multiplication by constants is closed under such operations and so itself a vector space which we will call $\mathbb{H}(t)$. That is, $\mathbb{H}_{\text{int}}$ is the space of possible values of $Q(t, \vec{x})$ at a particular $(t, \vec{x})$ whereas $\mathbb{H}(t)$ is the space of all configurations of the function $Q(t, \cdot)$ at a particular time.

Recall the defining feature of the Dirac delta:

$$f(\vec{x}) = \int_{\mathbb{S}} f(\vec{x}')\delta(\vec{x} - \vec{x}')d^d\vec{x}'. \tag{1.1.1}$$

That is, any function $f$ can be written as a linear combination (integral) over itself times the Dirac delta. If we view the function $f(\vec{x})$ as a vector in the vector space of (locally integrable) functions, then (1.1.1) says that any (locally integrable) function can be written as a linear combination of deltas with different centers (the center of a delta is the location of its peak) with appropriate coefficients (those coefficients being the value of the function in question at the center of the corresponding delta). As such, this means that the deltas form a basis of the vector space of (locally integrable) functions in space. We call the vector space of functions we are considering $\mathbb{H}_{\mathbb{S}}$ where $\mathbb{S}$ is for space. This motivates introducing so-called Dirac notation to make the linear algebra here more apparent. In Dirac notation, we write $|\vec{x}\rangle$ to mean a Dirac delta with center at location $\vec{x}$, a specific point in space, and $|f\rangle$ to mean a specific element $f(\vec{x})$ in the vector space of functions. Rewriting (1.1.1) in this notation:

$$|f\rangle = \int_{\mathbb{S}} f(\vec{x})|\vec{x}\rangle d^d\vec{x}. \tag{1.1.2}$$

This simply makes the thinking more clear: $|f\rangle$ is a linear combination of the basis vectors $|\vec{x}\rangle$ and the coefficients in the expansion are $f(\vec{x})$. $\vec{x}$ is like an index which is summed over, just like when we first developed classical field theory. This notation is also very useful for writing inner products. For two real

vectors $\vec{u}$ and $\vec{v}$ with components $u^a$ and $v^a$ in an orthonormal basis, the inner product is defined as:

$$\vec{u} \cdot \vec{v} = \sum_{a=1}^{D} u^a v^a. \tag{1.1.3}$$

This allows one to define $\vec{u}^T \stackrel{\text{def}}{=} \vec{u}\cdot$ as the "row vector", "dual vector", "covector", or "linear functional" (depending on your favorite flavor of name) as the object which when matrix multiplied onto $\vec{v}$ returns the inner product:

$$\vec{u}^T \vec{v} \stackrel{\text{def}}{=} \vec{u} \cdot \vec{v}. \tag{1.1.4}$$

In the Dirac notation, the dual vector to $|f\rangle$ is written as $\langle f|$. Because our vector space now has a continuous basis, the sum in (1.1.3) becomes continuous (and so is an integral). Hence the inner product on the vector space of real valued functions is defined as:

$$\langle f|g\rangle \stackrel{\text{def}}{=} \int_{\mathbb{S}} f(\vec{x})g(\vec{x})d^d\vec{x}. \tag{1.1.5}$$

For complex vectors, the mathematically convenient definition of the inner product is slightly different. The complex inner product in an orthonormal basis is defined by:

$$\vec{u} \cdot \vec{v} \stackrel{\text{def}}{=} \sum_{a=1}^{D} u_a^* v_a. \tag{1.1.6}$$

The space of complex numbers is equivalent to the plane of ordered pairs of real numbers (identifying the real and imaginary parts as the components of the ordered pair) along with a multiplication rule (the definition of complex multiplication). As such, we can think of a complex vector of complex dimension $D$ as a real vector of dimension $2D$. We can see why this definition of the complex inner product is natural by looking at what it does to two 1 dimensional complex numbers $a+ib$ and $c+id$:

$$(a+ib)\cdot(c+id) = (a-ib)(c+id) = ac+bd+i(ad-bc) = (a,b)\cdot(c,d) + i\det\begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{1.1.7}$$

Thus, the real part of the complex inner product matches the real inner product of the real vectors formed from the complex numbers and the imaginary part matches the signed area of the parallelogram spanned by those real vectors. Both the real and imaginary parts of the complex inner product carry geometrically meaningful information about the vectors!

For complex valued functions $f$ and $g$ then we define the inner product by:

$$\langle f|g\rangle \stackrel{\text{def}}{=} \int_{\mathbb{S}} f^*(\vec{x})g(\vec{x})d^d\vec{x}. \tag{1.1.8}$$

Following from the definition of the inner product and $|\vec{x}\rangle$:

$$\langle \vec{x}|\vec{x}'\rangle = \int_{\mathbb{S}} \delta(\vec{x}-\vec{x}'')\delta(\vec{x}'-\vec{x}'')d^d\vec{x}'' = \delta(\vec{x}-\vec{x}') \tag{1.1.9}$$

$$\langle \vec{x}|f\rangle = \langle \vec{x}|\int_{\mathbb{S}} f(\vec{x}')|\vec{x}'\rangle d^d\vec{x}' = \int_{\mathbb{S}} f(\vec{x}')\langle\vec{x}|\vec{x}'\rangle d^d\vec{x}' = \int_{\mathbb{S}} f(\vec{x}')\delta(\vec{x}-\vec{x}')d^d\vec{x}' = f(\vec{x}) \tag{1.1.10}$$

$$\int_{\mathbb{S}} |\vec{x}\rangle\langle\vec{x}|d^d\vec{x}|f\rangle = \int_{\mathbb{S}} |\vec{x}\rangle\langle\vec{x}|f\rangle d^d\vec{x} = \int_{\mathbb{S}} |\vec{x}\rangle f(\vec{x})d^d\vec{x} = |f\rangle. \tag{1.1.11}$$

Equation (1.1.9) is what we mean by an orthonormal basis on our vector space of functions $\mathbb{H}_{\mathbb{S}}$. Equation (1.1.10) gives a way to think about the explicit function $f(\vec{x})$ just as the inner product $\langle\vec{x}|f\rangle$. Equation (1.1.11) shows that acting $\int_{\mathbb{S}} |\vec{x}\rangle\langle\vec{x}|d^d\vec{x}$ on any element of $\mathbb{H}_{\mathbb{S}}$ space returns that element. Hence, it is the identity operator on $\mathbb{H}_{\mathbb{S}}$:

$$\int_{\mathbb{S}} |\vec{x}\rangle\langle\vec{x}|d^d\vec{x} = \mathbb{1}. \tag{1.1.12}$$

This is called the completeness relation.

We let $|Q(t)\rangle \in \mathbb{H}(t)$ be our field configuration at a specific time. When we look at that field configuration at a specific $\vec{x}$ (which in this linear algebraic perspective is done be considering its coefficient on a Dirac delta centered at that $\vec{x}$), the result is in arbitrary $Q(t,\vec{x}) \in \mathbb{H}_{\text{int}}$. This vector in $\mathbb{H}_{\text{int}}$ can be decomposed into its components in a basis of that space. An arbitrary $|Q(t)\rangle \in \mathbb{H}(t)$ thus is a linear combination of a product of basis elements of $\mathbb{H}_{\mathbb{S}}$ (Dirac deltas in space) with basis elements of $\mathbb{H}_{\text{int}}$. This is precisely the structure of a

tensor product! So, $\mathbb{H} = \mathbb{H}_{\mathbb{S}} \otimes \mathbb{H}_{\text{int}}$. At this point it is a good idea to revisit Orig. section 4.7.1 and Orig. problem 4.74 to better understand tensor products.

When the field $Q$ is real and has multiple components, $\langle Q|$ is still defined as a vector in the dual vector space so that the inner product matches the existing definition:

$$\langle Q| \text{ defined so that for all } |Q'\rangle : \langle Q|Q'\rangle = \int_{\mathbb{S}} Q^T(\vec{x})Q'(\vec{x})d^d\vec{x}. \tag{1.1.13}$$

Here the transpose is taken on the internal space. For complex valued fields this generalizes exactly as before:

$$\langle Q| : \langle Q|Q'\rangle = \int_{\mathbb{S}} (Q^T)^*(\vec{x})Q'(\vec{x})d^d\vec{x} = \int_{\mathbb{S}} Q^\dagger(\vec{x})Q'(\vec{x})d^d\vec{x}. \tag{1.1.14}$$

The complex conjugate of the transpose is clunky to write, so in the above we introduced the Hermitian adjoint or conjugate, notated with the dagger: $Q^\dagger = (Q^T)^*$.

In Dirac notation, we think of fields as vectors in an infinite dimensional vector space (there are infinitely many choices for $\vec{x}$) allowing us apply the results about the analytical mechanics of vectors. There are some significant technical details necessary to properly extend results of finite dimensional linear algebra to infinite dimensional vectors spaces that we won't cover. There will be some theorems and results for which we give a finite dimensional argument and then just assume the corresponding infinite dimensional result holds. This is because proving the infinite dimensional result in that case would be extremely technically challenging and for our purposes not worth the extra effort. See [6] for a proper introduction to functional analysis.

Since we gave our vector space an inner product, we can use it to define the norm $\|\psi\|$ or length of a vector $|\psi\rangle$ in the standard way by the by non-negative square root of the inner product of a vector with itself. Having a norm then allows us to define the distance $d(u, v)$ between two vectors $|u\rangle$ and $|v\rangle$ by the norm of their difference. That is:

$$\|\psi\| \stackrel{\text{def}}{=} \sqrt{\langle\psi|\psi\rangle}, \qquad d(u, v) \stackrel{\text{def}}{=} \|u - v\|. \tag{1.1.15}$$

We will find in the next section that we will naturally be considering fields of finite norm as our vector space of functions $\mathbb{H}_{\mathbb{S}}$. Such functions are said to be **square-integrable** because finite norm is thus equivalent to their magnitude squared having a finite integral over all space. A technical but mathematically essential property of the vector space of functions with finite norm is that any sequence of such functions so that the distance between neighboring functions in the sequence limits to 0 itself limits to a function which also has finite norm. Thus, said space is called "complete" as it is closed under this class of limits. A vector space which both has an inner product and is complete in this sense is called a **Hilbert space**.

**Theorem 1.1: Linear Algebra of Fields**

A field theory with configuration fields $Q(t, \vec{x})$ of finite norm can be represented as a vector $|Q(t)\rangle$ in a Hilbert space in the Dirac notation by:

$$|Q(t)\rangle = \int_{\mathbb{S}} Q(t, \vec{x})|\vec{x}\rangle d^d\vec{x} \tag{1.1.16}$$

where the vector $|\vec{x}\rangle$ corresponds to a Dirac delta centered at $\vec{x}$ and the inner product on the Hilbert space is given by:

$$\langle f|g\rangle = \int_{\mathbb{S}} f^\dagger(\vec{x})g(\vec{x})d^d\vec{x}. \tag{1.1.17}$$

The $|\vec{x}\rangle$ states satisfy the orthogonality and completeness relations:

$$\langle\vec{x}|\vec{x}'\rangle = \delta(\vec{x} - \vec{x}') \qquad \int_{\mathbb{S}} |\vec{x}\rangle\langle\vec{x}|d^d\vec{x} = \mathbb{1}. \tag{1.1.18}$$

## Exercises

**Problem 1.1: Practice with Dirac notation.** Consider a system in 1 spatial dimension with trivial internal space (and so with the Hilbert space of square-integrable functions of one real variable). In particular consider the following two vectors in $\mathbb{H}$:

$$|\psi_1\rangle = \int_{-\infty}^{\infty} N_1 e^{-(x-\mu_1)^2/(4\sigma^2)+ik_1x}|x\rangle dx, \qquad |\psi_2\rangle = \int_{-\infty}^{\infty} N_2 e^{-(x-\mu_2)^2/(4\sigma^2)+ik_2x}|x\rangle dx$$

for positive constants $\sigma$, $N_1$, $N_2$ and real constants $\mu_1, k_1, \mu_2, k_2$. Compute the following:

$$\|\psi_1\|, \qquad \|\psi_2\|, \qquad \langle\psi_1|\psi_2\rangle, \qquad d(\psi_1,\psi_2).$$

Identify what values of $N_1$ and $N_2$ make $|\psi_1\rangle$ and $|\psi_2\rangle$ have norm 1. Using geometric thinking about the nature of the inner product defined in the section, try to come up with a reasonable definition of the angle $\theta$ between the vectors $|\psi_1\rangle$ and $|\psi_2\rangle$. Find a formula for that angle and compute it for this example.

## 1.2 Schrödinger Equation

The argument of this section is based on [45]. Now that we have reformulated our field theory as a vector theory in an appropriate Hilbert space, let us study the general Hamiltonian for a vector of arbitrary dimension. We let $Q$ be our coordinate vector (on the configuration space, a vector space of dimension $D$, potentially infinite) and $P$ be our conjugate momentum (making phase space a vector space of dimension $2D$). Because we would ultimately like to find a wave theory of matter and the properties of matter do not inescapably grow with time, our theory should be stable in the sense that finite initial conditions should remain finite for arbitrarily long times. Consequently, this guarantees that our system has an equilibrium point (a point in phase space so that if we start at that location, we stay at that location forever) [50]. Let us define $(Q,P)$ so that $(0,0)$ is such an equilibrium of our Hamiltonian (if in some coordinates this were not true, we would define new canonical coordinates equal to the old ones with the coordinate values of such an equilibrium subtracted. Subtracting constants is a canonical transformation so this is always possible). Then, we can Taylor expand the Hamiltonian nearby equilibrium and apply the theorem of normal modes (Orig. Theorem 1.14). Doing so, we know that in appropriately chosen canonical coordinates our Hamiltonian can be written as a series of decoupled harmonic oscillators, up to corrections which are at least cubic in our canonical variables:

$$\mathcal{H}(Q,P) = \sum_{k=1}^{D}\left(\frac{a_k}{2}Q_k^2 + \frac{b_k}{2}P_k^2\right) + \text{Higher Order Terms} \tag{1.2.1}$$

for constants $a_k$ and $b_k$ with the property that $a_k b_k > 0$ for each $k$. Now define $\omega_k$ so that:

$$\omega_k = \sqrt{a_k b_k}\,\mathrm{sgn}(a_k) \tag{1.2.2}$$

and make the complex canonical transformation known as *John (Ralston)'s Quantum Map* to new coordinates $\phi_n$ and momenta $\varphi_n$ defined by:

$$\phi_n = \frac{1}{\sqrt{2}}\left(\left(\frac{a_n}{b_n}\right)^{\frac{1}{4}}Q_n + i\left(\frac{b_n}{a_n}\right)^{\frac{1}{4}}P_n\right), \qquad \varphi_n = \frac{1}{\sqrt{2}}\left(\left(\frac{b_n}{a_n}\right)^{\frac{1}{4}}P_n + i\left(\frac{a_n}{b_n}\right)^{\frac{1}{4}}Q_n\right) = i\phi_n^* \tag{1.2.3}$$

You can check quickly that these new coordinates $(\phi_n,\varphi_n)$ are indeed canonical (i.e. they satisfy the canonical Poisson brackets with each other). Writing the Hamiltonian in terms of them (dropping higher order terms):

$$\mathcal{H} = \frac{1}{i}\sum_n \omega_n\varphi_n\phi_n = \sum_n \omega_n\phi_n^*\phi_n = \sum_n \omega_n|\phi_n|^2. \tag{1.2.4}$$

By (1.2.3), $\varphi_n = i\phi_n^*$, the conjugate momentum can be written in terms of the complex conjugate and so the complex numbers $\phi_n$ determine the entire state.

Now define a $D$ dimensional complex Hilbert space with basis vectors $|n_\Omega\rangle$ written in components as column vectors so that $|n_\Omega\rangle$ has 0's in each entry except entry $n$ and has entry $n$ equal to 1. (The meaning of the subscript $\Omega$ in $|n_\Omega\rangle$ will be made clear shortly.) As well, define $|\phi\rangle$ as a vector in this Hilbert space and $\langle\varphi|$ as a vector in the dual Hilbert space by:

$$|\phi\rangle = \sum_n \phi_n|n_\Omega\rangle, \qquad \langle\varphi| = \sum_n \varphi_n\langle n_\Omega| = \sum_n i\phi_n^*\langle n_\Omega| = i\langle\phi|. \tag{1.2.5}$$

By defining the angular frequency operator we can give a basis independent expression for $\mathcal{H}$:

$$\hat{\Omega} = \sum_n \omega_n|n_\Omega\rangle\langle n_\Omega| = \hat{\Omega}^\dagger \implies \mathcal{H} = \frac{1}{i}\langle\varphi|\hat{\Omega}|\phi\rangle = \langle\phi|\hat{\Omega}|\phi\rangle. \tag{1.2.6}$$

It is now natural to see why the Hilbert space of fields with finite norm is of interest to us. If $\langle\phi|\phi\rangle$ is not finite, the $\mathcal{H}$ is liable to also be nonfinite as we could make $|\phi_n|^2$ arbitrarily large for a particular $n$. Thus, because we are physically interested in finite energy states we also take $\langle\phi|\phi\rangle$ to be finite. We can see clearly what is

the motivation for the subscript $\Omega$ on $|n_\Omega\rangle$. It is a notation to help remember that $|n_\Omega\rangle$ is an eigenvector of the $\hat{\Omega}$ operator:

$$\hat{\Omega}|n_\Omega\rangle = \omega_n|n_\Omega\rangle. \tag{1.2.7}$$

From our Hamiltonian, we can compute the Lagrangian:

$$\mathcal{L} = \sum_n \varphi_n\dot{\phi}_n - \mathcal{H} = i\langle\phi|\dot{\phi}\rangle - \langle\phi|\hat{\Omega}|\phi\rangle. \tag{1.2.8}$$

Hence, any stable Hamiltonian to quadratic order can be written in the form of (1.2.6) or correspondingly its Lagrangian can be written in the form of (1.2.8) as long as we make sufficient variable transformations to find the appropriate expression for $|\phi\rangle$. It is important to note that we started with the real configuration space vector space of dimension $D$, went to the real phase space vector space of dimension $2D$ and then mapped that phase space down into a complex vector space of complex dimension $D$. The appearance of complex numbers was completely unmysterious. We introduced them in a coordinate transformation to organize an equation nicely. With the Lagrangian and Hamiltonian we have found, we can now create a field theory for matter in a weak field limit about equilibrium in the correct coordinates. That is, our field theory for matter is defined by whatever $|\phi\rangle$ and $\hat{\Omega}$ are.

So far in this section we have worked under the assumption that the Hamiltonian $\mathcal{H}$ for our field would have time translational symmetry (because we believe that to be a true symmetry of the universe). However, when our field interacts with external forces we know from experience it may be interesting to consider an explicitly time dependent Hamiltonian to capture at least some of those external effects. So, sometimes we will find it useful for the angular frequency operator $\hat{\Omega}$ to be explicitly time dependent. When $\hat{\Omega}(t)$ is explicitly time dependent the eigenvectors $|n_\Omega(t)\rangle$ will generally also be time dependent. This makes them usually an inconvenient basis to use for computations. It will be important for us to understand how to make calculations in a generic constant basis of orthonormal vectors $|n\rangle$:

$$\langle n|n'\rangle = \delta_{nn'}. \tag{1.2.9}$$

Because the $|n\rangle$ form a Hilbert space basis, $|\phi\rangle$ can be written in terms of them for some coefficients $\phi_n$:

$$|\phi\rangle = \sum_n \phi_n|n\rangle \implies \langle n|\phi\rangle = \sum_{n'} \phi_{n'}\langle n|n'\rangle = \sum_{n'} \phi_{n'}\delta_{nn'} = \phi_n \tag{1.2.10}$$

$$\implies |\phi\rangle = \sum_n (\langle n|\phi\rangle)|n\rangle = \left(\sum_n |n\rangle\langle n|\right)|\phi\rangle \quad \forall|\phi\rangle \quad \implies \sum_n |n\rangle\langle n| = \mathbb{1}. \tag{1.2.11}$$

As we are now in a generic orthonormal basis, these $\phi_n$ are not the same $\phi_n$ used before in the $|n_\Omega\rangle$ basis. For the rest of this section whenever we write $\phi_n$ we will always mean $\phi_n$ as defined in this constant arbitrary basis. In an arbitrary basis we define the matrix elements of $\hat{\Omega}$ as:

$$\Omega_{nk}(t) = \langle n|\hat{\Omega}(t)|k\rangle. \tag{1.2.12}$$

Notice in (1.2.6) that by construction $\hat{\Omega}$ is equal to its own complex conjugate transpose (adjoint) and so is called **self-adjoint**. Because of this:

$$\Omega^*_{nk} = (\langle n|\hat{\Omega}|k\rangle)^\dagger = \langle k|\hat{\Omega}^\dagger|n\rangle = \langle k|\hat{\Omega}|n\rangle = \Omega_{kn}. \tag{1.2.13}$$

Using the completeness relation in (1.2.11) we have:

$$\langle\phi|\hat{\Omega}|\phi\rangle = \sum_n\sum_k \langle\phi|n\rangle\langle n|\hat{\Omega}|k\rangle\langle k|\phi\rangle = \sum_{nk} \Omega_{nk}\phi^*_n\phi_k. \tag{1.2.14}$$

With this our Lagrangian becomes:

$$\mathcal{L} = i\sum_n \phi^*_n\dot{\phi}_n - \sum_{nk} \Omega_{nk}\phi^*_n\phi_k. \tag{1.2.15}$$

This gives the equation of motion for the theory:

$$\frac{\partial\mathcal{L}}{\partial\phi^*_n} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\phi}^*_n} = i\dot{\phi}_n - \sum_k \Omega_{nk}\phi_k \stackrel{!}{=} 0 \implies i\dot{\phi}_n = \sum_k \Omega_{nk}\phi_k. \tag{1.2.16}$$

In terms of the Dirac notation this becomes:

$$i|\dot{\phi}\rangle = \sum_n i\dot{\phi}_n|n\rangle = \sum_{nk} |n\rangle\Omega_{nk}\phi_k = \sum_{nk} |n\rangle\langle n|\hat{\Omega}|k\rangle\langle k|\phi\rangle = \hat{\Omega}|\phi\rangle \implies i|\dot{\phi}\rangle = \hat{\Omega}|\phi\rangle. \tag{1.2.17}$$

The conjugate Euler-Lagrange equation gives the equation of motion for $\langle\phi|$:

$$\frac{\partial\mathcal{L}}{\partial\phi_n} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\phi}_n} = -\sum_k \phi_k^*\Omega_{kn} - i\dot{\phi}_n^* \stackrel{!}{=} 0 \implies i\dot{\phi}_n^* = -\sum_k \Omega_{kn}\phi_k^* \implies i\langle\dot{\phi}| = -\langle\phi|\hat{\Omega}. \tag{1.2.18}$$

We can put the whole theory into a bit more refined form using some properties of the equations of motion. Notice that our equation of motion always conserves the magnitude of the state vector in time:

$$\frac{d}{dt}\langle\phi|\phi\rangle = \frac{d\langle\phi|}{dt}|\phi\rangle + \langle\phi|\frac{d|\phi\rangle}{dt} = -\frac{1}{i}\langle\phi|\hat{\Omega}|\phi\rangle + \frac{1}{i}\langle\phi|\hat{\Omega}|\phi\rangle = 0. \tag{1.2.19}$$

Because this magnitude is constant, we can divide it out of the theory. Define **Planck's constant**[3] $\hbar$:

$$\hbar \stackrel{\text{def}}{=} \langle\phi|\phi\rangle = \sum_n |\phi_n|^2 > 0 \tag{1.2.20}$$

and the normalized state vector $|\psi\rangle$:

$$|\psi\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{\hbar}}|\phi\rangle \implies \langle\psi|\psi\rangle = 1. \implies \mathcal{L} = i\hbar\langle\psi|\dot{\psi}\rangle - \hbar\langle\psi|\hat{\Omega}|\psi\rangle. \tag{1.2.21}$$

From the first term of this Lagrangian and using that $\langle\psi|\psi\rangle = 1$, we can identify the units of $\hbar$:

$$[\mathcal{L}] = [\hbar][\langle\psi|\dot{\psi}\rangle] = \frac{\hbar}{T} \implies [\hbar] = [S] \tag{1.2.22}$$

where $T$ is the unit of time. That is, $\hbar$ has the same units as the action. We introduced the symbol $\hbar$ previously in Orig. problem 4.48, in Orig. (5.2.4), and in Orig. (5.12.93); this is not an accident. The $\hbar$ in Orig. (5.2.4) was allowed to be any positive constant with the units of the action, and we take from now on by definition that $\hbar$ to be chosen to be the same $\hbar$ appearing in the Schrödinger Lagrangian. Electromagnetism is a field theory with a quadratic Hamiltonian and so by the argument of this section can be put into the form of the Schrödinger equation (in fact the homework questions in Orig. section 4.4 basically do this) and so the $\hbar$ appearing in Orig. problem 4.48 enters the same way it does in the Schrödinger Lagrangian. It isn't obvious right now that the $\hbar$ introduced in Orig. problem 4.48 has the same value as the $\hbar$ introduced in the Schrödinger Lagrangian in general. At the level of classical field theory, there is no reason these values should be related, because $\hbar$ is capturing the overall magnitude of the field as an arbitrary initial condition. However, it will turn out that due to *quantum field theory*, everywhere we used $\hbar$ in the Orig. it really is the *same* $\hbar$. We will derive that fact in part at the end of this chapter and finish properly doing so next chapter.

For now, we just observe that we can divide out $\hbar$, some positive constant parameter of the theory, from the field magnitude and that doing so results in an overall multiplier on the Lagrangian. Having $|\psi\rangle$ set so that $\langle\psi|\psi\rangle = 1$ is said to "normalize" $|\psi\rangle$. Now then our Hamiltonian is given by:

$$\mathcal{H} = \langle\phi|\hat{\Omega}|\phi\rangle = \langle\psi|\hbar\hat{\Omega}|\psi\rangle = \langle\psi|\hat{H}|\psi\rangle \tag{1.2.23}$$

where we have defined the Hamiltonian operator:

$$\hat{H}(t) = \hbar\hat{\Omega}(t) \implies \hat{H}^\dagger = \hat{H}. \tag{1.2.24}$$

This gives us:

**Theorem 1.2: Lagrangian for the Schrödinger equation**

There exist coordinates in which a generic stable lowest order (i.e. quadratic) field theory is described by a complex field $\psi(t,\vec{x})$ in a Hilbert space satisfying:

$$\langle\psi|\psi\rangle = 1 \tag{1.2.25}$$

with Lagrangian given by:

$$\mathcal{L} = i\hbar\langle\psi|\dot{\psi}\rangle - \langle\psi|\hat{H}(t)|\psi\rangle \tag{1.2.26}$$

where $\hbar > 0$ is a constant and $\hat{H}$ is a self-adjoint operator called the Hamiltonian operator which defines the dynamics of the theory. Correspondingly the conjugate momentum $\pi$ and Hamiltonian $\mathcal{H}$

[3] Often $\hbar$ is called the *reduced Planck's constant* and $2\pi\hbar$ is called *Planck's constant*; we will not do this.

are given by:

$$\pi = \frac{\delta \mathcal{L}}{\delta \dot{\psi}} = i\hbar\psi^\dagger, \qquad \mathcal{H} = \langle \pi|\dot{\psi}\rangle - \mathcal{L} = \langle\psi|\hat{H}|\psi\rangle = \frac{1}{i\hbar}\langle\pi|\hat{H}|\psi\rangle. \tag{1.2.27}$$

The state $|\psi\rangle$ then satisfies the Schrödinger equation:

$$\frac{\partial \mathcal{L}}{\partial\langle\psi|} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial\langle\dot{\psi}|} = i\hbar|\dot{\psi}\rangle - \hat{H}|\psi\rangle \stackrel{!}{=} 0 \implies i\hbar|\dot{\psi}\rangle = \hat{H}|\psi\rangle \tag{1.2.28}$$

and its adjoint satisfies the adjoint Schrödinger equation:

$$\frac{\partial \mathcal{L}}{\partial|\psi\rangle} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial|\dot{\psi}\rangle} = -\langle\psi|\hat{H} - \frac{d}{dt}i\hbar\langle\psi| \stackrel{!}{=} 0 \implies i\hbar\langle\dot{\psi}| = -\langle\psi|\hat{H}. \tag{1.2.29}$$

The **Schrödinger equation** of motion (1.2.28) is *the* central equation of all of quantum mechanics. Going forward, we will generally call $|\psi\rangle$ the wave vector or state vector and $\psi(t,\vec{x})$ the wavefunction. We will not retain the use of the symbol $\phi$ as used in this section past this point and will consistently stick with only using $\psi$. You might object that the Lagrangian in theorem 1.2 is not real-valued, which does seem odd considering we arrived at it through canonical transformations of a system initially in real variables. However, it is actually real-valued up to a total time derivative and so can be made real-valued with an appropriate surface term. In particular:

$$\mathcal{L} = \frac{i}{2}\hbar\langle\psi|\dot{\psi}\rangle - \frac{i}{2}\hbar\langle\dot{\psi}|\psi\rangle - \langle\psi|\hat{H}|\psi\rangle + \frac{d}{dt}(i\hbar\langle\psi|\psi\rangle) \tag{1.2.30}$$

$$\simeq \frac{i\hbar}{2}(\langle\psi|\dot{\psi}\rangle - \langle\dot{\psi}|\psi\rangle) - \langle\psi|\hat{H}|\psi\rangle = -\hbar\mathrm{Im}(\langle\psi|\dot{\psi}\rangle) - \langle\psi|\hat{H}|\psi\rangle \tag{1.2.31}$$

which is now manifestly real-valued. (Here we recall the notation in the Orig. that $\simeq$ between Actions/Lagrangians/Hamiltonians indicates equivalence up to adding a surface term.) In order to arrive at the Schrödinger equation we dropped all possible terms in the Hamiltonian above quadratic order. We will show next chapter why in a truly fundamental formulation of quantum mechanics no such terms appear. But, single particle quantum mechanics, which is just an effective theory for that more fundamental formulation, can actually have nonlinear terms appear. We will even see some this chapter. Even when terms beyond quadratic in $\psi$, the $\psi$ coordinates are most often the most efficient and transparent coordinates in which to treat those higher terms as perturbations, and so remain useful. Even if you've seen some quantum mechanics before, you likely haven't seen it in Lagrangian form as in (1.2.26). As you can see here, quantum mechanics is not a replacement of analytical mechanics. It is a specific special case of it! The power of this perspective will make itself very clear over the coming sections.

### 1.2.1 Hermitian vs. Self-Adjoint

We have referred to the Hamiltonian operator as self-adjoint. However, when dealing with fields (and so infinite dimensional vector spaces instead of finite dimensional) self-adjointness is a delicate property. There is a related concept known as being Hermitian, which for finite dimensional vector spaces is the same thing. For infinite dimensional vector spaces they are different (being self-adjoint is a special case of being Hermitian) and many nice theorems we would like to use only apply when the operator to which we would apply them is specifically self-adjoint. For the rest of this subsection we are going to be careful in treating the concepts of Hermitian vs. self-adjoint but beyond this subsection we will not be so careful.

The subtlety in dealing with infinite dimensional Hilbert spaces comes from the fact that in such a space the inner product $\langle\psi_1|\psi_2\rangle$ between two vectors contains an infinite sum when expressed in components and the convergence properties of that sum can be quite delicate. Suppose we have a linear operator $\hat{\mathcal{O}}$ on the Hilbert space $\mathbb{H}$. Colloquially this means that when $\hat{\mathcal{O}}$ acts on a vector $|\psi_1\rangle \in \mathbb{H}$ the output is some other vector in $\mathbb{H}$. Let $|\psi_2\rangle$ be that vector:

$$|\psi_2\rangle = \hat{\mathcal{O}}|\psi_1\rangle. \tag{1.2.32}$$

$|\psi_1\rangle$, by being a vector in the Hilbert space, is required to have finite norm. However, depending on the particular matrix elements $\mathcal{O}_{nk}$ and components $\psi_{1n}$, it may be the case that $|\psi_2\rangle$ fails to have finite norm

and thus to be part of $\mathbb{H}$. Explicitly:

$$\langle\psi_2|\psi_2\rangle = \langle\psi_1|\hat{\mathcal{O}}^\dagger\hat{\mathcal{O}}|\psi_1\rangle = \sum_{nmk}\langle\psi_1|n\rangle\langle n|\hat{\mathcal{O}}^\dagger|m\rangle\langle m|\hat{\mathcal{O}}|k\rangle\langle k|\psi_1\rangle = \sum_{nmk}\psi^*_{1n}\mathcal{O}^*_{mn}\mathcal{O}_{mk}\psi_{1k}. \tag{1.2.33}$$

This infinite sum may not converge. In order to rigorously define the operator $\hat{\mathcal{O}}$ we have to, in addition to its matrix elements in some basis, supply a set $D(\hat{\mathcal{O}})$ called the domain of $\hat{\mathcal{O}}$ which is a linear subspace of the overall Hilbert space $\mathbb{H}$. This set $D(\hat{\mathcal{O}})$ is the set of vectors $|u\rangle$ in $\mathbb{H}$ on which we define $\hat{\mathcal{O}}|u\rangle$ to exist. Necessarily, it can only contain vectors $|u\rangle$ so that $\hat{\mathcal{O}}|u\rangle$ has finite norm. We will basically always require that whatever operators we are working with are densely defined in the Hilbert space. An operator $\hat{\mathcal{O}}$ is said to be densely defined if for any given vector $|v\rangle$ in $\mathbb{H}$ and $\varepsilon > 0$ there always exists a vector $|u\rangle \in D(\hat{\mathcal{O}})$ so that the distance between $|v\rangle$ and $|u\rangle$ is less than $\varepsilon$. This essentially means that every vector in $\mathbb{H}$ can be arbitrarily well approximated by a vector in $D(\hat{\mathcal{O}})$.

In order to rigorously define $\hat{\mathcal{O}}^\dagger$, suppose we have a densely defined operator $\hat{\mathcal{O}}$ on a domain $D(\hat{\mathcal{O}})$. The domain $D(\hat{\mathcal{O}}^\dagger)$ of the adjoint $\hat{\mathcal{O}}^\dagger$ is the set of all $|v\rangle \in \mathbb{H}$ so that there exists a $|w\rangle \in \mathbb{H}$ satisfying:

$$\langle v|(\hat{\mathcal{O}}|u\rangle) = \langle w|u\rangle \tag{1.2.34}$$

for all $|u\rangle \in D(\hat{\mathcal{O}})$. Because $\hat{\mathcal{O}}$ is assumed to be densely defined we can use the **Riesz representation theorem** to guarantee that such a $|w\rangle$ is unique if it exists and then we take by definition $\hat{\mathcal{O}}^\dagger$ to act on $|v\rangle$ according to:

$$\hat{\mathcal{O}}^\dagger|v\rangle = |w\rangle. \tag{1.2.35}$$

(The Riesz representation theorem is the formal theorem that proves that each vector $|u\rangle \in \mathbb{H}$ has a unique dual vector $\langle u|$ and vice versa. For finite dimensional vector spaces this is obvious; for the infinite dimensional case it is still true but requires a more complicated proof to show it.) Importantly, the domain of an operator and the domain of that operator's adjoint are generally not the same.

A densely defined operator $\hat{\mathcal{O}}$ is Hermitian if for all vectors $|u\rangle$ and $|v\rangle$ which are both in $D(\hat{\mathcal{O}})$:

$$\langle u|(\hat{\mathcal{O}}|v\rangle) = (\langle u|\hat{\mathcal{O}}^\dagger)|v\rangle. \tag{1.2.36}$$

The parentheses are important because they control the order in which infinite sums are taken which in general may affect the result. In order for this equation to be well defined, by definition it means $|v\rangle$ is in $D(\hat{\mathcal{O}}^\dagger)$. Therefore, if this equation holds for all $|v\rangle \in D(\hat{\mathcal{O}})$ then $D(\hat{\mathcal{O}})$ must be a subset of $D(\hat{\mathcal{O}}^\dagger)$. A Hermitian operator $\hat{\mathcal{O}}$ is said to be self-adjoint whenever additionally $D(\hat{\mathcal{O}}) = D(\hat{\mathcal{O}}^\dagger)$. Then, the above equation holds without the need for parentheses:

$$\langle u|\hat{\mathcal{O}}|v\rangle = \langle u|\hat{\mathcal{O}}^\dagger|v\rangle. \tag{1.2.37}$$

In physics we generally just assume the parentheses to be unnecessary as part of our bad habit of recklessly commuting limits. In the homework you will work through a simple concrete example of an operator which is Hermitian but not self-adjoint, due to domain issues.

### 1.2.2 Spectral Theorem

It will be extremely useful to know some basic properties of self-adjoint operators, since those will then be properties of our Hamiltonian operator. Consider a self-adjoint operator $\hat{\mathcal{O}}$ and an eigenstate $|\psi\rangle$ of that operator with eigenvalue $\lambda$:

$$\hat{\mathcal{O}}|\psi\rangle = \lambda|\psi\rangle. \tag{1.2.38}$$

Then:

$$\langle\psi|\hat{\mathcal{O}}|\psi\rangle = \lambda\langle\psi|\psi\rangle \implies \lambda = \frac{\langle\psi|\hat{\mathcal{O}}|\psi\rangle}{\langle\psi|\psi\rangle} \implies \lambda^* = \frac{\langle\psi|\hat{\mathcal{O}}^\dagger|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\langle\psi|\hat{\mathcal{O}}|\psi\rangle}{\langle\psi|\psi\rangle} = \lambda. \tag{1.2.39}$$

Hence, the eigenvalues of a self-adjoint operator are always real. Taking the adjoint of (1.2.38) gives:

$$\langle\psi|\hat{\mathcal{O}}^\dagger = \lambda^*\langle\psi| \implies \langle\psi|\hat{\mathcal{O}} = \lambda\langle\psi| \tag{1.2.40}$$

and so the left eigenstates and right eigenstates of $\hat{\mathcal{O}}$ are related by taking the adjoint and share the same eigenvalue. Further, consider another eigenstate $|\psi'\rangle$ of $\hat{\mathcal{O}}$ with eigenvalue of $\lambda' \neq \lambda$:

$$\langle\psi|\hat{\mathcal{O}}|\psi'\rangle = \left(\langle\psi|\hat{\mathcal{O}}\right)|\psi'\rangle = \langle\psi|\left(\hat{\mathcal{O}}|\psi'\rangle\right) \implies \langle\psi|\lambda|\psi'\rangle = \langle\psi|\lambda'|\psi'\rangle \tag{1.2.41}$$

$$\implies (\lambda - \lambda')\langle\psi|\psi'\rangle = 0 \implies \langle\psi|\psi'\rangle = 0. \tag{1.2.42}$$

Hence, eigenvectors of a self-adjoint operator corresponding to distinct eigenvalues are guaranteed to be orthogonal. In fact, there is a deep and general result from linear algebra which we will use regularly known as the Spectral Theorem. It holds not only for self-adjoint operators, but for a larger class of important operators which are called normal operators.

**Spectral Theorem:** Every normal operator $\hat{\mathcal{O}}$ on a Hilbert space $\mathbb{H}$ possesses an orthonormal basis of eigenvectors. (A normal operator is one which satisfies $[\hat{\mathcal{O}}, \hat{\mathcal{O}}^\dagger] = 0$.) Every operator which possesses an orthonormal basis of eigenvectors is normal.

Rigorously proving the spectral theorem for an infinite dimensional Hilbert space requires significant mathematics (functional analysis) we are not going to touch. Here we will provide a proof which is valid for finite dimensional Hilbert spaces, but the result does hold with appropriately rigorous careful treatment for infinite dimensions. To prove the spectral theorem we will need a few smaller results first.

**Lemma:** Every operator has an eigenvalue. Proof:

Consider an $N$ dimensional Hilbert space $\mathbb{H}$ and an operator $\hat{\mathcal{O}}$. For all vectors $|\psi\rangle \in \mathbb{H}$, we can consider the $N+1$ vectors $|\psi\rangle$, $\hat{\mathcal{O}}|\psi\rangle$, ..., $\hat{\mathcal{O}}^N|\psi\rangle$. Because we have $N+1$ vectors in an $N$ dimensional vector space they cannot all be linearly independent. Therefore, there must exist some coefficients $a_n$ such that:

$$\sum_{n=0}^{N} a_n \hat{\mathcal{O}}^n |\psi\rangle = 0. \tag{1.2.43}$$

Let $m$ be the largest value of $n$ for which $a_n \neq 0$. Every polynomial on the complex numbers can be factored by the fundamental theorem of algebra and so for some constants $c, \lambda_1, ..., \lambda_m$ we have for any $z$:

$$a_0 + a_1 z + ... + a_m z^m = c(z - \lambda_1)...(z - \lambda_m). \tag{1.2.44}$$

Returning this to our linear dependence sum we have:

$$\sum_{n=0}^{m} a_n \hat{\mathcal{O}}^n |\psi\rangle = c(\hat{\mathcal{O}} - \lambda_1 \mathbb{1})...(\hat{\mathcal{O}} - \lambda_m \mathbb{1})|\psi\rangle = 0. \tag{1.2.45}$$

Thus, at least one of these operators $\hat{\mathcal{O}} - \lambda_k \mathbb{1}$ for some $k$ must send a nonzero vector to 0. Let $K$ be such a value of $k$ and $|v_K\rangle$ be such a nonzero vector that $\hat{\mathcal{O}} - \lambda_K \mathbb{1}$ sends to 0. Then, automatically:

$$(\hat{\mathcal{O}} - \lambda_K \mathbb{1})|v_K\rangle = 0 \implies \hat{\mathcal{O}}|v_K\rangle = \lambda_K |v_L\rangle \tag{1.2.46}$$

and so $\lambda_K$ is an eigenvalue of $\hat{\mathcal{O}}$.

**Lemma:** Every operator has an upper triangular matrix representation. Proof:

(A matrix is called upper triangular if its only nonzero entries occur on or above its main diagonal, i.e. entries for which the column number is greater than or equal to the row number. Matrix representations here are only considered in orthonormal bases.) We proceed by induction. Consider a 1 dimensional Hilbert space $\mathbb{H}$. Trivially any operator $\hat{\mathcal{O}}$ on that space represented by a matrix is just a number and so is upper triangular. Let $S_n$ be the statement "Every operator on a Hilbert space of dimension less than or equal to $n$ has an upper triangular representation". We have already proven that $S_1$ is true. Now suppose that for a particular $n$ $S_n$ is true. Then, consider a Hilbert space $\mathbb{H}$ of dimension $n+1$ and an operator on that space $\hat{\mathcal{O}}$. Let $\lambda$ be an eigenvalue of $\hat{\mathcal{O}}$ (which is guaranteed to exist by our previous lemma) and consider for any vector $|\psi\rangle$:

$$\hat{\mathcal{O}}|\psi\rangle = (\hat{\mathcal{O}} - \lambda \mathbb{1})|\psi\rangle + \lambda|\psi\rangle. \tag{1.2.47}$$

Let $|v\rangle$ be a normalized ($\langle v|v\rangle = 1$) eigenvector of $\hat{\mathcal{O}}$ with eigenvalue $\lambda$. Any $|\psi\rangle \in \mathbb{H}$ can be written as a part parallel to $|v\rangle$ and a piece orthogonal to $|v\rangle$:

$$|\psi\rangle = A|v\rangle + |\psi_\perp\rangle, \qquad \langle v|\psi_\perp\rangle = 0. \tag{1.2.48}$$

(To see this simply define $|\psi_\perp\rangle = |\psi\rangle - \langle v|\psi\rangle|v\rangle$ and $A = \langle v|\psi\rangle$.) Now acting $\hat{\mathcal{O}}$ on $|\psi\rangle$:

$$\hat{\mathcal{O}}|\psi\rangle = \lambda|\psi\rangle + (\hat{\mathcal{O}} - \lambda\mathbb{1})|\psi_\perp\rangle. \tag{1.2.49}$$

The first term is $\lambda|\psi\rangle = \lambda\mathbb{1}|\psi\rangle$ and $\lambda\mathbb{1}$ is plainly upper triangular, because it is diagonal. The second term

can be decomposed into a piece parallel to $|v\rangle$ and a piece orthogonal to $|v\rangle$:

$$(\hat{\mathcal{O}} - \lambda \mathbb{1})|\psi_\perp\rangle = B|v\rangle + \Big((\hat{\mathcal{O}} - \lambda \mathbb{1})|\psi_\perp\rangle\Big)_\perp . \tag{1.2.50}$$

The second piece here is then some linear operator on $|\psi_\perp\rangle$ that returns a vector in $\mathbb{H}_v$, which is a Hilbert space of dimension $n$. Let $|u_1\rangle, ..., |u_n\rangle$ be an orthonormal basis of $\mathbb{H}_v$ on which $\hat{\mathcal{O}}$ has an upper triangular representation, which is guaranteed to exist by $S_n$. Then, we can complete a basis of $\mathbb{H}$ adding $|v\rangle$ to the end of the list. In that basis, both pieces of (1.2.50) are upper triangular, and thus the full expression for $\hat{\mathcal{O}}$ is upper triangular, showing $S_{n+1}$. By induction then, $S_n$ is true for all $n$.

**Proof of the Spectral Theorem:**

Consider a normal operator $\hat{\mathcal{O}}$ on a Hilbert space of dimension $N$ and an orthonormal basis $|n\rangle$ on which it is upper triangular, which is guaranteed to exist by the previous lemma:

$$\hat{\mathcal{O}}|n\rangle = \sum_{k=1}^{n} \mathcal{O}_{kn}|k\rangle \tag{1.2.51}$$

for coefficients $\mathcal{O}_{kn}$. Using the completeness relation:

$$\hat{\mathcal{O}} = \sum_{n=1}^{N}\sum_{k=1}^{n} \mathcal{O}_{kn}|k\rangle\langle n| \implies \hat{\mathcal{O}}^\dagger = \sum_{n=1}^{N}\sum_{k=1}^{n} \mathcal{O}^*_{kn}|n\rangle\langle k|. \tag{1.2.52}$$

Because $\hat{\mathcal{O}}$ is normal:

$$\|\hat{\mathcal{O}}|\psi\rangle\|^2 = \langle\psi|\hat{\mathcal{O}}^\dagger\hat{\mathcal{O}}|\psi\rangle = \langle\psi|\hat{\mathcal{O}}\hat{\mathcal{O}}^\dagger|\psi\rangle = \|\hat{\mathcal{O}}^\dagger|\psi|\rangle\|^2. \tag{1.2.53}$$

We can compute both sides easily for $|\psi\rangle = |n\rangle$:

$$\|\hat{\mathcal{O}}|n\rangle\|^2 = \sum_{k=1}^{n} |\mathcal{O}_{kn}|^2, \qquad \|\hat{\mathcal{O}}^\dagger|n\rangle\|^2 = \sum_{k=n}^{N} |\mathcal{O}_{nk}|^2 \tag{1.2.54}$$

and these two must be equal for each $n$. For $n = 1$ this requires:

$$|\mathcal{O}_{11}|^2 = \sum_{k=1}^{N} |\mathcal{O}_{1k}|^2 = |\mathcal{O}_{11}|^2 + \sum_{k=2}^{N} |\mathcal{O}_{1k}|^2. \tag{1.2.55}$$

This is only possible if $\mathcal{O}_{1k} = 0$ for $k > 1$. Consequently, the first row in $\hat{\mathcal{O}}$'s upper triangle is 0 after the first entry. Using this, we have for $n = 2$:

$$|\mathcal{O}_{22}|^2 = |\mathcal{O}_{22}|^2 + \sum_{k=3}^{N} |\mathcal{O}_{2k}|^2 \tag{1.2.56}$$

from which we get that $\mathcal{O}_{2k} = 0$ for $k > 2$. Iterating further, we get at each $n$ that $\mathcal{O}_{nk} = 0$ for each $k > n$. This leaves only diagonal entries in $\hat{\mathcal{O}}$. So, if $\hat{\mathcal{O}}$ is normal and written in a basis in which it is upper triangular then in that basis it is automatically diagonal. Because we used an orthonormal basis to diagonalize $\hat{\mathcal{O}}$, this proves that every normal operator has an orthonormal basis of eigenvectors.

The other direction is easier. Suppose $\hat{\mathcal{O}}$ has an orthonormal basis of eigenvectors. Then, in that basis:

$$\hat{\mathcal{O}} = \sum_n \lambda_n |n\rangle\langle n| \tag{1.2.57}$$

for eigenvalues $\lambda_n$. Automatically:

$$\hat{\mathcal{O}}^\dagger = \sum_n \lambda_n^* |n\rangle\langle n| \implies [\hat{\mathcal{O}}^\dagger, \hat{\mathcal{O}}] = \hat{\mathcal{O}}^\dagger\hat{\mathcal{O}} - \hat{\mathcal{O}}\hat{\mathcal{O}}^\dagger = \sum_n \lambda_n^*\lambda_n |n\rangle\langle n| - \sum_n \lambda_n\lambda_n^* |n\rangle\langle n| = 0. \tag{1.2.58}$$

Thus, if $\hat{\mathcal{O}}$ has an orthonormal basis of eigenvectors then it is a normal operator, completing our proof of the spectral theorem. ■

Self-adjoint operators are of course normal as $[\hat{\mathcal{O}}, \hat{\mathcal{O}}^\dagger] = [\hat{\mathcal{O}}, \hat{\mathcal{O}}] = 0$. But, there are many normal operators which are not self-adjoint. Another important class of such operators we will see often is that of **unitary** operators. A unitary operator $\hat{U}$ is an operator whose adjoint is its inverse:

$$\hat{U}^{-1} = \hat{U}^\dagger. \tag{1.2.59}$$

In finite dimensional vector spaces the left inverse and right inverse of an operator are the same. In the infinite dimensional case an operator may have a left inverse but not a right inverse and vice versa. Because

of this, the notation $\hat{U}^{-1}$ for an infinite dimensional Hilbert space is only defined when the operator has both a left inverse and a right inverse and when those two inverses are the same operator. Unitarity can equivalently be written as:

$$\hat{U}^\dagger \hat{U} = \hat{U}\hat{U}^\dagger = \mathbb{1}. \tag{1.2.60}$$

Unitary operators are automatically normal because: $[\hat{U}, \hat{U}^\dagger] = \hat{U}\hat{U}^\dagger - \hat{U}^\dagger\hat{U} = \mathbb{1} - \mathbb{1} = 0$. There are also many normal operators which are neither self-adjoint nor unitary.

Consider a normal operator $\hat{\mathcal{O}}$ with orthonormal eigenvectors $|n\rangle$ and corresponding eigenvalues $\mathcal{O}_n$. The eigenvector-eigenvalue equation, together with the completeness relation for the basis, gives us the so-called spectral expansion of $\hat{\mathcal{O}}$ in terms of its eigenbasis:

$$\hat{\mathcal{O}}|n\rangle = \lambda_n|n\rangle \implies \hat{\mathcal{O}}\sum_n |n\rangle\langle n| = \sum_n \lambda_n |n\rangle\langle n| \implies \hat{\mathcal{O}} = \sum_n \lambda_n |n\rangle\langle n| \tag{1.2.61}$$

and a correspondingly simple expression for $\langle\psi|\hat{\mathcal{O}}|\psi\rangle$:

$$\langle\psi|\hat{\mathcal{O}}|\psi\rangle = \langle\psi \sum_n \lambda_n |n\rangle\langle n|\psi\rangle = \sum_n \lambda_n \langle\psi|n\rangle\langle n|\psi\rangle = \sum_n \lambda_n |\psi_n|^2. \tag{1.2.62}$$

This final sum is like a weighted combination of the eigenvalues $\lambda_n$ with (nonnegative) weights $|\psi_n|^2$ (that sum to 1 if $|\psi\rangle$ is normalized). The set of eigenvalues of an operator is known as the spectrum of that operator[4].

It will be useful to understand how to change between different orthonormal bases. Consider one basis labeled as $|n_1\rangle$ and another labeled as $|n_2\rangle$. Consider a vector expanded in terms of the first basis:

$$|\psi\rangle = \sum_n \psi_n |n_1\rangle \tag{1.2.63}$$

and another vector expanded in terms of the second basis but with the same expansion coefficients:

$$|\psi'\rangle = \sum_n \psi_n |n_2\rangle. \tag{1.2.64}$$

This second vector is then related to the first some change of basis from the first to the second basis in which orthonormality is preserved. Then:

$$|\psi'\rangle = \sum_n \psi_n |n_2\rangle = \sum_n \langle n_1|\psi\rangle |n_2\rangle = \left(\sum_n |n_2\rangle\langle n_1|\right)|\psi\rangle. \tag{1.2.65}$$

And so, $|\psi'\rangle$ and $|\psi\rangle$ are related by the operator $\hat{U}$:

$$\hat{U} = \sum_n |n_2\rangle\langle n_1|, \qquad |\psi'\rangle = \hat{U}|\psi\rangle. \tag{1.2.66}$$

It immediately follows that $\hat{U}$ is unitary:

$$\hat{U}^\dagger\hat{U} = \sum_n |n_1\rangle\langle n_2| \sum_{n'} |n_2'\rangle\langle n_1'| = \sum_{nn'} |n_1\rangle\langle n_2|n_2'\rangle\langle n_1'| = \sum_{nn'} |n_1\rangle\delta_{nn'}\langle n_1'| = \sum_n |n_1\rangle\langle n_1| = \mathbb{1} \tag{1.2.67}$$

$$\hat{U}\hat{U}^\dagger = \sum_n |n_1\rangle\langle n_1| \sum_{n'} |n_1'\rangle\langle n_2'| = \sum_{nn'} |n_2\rangle\langle n_1|n_1'\rangle\langle n_2'| = \sum_{nn'} |n_2\rangle\delta_{nn'}\langle n_2'| = \sum_n |n_2\rangle\langle n_2| = \mathbb{1}. \tag{1.2.68}$$

Thus **a change of basis between orthonormal bases is represented by a unitary operator**. In the other direction, given any orthonormal basis $|n_1\rangle$ and a unitary operator $\hat{U}$ we can define as second basis $|n_2\rangle$ by $|n_2\rangle = \hat{U}|n_1\rangle$ from which it immediately follows that $|n_2\rangle$ is also an orthonormal basis:

$$\langle n_2|n_2'\rangle = \langle n_1|\hat{U}^\dagger\hat{U}|n_1'\rangle = \langle n_1|\mathbb{1}|n_1'\rangle = \delta_{nn'} \tag{1.2.69}$$

$$\sum_n |n_2\rangle\langle n_2| = \sum \hat{U}|n_1\rangle\langle n_1|\hat{U}^\dagger = \hat{U}\,\mathbb{1}\hat{U}^\dagger = \mathbb{1}. \tag{1.2.70}$$

Thus, every unitary operator is equivalent to an orthonormal change of basis.

Sometimes we will have two normal operators $\hat{A}$ and $\hat{B}$ to work with at the same time and it would be nice for calculations if we could make them both diagonal. Suppose that it is possible to find an orthonormal

[4]This definition of spectrum is valid for operators on finite dimensional operators. It is not quite correct in the infinite dimensional case but does capture the correct intuition. For our purposes, the definition given here will suffice.

basis $|n\rangle$ so that in that basis both $\hat{A}$ and $\hat{B}$ are diagonal:

$$\hat{A} = \sum_n A_n |n\rangle\langle n|, \qquad \hat{B} = \sum_n B_n |n\rangle\langle n|. \tag{1.2.71}$$

Then, automatically:

$$\hat{A}\hat{B} - \hat{B}\hat{A} = \sum_n (A_n B_n - B_n A_n)|n\rangle\langle n| = 0. \tag{1.2.72}$$

Therefore, if two normal operators are simultaneously diagonalizable then they commute. Conversely, assume that $|n\rangle$ is an orthonormal basis so that in that basis $\hat{A}$ is diagonal but $\hat{B}$ may not be:

$$\hat{A} = \sum_n A_n |n\rangle\langle n|, \qquad \hat{B} = \sum_{nm} B_{nm} |n\rangle\langle m|. \tag{1.2.73}$$

Now let us assume that $\hat{A}$ and $\hat{B}$ commute:

$$\langle n|[\hat{A}, \hat{B}]|m\rangle = (A_n - A_m)B_{nm} \stackrel{!}{=} 0. \tag{1.2.74}$$

Therefore, for each $n$ either $A_n = A_m$ or $B_{nm}$. If $A_n$ has no repeated eigenvalues then $A_n = A_m$ if and only if $n = m$. So, in that case this equality requires $B_{nm} = 0$ for all $n \neq m$ in which case $\hat{B}$ is diagonal. Therefore, if two normal operators commute and one of them has no repeated eigenvalues then they are simultaneously diagonalizable. If $\hat{A}$ does have repeated eigenvalues, then on each eigenspace $\hat{A}$ is proportional to the identity operator and we can redefine the eigenvectors within that space to diagonalize $\hat{B}$ on that subspace. Whatever choices we make $\hat{A}$ will remain diagonal on the space due to being proportional to the identity. Therefore, even if the operators have repeated eigenvalues, if they are normal and commute then they are simultaneously diagonalizable. Altogether then, **two normal operators commute if and only if they are simultaneously diagonalizable**.

It is important to understand that the sum of two self-adjoint operators is guaranteed to be self-adjoint but the product is not:

$$(\hat{A} + \hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger = \hat{A} + \hat{B} \tag{1.2.75}$$

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger = \hat{B}\hat{A} \neq \hat{A}\hat{B} \text{ (Unless } \hat{A} \text{ and } \hat{B} \text{ commute)}. \tag{1.2.76}$$

(The order of multiplication is reversed under the adjoint due to the transpose that it takes!) However, the symmetrized product of $\hat{A}$ and $\hat{B}$ is self-adjoint whenever $\hat{A}$ and $\hat{B}$ are self-adjoint:

$$\mathrm{Sym}(\hat{A}, \hat{B}) \stackrel{\text{def}}{=} \frac{\hat{A}\hat{B} + \hat{B}\hat{A}}{2} \tag{1.2.77}$$

$$\implies \mathrm{Sym}(\hat{A}, \hat{B})^\dagger = \frac{\hat{B}^\dagger \hat{A}^\dagger + \hat{A}^\dagger \hat{B}^\dagger}{2} = \frac{\hat{B}\hat{A} + \hat{A}\hat{B}}{2} = \mathrm{Sym}(\hat{A}, \hat{B}). \tag{1.2.78}$$

The overall factor of $\frac{1}{2}$ is put in so that when $\hat{A}$ and $\hat{B}$ commute, $\mathrm{Sym}(\hat{A}, \hat{B})$ reduces to simply $\hat{A}\hat{B}$. It will be useful later to have defined the symmetric product of $N$ operators as the sum of over all $N!$ orderings of the operators with a $\frac{1}{N!}$ normalization. Let $\sigma : \{1, ..., N\} \to \{1, ..., N\}$ be a permutation of the numbers $1, ..., N$. Then, the symmetric product in general is defined as:

$$\mathrm{Sym}(\hat{\mathcal{O}}_1, ..., \hat{\mathcal{O}}_N) \stackrel{\text{def}}{=} \frac{1}{N!} \sum_{\text{All } \sigma} \hat{\mathcal{O}}_{\sigma(1)} \hat{\mathcal{O}}_{\sigma(2)} ... \hat{\mathcal{O}}_{\sigma(N)}. \tag{1.2.79}$$

This definition is convenient mathematically because it is automatically self-adjoint if all of the input operators are self-adjoint, treats no order of multiplication of the input operators preferentially, is linear in the input operators, and reduces to the simple product of the operators in the case that all the input operators commute. In fact, this is the unique function of $N$ operators which satisfies all of these properties simultaneously.

Finally, the commutator of self-adjoint operators is guaranteed to be anti-self-adjoint:

$$\left([\hat{A}, \hat{B}]\right)^\dagger = \left(\hat{A}\hat{B} - \hat{B}\hat{A}\right)^\dagger = \hat{B}^\dagger \hat{A}^\dagger - \hat{A}^\dagger \hat{B}^\dagger = [\hat{B}, \hat{A}] = -[\hat{A}, \hat{B}]. \tag{1.2.80}$$

Usually it will be nicer to express this by saying that $\frac{1}{i}[\hat{A}, \hat{B}]$ is self-adjoint if both $\hat{A}$ and $\hat{B}$ are:

$$\left(\frac{1}{i}[\hat{A}, \hat{B}]\right)^\dagger = \frac{1}{i}[\hat{A}, \hat{B}]. \tag{1.2.81}$$

### 1.2.3 Adiabatic Invariants in Quantum Mechanics

If $\hat{H}$ has no explicit time dependence, then the Schrödinger Hamiltonian is integrable. (This should be a surprise, because we started with a bunch of harmonic oscillators and changed coordinates to get the Schrödinger equation, and the harmonic oscillator is integrable.) Consider a time independent Hamiltonian operator with orthonormal eigenvectors $|n\rangle$ and eigenvalues $E_n$:

$$\hat{H}|n\rangle = E_n|n\rangle, \qquad \langle n|n'\rangle = \delta_{nn'}, \qquad \sum_n |n\rangle\langle n| = \mathbb{1}. \tag{1.2.82}$$

Expressing $|\psi\rangle$ in that basis and returning the result to the Schrödinger Lagrangian:

$$|\psi\rangle = \sum_n \psi_n|n\rangle \implies \psi_n = \langle n|\psi\rangle \implies \mathcal{L} = \sum_n \left(i\hbar\psi_n^*\dot{\psi}_n - E_n\psi_n^*\psi_n\right). \tag{1.2.83}$$

Legendre transforming this to the Hamiltonian gives:

$$\pi_n = \frac{\partial\mathcal{L}}{\partial\dot{\psi}_n} = i\hbar\psi_n^* \implies \mathcal{H} = \frac{1}{i\hbar}\sum_n E_n\pi_n\psi_n. \tag{1.2.84}$$

Each $\psi_n$ then evolves quite simply as:

$$\dot{\psi}_n = \frac{\partial\mathcal{H}}{\partial\pi_n} = \frac{1}{i\hbar}E_n\psi_n \implies \psi_n(t) = e^{-\frac{i}{\hbar}E_n t}\psi_n(0). \tag{1.2.85}$$

That is, each $\psi_n$ (and consequently $\pi_n$) is periodic with period $\tau_n = \frac{2\pi\hbar}{E_n}$ and does not interact with the other coordinates. Immediately then, the theory has adiabatic invariants:

$$J_n = \frac{1}{2\pi}\oint \pi_n d\psi_n = \frac{1}{2\pi}\int_0^{\frac{2\pi\hbar}{E_n}} i\hbar\psi_n^*\frac{d\psi_n}{dt}dt = \hbar|\psi_n(0)|^2. \tag{1.2.86}$$

Because $\hbar$ is simply a constant, the quantities $|\psi_n(0)|^2 = |\psi_n(t)|^2$ for each $n$ are adiabatically invariant in quantum mechanics. This means that when the Schrödinger theory is interacted with by an external force (which could make $\hat{H}$ develop explicit $t$ dependence), the $|\psi_n(t)|^2$ evolve much more slowly and stably in time than the other degrees of freedom in the theory (such as the phase of each of the $\psi_n(t)$ or the magnitudes of the components of $|\psi\rangle$ in bases other than the eigenbasis of the Hamiltonian operator). Because of this, $|\psi_n(t)|^2$ **is stable enough on experimental time scales to be directly relevant to observations**. We can also express this quantity as:

$$|\psi_n(t)|^2 = |\langle n|\psi\rangle|^2. \tag{1.2.87}$$

We call this the **intensity** of the projection of $|\psi\rangle$ onto $|n\rangle$. In general in quantum mechanics we call the quantity $|\langle\phi|\psi\rangle|^2$ the **projective intensity** of $\phi$ onto $|\psi\rangle$ or vice versa. In a very literal sense, these intensities often measure the signal strength of various experimental quantities, such as the amount of incident electric charge on a screen from a quantum system.

Instead of using the word intensity, most other sources on quantum mechanics call the above quantity a probability. In several cases these intensities do manifest experimentally as probabilities. This is because in an experiment in which a single microscopic quantum system is measured by a macroscopic measurement device, due to the thermal background of the experiment and the interaction of the quantum system with the macroscopic measurement apparatus, the system becomes chaotic and the combination of imperfect knowledge of the exact physical state of the full interacting system with the time averaging present in measurement necessitates a statistical mechanical treatment of the quantum system. In giving such a treatment, one finds that due to the adiabatic invariance of $|\psi_n|^2$, the system approximately behaves as an eigenstate of the interaction Hamiltonian operator and which eigenstate is selected is effectively random from the perspective of the experiment, with the probability of selecting state $|n\rangle$ given by $|\psi_n|^2$ at the time the measurement interaction begins. This is the Born rule. We will describe how it is applied in more detail in detail in subsection 1.5.5 and how it arises from statistical mechanics in section 2.4. While the projective intensities sometimes manifest as probabilities, that is a derived result and not their only use, so I believe for clarity (and so they are not interpreted as probabilities in inappropriate contexts) they deserve their own name.

It is an unfortunate result of history that while almost every quantity in classical mechanics is named very intuitively and accurately (force, energy, power), many quantities in quantum mechanics are traditionally named poorly, or even misleadingly. As a result, we will make several other particular language choices to avoid confusion. For example, a quantum system with a state vector $|\psi\rangle$ that lives in a Hilbert space of dimension $N$ we will always call an $N$ dimensional quantum system (because that is the dimension of the

Hilbert space) whereas one is often likely to hear it called an "$N$ state quantum system". The "$N$ state" description is literally incorrect as there are an infinite number of different state vectors belonging to an $N$ dimensional Hilbert space. The "$N$ state" description is used because any vector in the Hilbert space can be written as a linear combination of at most $N$ linearly independent terms; it is just an inconsistent use of the term. There will never be confusion between the Hilbert space dimension of a system in question and how many spatial dimensions we are considering because we will always say spatial dimension when referring to spatial dimensions. We will also always call $\hat{H}$ the Hamiltonian operator of the system and never just the Hamiltonian. The Hamiltonian of the system is $\langle\psi|\hat{H}|\psi\rangle$, as we have established. However, it is a very common bad habit in other sources to incorrectly simply call $\hat{H}$ the Hamiltonian.

## Exercises

**Problem 1.2: Functions of an Operator.** We will often use the notion of a function of an operator. For example, $f(\hat{H})$. For an analytic function $f(x)$ and an operator $\hat{\mathcal{O}}$, the notation $f(\hat{\mathcal{O}})$ is defined by applying the Taylor series expansion of $f(x)$ to $\hat{\mathcal{O}}$. That is:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \implies f(\hat{\mathcal{O}}) = \sum_{n=0}^{\infty} a_n \hat{\mathcal{O}}^n.$$

Consider an analytic function $f(x)$ and a normal operator $\hat{\mathcal{O}}$. Let $|n\rangle$ be an orthonormal basis of eigenvectors of $\hat{\mathcal{O}}$ and let $\mathcal{O}_n$ be the eigenvalue corresponding to $|n\rangle$. Show that:

$$f(\hat{\mathcal{O}})|n\rangle = f(\mathcal{O}_n)|n\rangle, \qquad f(\hat{\mathcal{O}}) = \sum_n f(\mathcal{O}_n)|n\rangle\langle n|.$$

This second expression is used to define the notation $f(\hat{\mathcal{O}})$ when $f$ is not analytic but $\hat{\mathcal{O}}$ is normal.

**Problem 1.3: ⋆ Silver Bullet Solution to the Schrödinger equation.** Consider a quantum system with a time independent Hamiltonian operator $\hat{H}$. Assume that you have solved the eigenvector problem for $\hat{H}$ and thus identified an orthonormal basis of eigenvectors $|n\rangle$ of $\hat{H}$ with vector $|n\rangle$ having eigenvalue $E_n$. Show that the solution to the Schrödinger equation for arbitrary initial conditions can be written as:

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}t\hat{H}}|\psi(0)\rangle = \sum_n e^{-\frac{i}{\hbar}tE_n}|n\rangle\langle n|\psi(0)\rangle.$$

Thus, solving the dynamics of a time translationally symmetric quantum system is equivalent to solving the eigenvector/eigenvalue problem for the Hamiltonian operator. Notice that the angular frequency of oscillation of the component of $|\psi\rangle$ in the $|n\rangle$ direction is $\frac{E_n}{\hbar}$. Show as well that the total energy of the system is:

$$\mathrm{E} = \sum_n E_n |\langle n|\psi\rangle|^2.$$

**Problem 1.4: Practice Diagonalizing Hamiltonians.** Consider a two dimensional Hilbert space $\mathbb{H}$ and a quantum system with Hamiltonian operator:

$$\hat{H} = \begin{pmatrix} v_z & v_x - iv_y \\ v_x + iv_y & -v_z \end{pmatrix}$$

for some real constants $v_x, v_y, v_z$. This notation is defined in a basis we will label $|u_1\rangle$ and $|u_2\rangle$ as:

$$|u_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |u_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \langle u_1| = \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad \langle u_2| = \begin{pmatrix} 0 & 1 \end{pmatrix}.$$

(a) Show that Hamiltonian operator given is self-adjoint.

(b) Let $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$. Show that the eigenvalues of the Hamiltonian operator are $\pm v$.

(c) For the rest of this question it is useful to treat $(v_x, v_y, v_z)$ as a vector in a 3 dimensional space and to introduce standard spherical coordinates so that $v_x = v\sin\theta\cos\varphi$, $v_y = v\sin\theta\sin\varphi$, $v_z = v\cos\theta$. Let $|+\rangle$ be the eigenvector of $\hat{H}$ with eigenvalue $v$ and $|-\rangle$ be the eigenvector of $\hat{H}$ with the eigenvalue $-v$. Show that they can be chosen to be:

$$|+\rangle = \begin{pmatrix} e^{-\frac{i\varphi}{2}}\cos\frac{\theta}{2} \\ e^{\frac{i\varphi}{2}}\sin\frac{\theta}{2} \end{pmatrix}, \qquad |-\rangle = \begin{pmatrix} -e^{-\frac{i\varphi}{2}}\sin\frac{\theta}{2} \\ e^{\frac{i\varphi}{2}}\cos\frac{\theta}{2} \end{pmatrix}.$$

(d) Show that in the $|\pm\rangle$ basis $\hat{H}$ can be written:

$$\hat{H} = v|+\rangle\langle+| - v|-\rangle\langle-|$$

and verify the orthogonality properties and completeness relation of the $|\pm\rangle$ basis.

(e) Show that the solution to the Schrödinger equation for arbitrary initial conditions is:

$$|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle \text{ where: } \hat{U}(t) = \cos\left(\frac{vt}{\hbar}\right)\mathbb{1} - \frac{i}{v}\sin\left(\frac{vt}{\hbar}\right)\hat{H}.$$

(f) Show by explicit calculation that $\hat{U}$ is unitary.

**Problem 1.5: Hermitian but Not Self-Adjoint.** Consider a system in 1 spatial dimension on an interval of length 1 so that the Hilbert space $\mathbb{H}$ is the space of complex valued functions $\psi(x)$ of a single real variable $x$ so that $\int_0^1 |\psi(x)|^2 dx$ exists and is finite. Define the operator $\hat{p}_1$ as the operator so that:

$$\langle x|\hat{p}_1|\psi\rangle = -i\frac{\partial\psi}{\partial x}$$

and so that the domain of $\hat{p}_1$ (meaning the set of $|\psi\rangle \in \mathbb{H}$ on which it is defined, written $D(\hat{p}_1)$) is the set of periodic functions (meaning $\psi(x) = \psi(x+1)$). Define the operator $\hat{p}_2$ as the operator so that:

$$\langle x|\hat{p}_2|\psi\rangle = -i\frac{\partial\psi}{\partial x}$$

and so that $D(\hat{p}_2)$ is the set of functions with Dirichlet boundary conditions (meaning $\psi(0) = \psi(1) = 0$).

(a) An operator $\hat{L}$ is said to be Hermitian if for all $|u\rangle \in D(\hat{L})$ and $|v\rangle \in D(\hat{L})$:

$$\langle u|(\hat{L}|v\rangle) = \left(\langle v|(\hat{L}|u\rangle)\right)^*.$$

Show by direct calculation that both $\hat{p}_1$ and $\hat{p}_2$ are Hermitian.

(b) The adjoint domain $D(\hat{L}^\dagger)$ is defined as the set of all $|v\rangle \in \mathbb{H}$ such that there exists a $|w\rangle \in \mathbb{H}$ satisfying:

$$\langle v|(\hat{L}|u\rangle) = \langle w|u\rangle$$

for all $|u\rangle \in D(\hat{L})$. Then, the adjoint $\hat{L}^\dagger$ is defined so that for such $|v\rangle$ and $|w\rangle$:

$$\hat{L}^\dagger|v\rangle = |w\rangle.$$

Show that $D(\hat{p}_1^\dagger)$ is the set of periodic functions and so is equal to $D(\hat{p}_1)$. As well, show that:

$$\langle x|\hat{p}_1^\dagger|v\rangle = -i\frac{\partial v}{\partial x}.$$

Thus, $\hat{p}_1$ is Hermitian and has the same domain as its adjoint, i.e. $\hat{p}_1$ is self-adjoint ($\hat{p}_1 = \hat{p}_1^\dagger$). Show that $D(\hat{p}_2^\dagger)$ is simply $\mathbb{H}$ and so is larger than $D(\hat{p}_2)$. Because $D(\hat{p}_2) \neq D(\hat{p}_2^\dagger)$, even though $\hat{p}_2$ is Hermitian, it is not self-adjoint.

**Problem 1.6: Hilbert's Hotel.** For a finite dimensional vector space, if an operator has a left inverse then it has a right inverse and they are equal. This is not true for infinite dimensional vector spaces. Consequently, in order for any operator $\hat{U}$ to be unitary it is not enough to check $\hat{U}^\dagger\hat{U} = \mathbb{1}$. One must also show $\hat{U}\hat{U}^\dagger = \mathbb{1}$. We reserve the operator inverse notation only for operators which have both a left and right inverse and whose left and right inverses coincide. Consider an infinite dimensional Hilbert space with orthonormal basis $|0\rangle, |1\rangle, |2\rangle, ...$ Define the shift operator:

$$\hat{S}|n\rangle = |n+1\rangle.$$

(a) Show that $\hat{S}$ is given by:

$$\hat{S} = \sum_{n=0}^{\infty} |n+1\rangle\langle n|.$$

(b) Show that $S^\dagger$ is the left inverse of $\hat{S}$ (and so you might suspect $\hat{S}$ to be unitary):

$$\hat{S}^\dagger\hat{S} = \mathbb{1}.$$

(c) Show that $\hat{S}^\dagger$ is not the right inverse of $\hat{S}$ and thus that $\hat{S}$ is not unitary:

$$\hat{S}\hat{S}^\dagger = \mathbb{1} - |0\rangle\langle 0|.$$

The classic story around this shift operator is to imagine an infinitely tall hotel that has one room on each floor and is completely full. The shift up operator, corresponding in this analogy to $\hat{S}$, makes each guest move up one floor to a new room. Because the hotel is infinite, every guest is able to go to that room above them while simultaneously leaving the room on floor 0 empty, free for a new guest to take. But, the shift down operator, corresponding in this analogy to $\hat{S}^\dagger$, makes each guest on floors 1 and up move down a room, and keeps the guest on floor 0 out of the hotel. Shifting up and then down ($\hat{S}^\dagger\hat{S}$) returns the hotel to its original state ($\mathbb{1}$) after rudely admitting and then ejecting a new guest. Shifting down and then up ($\hat{S}\hat{S}^\dagger$) kicks a guest who is already there out and leaves room 0 empty ($\mathbb{1} - |0\rangle\langle 0|$).

## 1.3 Wigner's Theorem

We found in theorem 1.2 that the Lagrangian for quantum mechanics takes a very specific form. Only a very restricted class of canonical transformations can preserve that form and thus only a limited set of transformations can be symmetries in quantum mechanics. The Schrödinger Hamiltonian depends directly on exactly 1 power of $|\psi\rangle$ and exactly 1 power of $\langle\psi|$ (such a function is said to be a sesquilinear form in $\psi^*, \psi$). If we rewrite that Hamiltonian in terms of the real and imaginary parts of $\psi$, then it becomes a real quadratic form in terms of them (which is where we started in (1.2.1)). Suppose we make some transformation of our field at a single time into a new field also at a single time $\psi \to \psi'$ so that the Schrödinger Hamiltonian has the same form in terms of $\psi'$ as it does in terms of $\psi$ (possibly with a new Hamiltonian operator). We could express that transformation in terms of how the real and imaginary parts of $\psi'$ depend on the real and imaginary parts of $\psi$. If the real and imaginary parts of $\psi'$ depend linearly on the real and imaginary parts of $\psi$, then the new Hamiltonian will automatically be a quadratic form in the real and imaginary parts of $\psi'$, just as we want. If the real and imaginary parts of $\psi'$ depend nonlinearly on the real and imaginary parts of $\psi$ then the new Hamiltonian will not be a quadratic form in the real and imaginary parts of $\psi'$. Instead, all of the nonlinear terms (if we were to express the real and imaginary parts of $\psi'$ in terms of a Taylor series on the real and imaginary parts of $\psi$) would generate cubic and higher terms in the Hamiltonian. But, these are precisely terms we are insensitive to due to our truncation in section 1.2. And so, we are not interested in them. Because of this, we need not consider nonlinear transformations of the real and imaginary parts of $|\psi\rangle$.

So, the only relevant transformations $\psi \to \psi'$ at a single time are those for which $\psi'$ is a linear combination of the real and imaginary parts of our untransformed $\psi$, or equivalently a linear combination of $\psi$ and $\psi^*$ (because the real and imaginary parts of a complex number can be written as linear combinations of the complex number and its conjugate):

$$|\psi'\rangle = \hat{U}|\psi\rangle + \hat{V}|\psi^*\rangle \tag{1.3.1}$$

where $\hat{U}$ and $\hat{V}$ are as yet unspecified linear operators. It will be useful to note that any complex number is by definition equal to its own transpose and therefore for any operator $\hat{\mathcal{O}}$ and vectors $|\psi\rangle$ and $|\phi\rangle$:

$$\langle\psi|\hat{\mathcal{O}}|\phi\rangle = \langle\psi|\hat{\mathcal{O}}|\phi\rangle^T = (|\phi\rangle)^T\hat{\mathcal{O}}^T(\langle\psi|^T) = (|\phi\rangle^*)^\dagger\hat{\mathcal{O}}^T(\langle\psi|^*)^\dagger = \langle\phi^*|\hat{\mathcal{O}}^T|\psi^*\rangle. \tag{1.3.2}$$

By definition $\langle\psi|\psi\rangle = 1$ and so to preserve this:

$$\langle\psi'|\psi'\rangle = \left(\langle\psi|\hat{U}^\dagger + \langle\psi^*|\hat{V}^\dagger\right)\left(\hat{U}|\psi\rangle + \hat{V}|\psi^*\rangle\right) \tag{1.3.3}$$

$$= \langle\psi|\left(\hat{U}^\dagger\hat{U} + \hat{V}^T\hat{V}^*\right)|\psi\rangle + \langle\psi^*|\hat{V}^\dagger\hat{U}|\psi\rangle + \langle\psi|\hat{V}^T\hat{U}^*|\psi^*\rangle \tag{1.3.4}$$

$$\stackrel{!}{=} \langle\psi|\psi\rangle \qquad \forall|\psi\rangle \tag{1.3.5}$$

$$\implies \langle\psi|\left(\hat{U}^\dagger\hat{U} + \hat{V}^T\hat{V}^* - \mathbb{1}\right)|\psi\rangle + \langle\psi^*|\hat{V}^\dagger\hat{U}|\psi\rangle + \langle\psi|\hat{V}^T\hat{U}^*|\psi^*\rangle = 0. \tag{1.3.6}$$

Let us put (1.3.6) in components by using a real orthonormal basis $|n\rangle$ and introducing a few quantities:

$$|\psi\rangle = \sum_n \psi_n|n\rangle, \quad \hat{A} = \hat{U}^\dagger\hat{U} + \hat{V}^T\hat{V}^* - \mathbb{1}, \quad \hat{B} = \hat{V}^\dagger\hat{U}, \quad A_{nm} = \langle n|\hat{A}|m\rangle, \quad B_{nm} = \langle n|\hat{B}|m\rangle. \tag{1.3.7}$$

Then, define $F(\psi, \psi^*)$ by:

$$F(\psi,\psi^*) = \langle\psi|\hat{A}|\psi\rangle + \langle\psi^*|\hat{B}|\psi\rangle + \langle\psi|\hat{B}^*|\psi^*\rangle = \sum_{nm}\left(A_{nm}\psi_n^*\psi_m + B_{nm}\psi_n\psi_m + B_{nm}^*\psi_n^*\psi_m^*\right). \tag{1.3.8}$$

(1.3.6) requires $F$ to vanish fro all $\psi$. So, all of its derivatives must vanish for all $\psi$ as well. In particular:

$$\frac{\partial^2 F}{\partial\psi_n\partial\psi_m} = B_{nm} + B_{mn} = 0 \implies \hat{B}^T = -\hat{B}. \tag{1.3.9}$$

Thus, $\hat{B}$ must be antisymmetric. Then by (1.3.2):

$$\langle\psi^*|\hat{B}|\psi\rangle = \langle\psi^*|\hat{B}^T|\psi\rangle = -\langle\psi^*|\hat{B}|\psi\rangle \implies \langle\psi^*|\hat{B}|\psi\rangle = 0. \tag{1.3.10}$$

Further requiring the derivatives of $F$ to vanish:

$$\frac{\partial^2 F}{\partial\psi_n^*\psi_m} = A_{nm} = 0 \implies \hat{A} = 0 \implies F = 0. \tag{1.3.11}$$

Thus, we must have:

$$\hat{U}^\dagger\hat{U} + \hat{V}^T\hat{V}^* = \mathbb{1}, \qquad \hat{V}^\dagger\hat{U} = -\hat{U}^T\hat{V}^*. \tag{1.3.12}$$

Now let us transform everything in our Lagrangian, even allowing time to be transformed:

$$\mathcal{L}' = i\hbar\langle\psi'|\dot{\psi}'\rangle - \langle\psi'|\hat{H}'|\psi'\rangle \tag{1.3.13}$$

$$\begin{aligned}
=& i\hbar\frac{dt}{dt'}\left(\langle\psi|\hat{U}^\dagger\hat{U}|\dot{\psi}\rangle + \langle\psi|\hat{U}^\dagger\dot{\hat{U}}|\psi\rangle + \langle\psi^*|\hat{V}^\dagger\hat{U}|\dot{\psi}\rangle + \langle\psi^*|\hat{V}^\dagger\dot{\hat{U}}|\psi\rangle\right) \\
&+ i\hbar\frac{dt}{dt'}\left(\langle\dot{\psi}|\hat{V}^T\hat{U}^*|\psi^*\rangle + \langle\psi|\hat{U}^\dagger\dot{\hat{V}}|\psi^*\rangle + \langle\dot{\psi}|\hat{V}^T\hat{V}^*|\psi\rangle + \langle\psi|\dot{\hat{V}}^T\hat{V}^*|\psi\rangle\right) \\
&- \langle\psi|\hat{U}^\dagger\hat{H}'\hat{U}|\psi\rangle - \langle\psi^*|\hat{V}^\dagger\hat{H}'\hat{U}|\psi\rangle - \langle\psi|\hat{U}^\dagger\hat{H}'\hat{V}|\psi^*\rangle - \langle\psi|\hat{V}^T\hat{H}'^T\hat{V}^*|\psi\rangle.
\end{aligned} \tag{1.3.14}$$

Let $\sigma$ be the sign of $\frac{dt'}{dt}$ so that $\frac{dt'}{dt} = \sigma|\frac{dt'}{dt}|$. Under a canonical transformation in the Lagrangian, the new Lagrangian is given by a Jacboian factor $|\frac{dt}{dt'}|$ times the original Lagrangian plus a total time derivative (Orig. (1.9.101)). So, in order for our transformed Lagrangian to be of the same form as our untransformed Lagrangian we must have, for some function $f(t, |\psi\rangle, \langle\psi|)$:

$$\mathcal{L}' = \sigma\frac{dt}{dt'}\left(i\hbar\langle\psi|\dot{\psi}\rangle - \langle\psi|\hat{H}|\psi\rangle + \frac{\partial f}{\partial t} + \sum_n\frac{\partial f}{\partial\psi_n}\dot{\psi}_n + \sum_n\frac{\partial f}{\partial\psi_n^*}\dot{\psi}_n^*\right). \tag{1.3.15}$$

For convenience of notation, define:

$$\frac{\partial f}{\partial|\psi\rangle} = \sum_n\frac{\partial f}{\partial\psi_n}\langle n|, \qquad \frac{\partial f}{\partial\langle\psi|} = \sum_n\frac{\partial f}{\partial\psi_n^*}|n\rangle \tag{1.3.16}$$

$$\implies \mathcal{L}' = \sigma\frac{dt}{dt'}\left(i\hbar\langle\psi|\dot{\psi}\rangle - \langle\psi|\hat{H}|\psi\rangle + \frac{\partial f}{\partial t} + \frac{\partial f}{\partial|\psi\rangle}|\dot{\psi}\rangle + \langle\dot{\psi}|\frac{\partial f}{\partial\langle\psi|}\right). \tag{1.3.17}$$

We can then match terms from (1.3.14) to (1.3.17). If $t'$ has any dependence on $|\psi\rangle$ or $\langle\psi|$ then there are nonlinear terms $\dot{\psi}$ in (1.3.17). There are no such terms to match against in (1.3.14) and therefore $t'$ must be a function of $t$ only. Matching the $\langle\dot{\psi}|$ terms in both equations:

$$\frac{\partial f}{\partial\langle\psi|} = i\hbar\sigma\hat{V}^T\hat{V}^*|\psi\rangle + i\hbar\sigma\hat{V}^T\hat{U}^*|\psi^*\rangle. \tag{1.3.18}$$

In components and making use of (1.3.12):

$$\frac{\partial^2 f}{\partial\psi_n^*\partial\psi_m^*} = \langle n|\hat{V}^T\hat{U}^*|m\rangle \implies \langle n|\hat{V}^T\hat{U}^*|m\rangle = \langle m|\hat{V}^T\hat{U}^*|n\rangle \implies \hat{V}^T\hat{U}^* = (\hat{V}^T\hat{U}^*)^T = \hat{U}^\dagger\hat{V} \tag{1.3.19}$$

$$\implies V^\dagger\hat{U} = -\hat{U}^T\hat{V}^* = -\left(\hat{V}^TU^*\right)^\dagger = -\left(\hat{U}^\dagger\hat{V}\right) = -\hat{V}^\dagger\hat{U} \implies \hat{V}^\dagger\hat{U} = 0. \tag{1.3.20}$$

With this (1.3.18) becomes:

$$\frac{\partial f}{\partial\langle\psi|} = i\hbar\sigma\hat{V}^T\hat{V}^*|\psi\rangle \implies f(t,\psi,\psi^*) = i\hbar\sigma\langle\psi|\hat{V}^T\hat{V}^*|\psi\rangle + g(t,\psi). \tag{1.3.21}$$

Next matching the $|\dot{\psi}\rangle$ terms between (1.3.14) and (1.3.17):

$$\frac{\partial g}{\partial|\psi\rangle}(t,\psi) = i\hbar\sigma\langle\psi|\left(\hat{U}^\dagger\hat{U} - \hat{V}^T\hat{V}^* - \sigma\mathbb{1}\right). \tag{1.3.22}$$

The left hand side is a function of $\psi$ only. Therefore, its $\psi^*$ derivative must vanish. Taking the $\psi^*$ derivative

of both sides then gives:

$$\hat{U}^\dagger\hat{U} - \hat{V}^T\hat{V}^* - \sigma\mathbb{1} = 0. \tag{1.3.23}$$

Using (1.3.12) then gives:

$$\hat{U}^\dagger\hat{U} = \frac{1+\sigma}{2}\mathbb{1}, \qquad \hat{V}^\dagger\hat{V} = \frac{1-\sigma}{2}\mathbb{1}. \tag{1.3.24}$$

Returning this result, it follows that:

$$\frac{\partial g}{\partial|\psi\rangle} = 0 \implies g(t,\psi) = g(t). \tag{1.3.25}$$

So, $g(t)$ appears only in the transformed Lagrangian through the $\frac{\partial f}{\partial t}$ term in (1.3.17). $\frac{\partial g}{\partial t}$ does not depend on $\psi$ or $\psi^*$ and so there is no other term to match it against. Thus, it must be 0. As such, $g(t)$ must just be some constant number. However, because the surface term only enters the Lagrangian through its derivatives, we can choose such a constant to be 0, giving:

$$f(t,\psi,\psi^*) = i\hbar\sigma\langle\psi|\hat{V}^T\hat{V}^*|\psi\rangle. \tag{1.3.26}$$

Now let us do some case analysis. First, assume time is not reversed so that $\sigma = 1$. Then by (1.3.24):

$$\hat{U}^\dagger\hat{U} = \mathbb{1}, \qquad \hat{V}^\dagger\hat{V} = 0. \tag{1.3.27}$$

So, for all $|\psi\rangle$:

$$\langle\psi|\hat{V}^\dagger\hat{V}|\psi\rangle = 0 \implies \left\|\hat{V}|\psi\rangle\right\|^2 = 0 \implies \hat{V}|\psi\rangle = 0 \implies \hat{V} = 0. \tag{1.3.28}$$

The last step only following because this is true for all $|\psi\rangle$. Because $\hat{V}$ is 0, we then have:

$$|\psi'\rangle = \hat{U}|\psi\rangle. \tag{1.3.29}$$

Because a coordinate transformation is by definition invertible, $\hat{U}^{-1}$ must exist and by (1.3.27):

$$\hat{U}^{-1} = \hat{U}^\dagger. \tag{1.3.30}$$

So, $\hat{U}$ must be unitary. Then, $f = 0$ and the remaining terms in matching (1.3.14) to (1.3.17) give:

$$i\hbar\frac{d}{dt'}\langle\psi|\hat{U}^\dagger\dot{\hat{U}}|\psi\rangle - \langle\psi|\hat{U}^\dagger\hat{H}'\hat{U}|\psi\rangle = -\frac{dt}{dt'}\langle\psi|\hat{H}|\psi\rangle \quad \forall\ \psi \quad \implies \hat{U}^\dagger\hat{H}'\hat{U} = \frac{dt}{dt'}\left(\hat{H} + i\hbar\hat{U}^\dagger\dot{\hat{U}}\right) \tag{1.3.31}$$

Using the unitarity of $\hat{U}$ we can see that:

$$\hat{U}^\dagger\hat{U} = \mathbb{1} \implies \dot{\hat{U}}^\dagger\hat{U} + \hat{U}^\dagger\dot{\hat{U}} = 0 \implies \dot{\hat{U}} = -\hat{U}\dot{\hat{U}}^\dagger\hat{U} \implies \hat{H}' = \frac{dt}{dt'}\hat{U}\left(\hat{H} - i\hbar\frac{\partial}{\partial t}\right)\hat{U}^\dagger. \tag{1.3.32}$$

Now let us analyze the case in which time is reversed so that $\sigma = -1$. Then by (1.3.24):

$$\hat{U}^\dagger\hat{U} = 0, \qquad \hat{V}^\dagger\hat{V} = \mathbb{1}. \tag{1.3.33}$$

So, for all $|\psi\rangle$:

$$\langle\psi|\hat{U}^\dagger\hat{U}|\psi\rangle = 0 \implies \left\|\hat{U}|\psi\rangle\right\|^2 = 0 \implies \hat{U}|\psi\rangle = 0 \implies \hat{U} = 0. \tag{1.3.34}$$

Again with the last step following because this was true for all $|\psi\rangle$. Because $\hat{U}$ is 0, we then have:

$$|\psi'\rangle = \hat{V}|\psi^*\rangle. \tag{1.3.35}$$

A coordinate transformation is by definition invertible so $\hat{V}^{-1}$ must exist and by (1.3.33) we see that:

$$\hat{V}^{-1} = \hat{V}^\dagger \tag{1.3.36}$$

and so $\hat{V}$ is also unitary. We have $f = -i\hbar\langle\psi|\psi\rangle$ and finishing matching (1.3.14) to (1.3.17) gives:

$$i\hbar\frac{dt}{dt'}\langle\psi|\dot{\hat{V}}^T\hat{V}^*|\psi\rangle - \langle\psi|\hat{V}^T\hat{H}'^T\hat{V}^*|\psi\rangle = \frac{dt}{dt'}\langle\psi|\hat{H}|\psi\rangle \quad \forall\psi \quad \implies \hat{H}' = \left|\frac{dt}{dt'}\right|\hat{V}\left(\hat{H}^T + i\hbar\frac{\partial}{\partial t}\right)\hat{V}^\dagger. \tag{1.3.37}$$

These two cases together give us:

**Theorem 1.3: Wigner's Theorem**

Any canonical transformation in quantum mechanics at a single time in which time is not reversed is given by some unitary operator $\hat{U}$ under which $|\psi\rangle$ and $\hat{H}$ transform as:

$$|\psi'\rangle = \hat{U}|\psi\rangle, \qquad \hat{H}' = \frac{dt}{dt'}\hat{U}\left(\hat{H} - i\hbar\frac{\partial}{\partial t}\right)\hat{U}^\dagger. \tag{1.3.38}$$

As well, no surface term can be added to the action under such a transformation. Any canonical transformation in quantum mechanics at a single time in which time is reversed is given by some unitary operator $\hat{V}$ under which $|\psi\rangle$ and $\hat{H}$ transform as:

$$|\psi'\rangle = \hat{V}|\psi^*\rangle, \qquad \hat{H}' = \left|\frac{dt}{dt'}\right|\hat{V}\left(\hat{H}^T + i\hbar\frac{\partial}{\partial t}\right)\hat{V}^\dagger. \tag{1.3.39}$$

As well, the surface term added to the action must be $f = -i\hbar\langle\psi|\psi\rangle$ under such a transformation.

Importantly, in the above the surface term $f = -i\hbar\langle\psi|\psi\rangle$ should not be set to $f = -i\hbar$ by using $\langle\psi|\psi\rangle = 1$ in the Lagrangian, because $\langle\psi|\psi\rangle = 1$ is actually a constraint on the initial conditions. In order to use it as a constraint in the Lagrangian, you are required first to add it as a Lagrange multiplier. If you do not use it in the Lagrangian and instead simply require it of the initial conditions used when actually solving the equations of motion, then dynamically it will be preserved.

When we studied time reversal as a subset of canonical transformations in the Orig., we identified that all time reversing canonical transformations can be understood as a composition of a time nonreversing (orthochronous) canonical transformation together with a simple time reversal. In a simple time reversal, $t' = -t$, $Q' = Q$. Consequently we found $P' = -P$ and $\mathcal{H}' = \mathcal{H}$. These considerations from Orig. chapter 1 were done assuming real valued canonical variables. In section 1.2 each component of $|\psi\rangle$ was defined so that a previous canonical coordinate was its real part (up to a constant multiplier) and the corresponding previous canonical momentum was its imaginary part. So, under a simple time reversal the real part should be unchanged but the imaginary part inverted, which is precisely complex conjugation. Therefore, the quantum mechanical special case of a simple time reversal is the transformation:

$$|\psi'\rangle = |\psi^*\rangle. \tag{1.3.40}$$

Careful readers may object to the notation $|\psi^*\rangle$ because it requires a meaningful notion of taking the complex conjugate of the coefficients of $\psi$ without also taking the Hermitian adjoint of the basis elements. This is in general a basis dependent procedure, therefore ambiguous unless a natural basis is specified. This is okay, one just needs to understand that $|\psi^*\rangle$ is defined with respect to some basis. This is not surprising as the notion of a simple time reversal is coordinate dependent. (A simple time reversal in one set of coordinates is not a simple time reversal in another set of canonical coordinates.) In fact this property is essential because we need all time reversing transformations to be related to any designated simple time reversal by an orthochronous transformation, which is precisely a change of coordinates/basis.

While we motivated considering only transformations so that $\psi'$ depends linearly on $\psi$ and $\psi^*$ using the structure of the Schrödinger Hamiltonian, it is possible with significant extra effort to prove rigorously that this linearity is forced upon us by other structures as well. For example, we can also prove the results of this section by instead requiring intensities $|\langle\psi_1|\psi_2\rangle|$ to be preserved instead of the form of the Hamiltonian. This is reasonable because such intensities are generally directly experimentally measurable. For a detailed proof of Wigner's theorem starting from this perspective, see [62].

For now our interest will be nonrelativistic quantum mechanics, in which the only transformations of interest that change time are time translation and time reversal, both of which map the wavefunction at a single time to itself at a single other time, consistent with the hypotheses of this section. Later we will consider Lorentz boosts which do **NOT** satisfy the single time assumption. Consider, for example, a Lorentz scalar $\psi(t, \vec{x})$. Under Lorentz boost by $\vec{v}$ about the origin then:

$$\psi'(t', \vec{x}') = \psi(t, \vec{x}) \implies \psi'(t, \vec{x}) = \psi(\gamma(t - \vec{v}\cdot\vec{x}), \vec{x} + (\gamma - 1)\vec{\imath}_v\vec{\imath}_v\cdot\vec{x} - \gamma\vec{v}t). \tag{1.3.41}$$

Plainly, this mixes $|\psi\rangle$ at many different times and so isn't a map from the Hilbert space at one time to the Hilbert space at another, avoiding the conditions of Wigner's theorem entirely. Consequently, this map may not be unitarity. We will return to this when we study relativistic quantum mechanics and quantum field theory.

## Exercises

**Problem 1.7: ⋆ Time Evolution Operator.** The Schrödinger equation for given initial conditions has a solution which exists and that solution is unique (for a finite dimensional Hilbert space this is easily proved because the Schrödinger equation is linear; for infinite dimensional Hilbert spaces it is more difficult to prove but still true). Because the Schrödinger equation is linear, the solution at time $t$ is related to the solution at time $t_0$ linearly. Thus, there exists a time operator $\hat{U}(t, t_0)$ defined so that:

$$|\psi(t)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle.$$

(a) Show that the time evolution operator satisfies the identity condition:

$$\hat{U}(t_0, t_0) = \mathbb{1},$$

the composition condition:

$$\hat{U}(t, t_1)\hat{U}(t_1, t_0) = \hat{U}(t, t_0),$$

and the operator-Schrödinger equation:

$$i\hbar\frac{\partial\hat{U}(t, t_0)}{\partial t} = \hat{H}(t)\hat{U}(t, t_0).$$

(b) Using the operator-Schrödinger equation, show that $\hat{U}(t, t_0)$ is unitary for all $t$. That is, show that:

$$\hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = \mathbb{1}.$$

(c) Using the composition condition, show that $\hat{U}(t, t_0)$ also satisfies:

$$i\hbar\frac{\partial\hat{U}(t, t_0)}{\partial t_0} = -\hat{U}(t, t_0)\hat{H}(t_0).$$

Because $\hat{U}$ is unitary and gives all quantum mechanical time evolution, by Wigner's theorem we see that once again time evolution is just a particular canonical transformation.

**Problem 1.8: Hamilton-Jacobi Solution to the Schrödinger Equation.** Solve the Hamilton-Jacobi equation for the Schrödinger equation to show that its on-shell action is given by:

$$S_c(t_f, t_0, \psi_f, \psi_0) = \langle\pi_0|\hat{U}^\dagger(t_f, t_0)|\psi_f\rangle - \langle\pi_0|\psi_0\rangle$$

in terms of the time evolution operator $\hat{U}(t, t_0)$. Verify all of the Hamilton-Jacobi conditions for this function. Funnily, it happens to take the numerical value 0.

**Problem 1.9: A Hint of Decoherence.** Consider a time dependent Hamiltonian operator $\hat{H}(t)$ so that at time $t$ it has orthonormal basis of eigenvectors $|n(t)\rangle$ and eigenvalues $E_n(t)$:

$$\hat{H}(t)|n(t)\rangle = E_n(t)|n(t)\rangle.$$

Assume none of the eigenvalues of $\hat{H}(t)$ are repeated (such an operator is called nondegenerate).

(a) Take the time derivative of the eigenvalue equation. Show, for $m \neq n$, that:

$$\dot{E}_n = \langle n|\dot{\hat{H}}|n\rangle, \qquad \langle m|\dot{n}\rangle = \frac{\langle m|\dot{\hat{H}}|n\rangle}{E_n - E_m}$$

(b) Let $\psi_n(t) = \langle n|\psi(t)\rangle$. Show that these components evolve in time according to:

$$\dot{\psi}_n = -i\frac{E_n}{\hbar}\psi_n + \sum_{m\neq n}\frac{\langle n|\dot{\hat{H}}|m\rangle}{E_n - E_m}\psi_m.$$

(c) Let $p_n(t) = |\psi_n|^2$ and $\varphi_n(t)$ be defined so that $\psi_n(t) = \sqrt{p_n(t)}e^{i\varphi_n(t)}$. Further, let $V_{nm}(t)$ and $\phi_{nm}(t)$ be real valued functions so that:

$$\langle n|\dot{\hat{H}}|m\rangle = V_{nm}e^{i\phi_{nm}}.$$

Using that $\hat{H}$ is self-adjoint, show that $V_{nm} = V_{mn}$ and $\phi_{nm} = -\phi_{mn}$. Show that $p_n(t)$ and $\varphi_n(t)$ evolve according to:

$$\dot{p}_n = \sum_{m\neq n}\frac{2V_{nm}}{E_n - E_m}\sqrt{p_n p_m}\cos\left(\varphi_m - \varphi_n + \phi_{nm}\right)$$

$$\dot{\varphi}_n = -\frac{E_n}{\hbar} + \sum_{m\neq n} \frac{V_{nm}}{E_n - E_m}\sqrt{\frac{p_m}{p_n}}\sin\left(\varphi_m - \varphi_n + \phi_{nm}\right).$$

Explain how these equations show that if $\hat{H}(t)$ changes slowly with time compared to the periods $\tau_{mn} = \frac{2\pi\hbar}{|E_m - E_n|}$ then $p_n$ is approximately constant and $\varphi_n$ varies very quickly with time. Explain how this establishes further the adiabatic invariance of $|\langle n|\psi\rangle|^2$. The fact that these phases oscillate very rapidly in many physical scenarios causes them to be difficult to access experimentally while the slow evolution of the $p_n$ causes them to be directly experimentally meaningful.

## 1.4 Continuous Transformations in Quantum Mechanics

This section is based on [25]. As was true in classical physics, we are most often interested in continuous transformations. A continuous transformation is necessarily continuously connected to the identity transformation and of course the identity transformation does not reverse time. So, due to Wigner's theorem, in quantum mechanics every continuous transformation at a single time is represented by a unitary transformation:

$$|\psi'\rangle = \hat{U}|\psi\rangle, \qquad \hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = \mathbb{1}. \tag{1.4.1}$$

From Noether's theorem, the generator $G$ of a continuous transformation is given by:

$$G = \langle\pi|\mathscr{D}|\psi\rangle - \mathcal{H}\mathscr{D}t - \mathscr{D}f \tag{1.4.2}$$

where $\mathscr{D}$ represents the derivative at the identity with respect to the transformation parameter $\varepsilon$ and $f$ is the surface function added to the action by the transformation. As established in Wigner's theorem, $f = 0$ and so $\mathscr{D}f = 0$. If we consider transformations which leave time untransformed, $\mathscr{D}t = 0$ and so:

$$G = \langle\pi|\mathscr{D}|\psi\rangle = i\hbar\langle\psi|\mathscr{D}|\psi\rangle. \tag{1.4.3}$$

Now using (1.4.1):

$$\mathscr{D}|\psi\rangle = \left.\frac{d|\psi'\rangle}{d\varepsilon}\right|_{\varepsilon=0} = \left.\frac{d(\hat{U}|\psi\rangle)}{d\varepsilon}\right|_{\varepsilon=0} = \left.\frac{d\hat{U}}{d\varepsilon}\right|_{\varepsilon=0}|\psi\rangle = (\mathscr{D}\hat{U})|\psi\rangle. \tag{1.4.4}$$

Thus, $G$ becomes:

$$G = i\hbar\langle\psi|(\mathscr{D}\hat{U})|\psi\rangle \tag{1.4.5}$$

and so a generator in quantum mechanics is defined by some linear operator $\hat{G} = i\hbar(\mathscr{D}\hat{U})$. Let us see what unitarity implies near the identity:

$$\frac{d}{d\varepsilon}\left(\hat{U}^\dagger\hat{U}\right)\Big|_{\varepsilon=0} = \left.\frac{d\mathbb{1}}{d\varepsilon}\right|_{\varepsilon=0} \implies \left.\left(\frac{d\hat{U}^\dagger}{d\varepsilon}\hat{U} + \hat{U}^\dagger\frac{d\hat{U}}{d\varepsilon}\right)\right|_{\varepsilon=0} = 0 \tag{1.4.6}$$

$$\implies (\mathscr{D}\hat{U})^\dagger\mathbb{1} + \mathbb{1}(\mathscr{D}\hat{U}) = 0 \implies (\mathscr{D}\hat{U})^\dagger = -\mathscr{D}\hat{U} \implies \hat{G}^\dagger = -i\hbar(\mathscr{D}\hat{U})^\dagger = i\hbar(\mathscr{D}\hat{U}) = \hat{G}. \tag{1.4.7}$$

Thus, every generator of continuous transformations of the wavefunction in quantum mechanics is of the form:

$$G = \langle\psi|\hat{G}|\psi\rangle \tag{1.4.8}$$

where $\hat{G}$ is a self-adjoint operator. One tends to also call $\hat{G}$ the generator of the transformation, because there is a one-to-one correspondence between $G$ and $\hat{G}$, however this is a bit of abuse of language. Many sources call both these $G$ and $\hat{G}$ "observables" rather than generators. I will avoid this terminology as it is overly metaphysical and seems to suggest that all experimentally observable quantities are of this form and that all quantities of this form are experimentally observable, neither of which is true. (For examples, the parameters of the Hamiltonian, such as a particle mass, can be extracted from experimental observables and so are observable but not of this form and there are many quantities one could write down of this form which are not gauge-invariant and so not observable in that they aren't even experimentally well-defined.) From the self-adjointness of $\hat{G}$ it immediately follows that $G$ is real valued:

$$G^* = \left(\langle\psi|\hat{G}|\psi\rangle\right)^* = \left(\langle\psi|\hat{G}|\psi\rangle\right)^\dagger = \langle\psi|\hat{G}^\dagger|\psi\rangle = \langle\psi|\hat{G}|\psi\rangle = G. \tag{1.4.9}$$

How let us consider the Poisson bracket corresponding to the Schrödinger Hamiltonian in a generic

orthonormal basis $|n\rangle$, recalling (1.2.84). It takes the form:

$$\{A,B\} = \sum_n \left( \frac{\partial A}{\partial \psi_n}\frac{\partial B}{\partial \pi_n} - \frac{\partial B}{\partial \psi_n}\frac{\partial A}{\partial \pi_n} \right) = \frac{1}{i\hbar}\sum_n \left( \frac{\partial A}{\partial \psi_n}\frac{\partial B}{\partial \psi_n^*} - \frac{\partial B}{\partial \psi_n}\frac{\partial A}{\partial \psi_n^*} \right). \tag{1.4.10}$$

Let us use this to see how $|\psi\rangle$ evolves under the transformation generated by $G$ using Orig. (1.4.1) (which told us that for any quantity $A$ under a continuous transformation with parameter $\varepsilon$ by generator $G$, $\frac{dA}{d\varepsilon} = \{A, G\}$):

$$\frac{\partial |\psi\rangle}{\partial \psi_n} = \frac{\partial}{\partial \psi_n}\sum_m \psi_m |m\rangle = |n\rangle, \qquad \frac{\partial |\psi\rangle}{\partial \psi_n^*} = \frac{\partial}{\psi_n^*}\sum_m \psi_m |m\rangle = 0 \tag{1.4.11}$$

$$\frac{\partial G}{\partial \psi_n} = \frac{\partial \langle\psi|}{\partial \psi_n}\hat{G}|\psi\rangle + \langle\psi|\hat{G}\frac{\partial |\psi\rangle}{\partial \psi_n} = \left(\frac{\partial |\psi\rangle}{\partial \psi_n^*}\right)^\dagger \hat{G}|\psi\rangle + \langle\psi|\hat{G}|n\rangle = \langle\psi|\hat{G}|n\rangle, \qquad \frac{\partial G}{\partial \psi_n^*} = \left(\frac{\partial G}{\partial \psi_n}\right)^\dagger = \langle n|\hat{G}|\psi\rangle \tag{1.4.12}$$

$$\implies \frac{d|\psi\rangle}{d\varepsilon} = \{|\psi\rangle, G\} = -\frac{i}{\hbar}\hat{G}|\psi\rangle \tag{1.4.13}$$

This allows us to compute the $n^{\text{th}}$ derivative of $|\psi\rangle$ with respect to $\varepsilon$:

$$\frac{d^n|\psi\rangle}{d\varepsilon^n} = \left(\frac{d}{d\varepsilon}\right)^{n-1}\left(-\frac{i}{\hbar}\hat{G}|\psi\rangle\right) = -\frac{i}{\hbar}\hat{G}\frac{d^{n-1}|\psi\rangle}{d\varepsilon^{n-1}} = \ldots = \left(-\frac{i}{\hbar}\hat{G}\right)^n |\psi\rangle \tag{1.4.14}$$

and so applying the Taylor expansion for $|\psi\rangle$ in terms of $\varepsilon$:

$$|\psi'\rangle = |\psi(\varepsilon)\rangle = \sum_{n=0}\frac{\varepsilon^n}{n!}\left(\frac{d^n|\psi\rangle}{d\varepsilon^n}\right)\Big|_{\varepsilon=0} = \sum_{n=0}\frac{\varepsilon^n}{n!}\left(-\frac{i}{\hbar}\varepsilon\hat{G}\right)^n|\psi\rangle = e^{-\frac{i}{\hbar}\hat{G}}|\psi\rangle. \tag{1.4.15}$$

Comparing (1.4.15) with (1.4.1) we can see that the unitary transformation operator $\hat{U}$ is related to the transformation parameter $\varepsilon$ and the generating operator $\hat{G}$ by:

$$\hat{U} = e^{-\frac{i}{\hbar}\varepsilon\hat{G}}. \tag{1.4.16}$$

Thus, we have:

**Theorem 1.4: Continuous Transformations in Quantum Mechanics**

Any continuous transformation which leaves time unchanged in quantum mechanics is generated by a function $G$ of the form (with $\hat{G}$ a self-adjoint operator):

$$G = \langle\psi|\hat{G}|\psi\rangle \tag{1.4.17}$$

Under such a transformation with transformation parameter $\varepsilon$, $|\psi\rangle$ changes as:

$$|\psi'\rangle = e^{-\frac{i}{\hbar}\varepsilon\hat{G}}|\psi\rangle. \tag{1.4.18}$$

The exponential relationship between $\hat{G}$ and $\hat{U}$ is not incidental. Let us now establish an important fact: **for an operator $\hat{U}$, there exists a self-adjoint $\hat{F}$ so that $\hat{U} = e^{-i\hat{F}}$ if and only if $\hat{U}$ is unitary.** Proof: Suppose $\hat{U}$ can be written in the form $\hat{U} = e^{-i\hat{F}}$ for a self-adjoint $\hat{F}$. Then, because $\hat{F}$ is self-adjoint it is diagonalizable and we can identify an orthonormal eigenbasis $|n\rangle$ for $\hat{F}$. Automatically then by the definition of the operator exponential:

$$\hat{F} = \sum_n F_n|n\rangle\langle n|, \qquad \hat{U} = e^{-i\hat{F}} \implies \hat{U} = \sum_n e^{-iF_n}|n\rangle\langle n| \tag{1.4.19}$$

and the $F_n$ coefficients are real. With this, we immediately have that $\hat{U}$ is unitary:

$$\hat{U}\hat{U}^\dagger = \sum_{nn'} e^{-iF_n}e^{iF_{n'}}|n\rangle\langle n|n'\rangle\langle n'| = \sum_n |n\rangle\langle n| = \mathbb{1} \tag{1.4.20}$$

$$\hat{U}^\dagger\hat{U} = \sum_{nn'} e^{iF_n}e^{-iF_{n'}}|n\rangle\langle n|n'\rangle\langle n'| = \sum_n |n\rangle\langle n| = \mathbb{1}. \tag{1.4.21}$$

Now for the other direction. Suppose that $\hat{U}$ is unitary. Then, it is diagonalizable by the spectral theorem with some orthonormal eigenbasis $|n\rangle$. Let $u_n$ be the eigenvalues of $\hat{U}$:

$$\hat{U}|n\rangle = u_n|n\rangle. \tag{1.4.22}$$

Take the norm of both sides:

$$\|\hat{U}|n\rangle\|^2 = \|u_n|n\rangle\|^2 \implies \langle n|\hat{U}^\dagger\hat{U}|n\rangle = u_n^* u_n\langle n|n\rangle \implies |u_n|^2 = 1 \tag{1.4.23}$$

by using that $\hat{U}$ is unitary. So, for each $n$ we can define real numbers $F_n$ so that $u_n = e^{-iF_n}$. From these real numbers define $\hat{F}$ by:

$$\hat{F} = \sum F_n|n\rangle\langle n|. \tag{1.4.24}$$

Automatically it follows that:

$$\hat{U} = e^{-i\hat{F}} \tag{1.4.25}$$

as desired.■

Because the form $\langle\psi|\hat{G}|\psi\rangle$ will continue to appear often, it is useful to give it a name. For an operator $\hat{\mathcal{O}}$, the quantity $\langle\psi|\hat{\mathcal{O}}|\psi\rangle$ is called the **operator summary** or "operator expectation" of $\hat{\mathcal{O}}$. I will mostly avoid the term operator expectation because the word expectation invokes concepts from probability and statistics which are not intended in general. There is an emergent connection between operator summaries and statistical expectations that is experimentally important, but the mechanism behind that is statistical mechanics and not always appropriate. We can see from theorem 1.2 that even the Hamiltonian is the operator summary of a self-adjoint operator (the Hamiltonian operator). We also introduce a short hand notation $\langle\cdot\rangle$ for operator summaries:

$$G = \langle\hat{G}\rangle = \langle\psi|\hat{G}|\psi\rangle. \tag{1.4.26}$$

The Poisson bracket of two operator summaries has a very nice structure. Consider $A = \langle\hat{A}\rangle$ and $B = \langle\hat{B}\rangle$ for some operators $\hat{A}$ and $\hat{B}$. Then:

$$\{A, B\} = \{\langle\hat{A}\rangle, \langle\hat{B}\rangle\} = \frac{1}{i\hbar}\sum_n \left(\frac{\partial A}{\partial\psi_n}\frac{\partial B}{\partial\psi_n^*} - \frac{\partial B}{\partial\psi_n}\frac{\partial A}{\partial\psi_n^*}\right) \tag{1.4.27}$$

$$= \frac{1}{i\hbar}\sum_n \Big(\langle\psi|\hat{A}|n\rangle\langle n|\hat{B}|\psi\rangle - \langle\psi|\hat{B}|n\rangle\langle n|\hat{A}|\psi\rangle\Big) \tag{1.4.28}$$

$$= \frac{1}{i\hbar}\Big(\langle\psi|\hat{A}\hat{B}|\psi\rangle - \langle\psi|\hat{B}\hat{A}|\psi\rangle\Big) \tag{1.4.29}$$

$$= \langle\psi\frac{1}{i\hbar}[\hat{A}, \hat{B}]|\psi\rangle. \tag{1.4.30}$$

This argument is due to [44]. Hence, the Poisson bracket of two operator summaries is itself an operator summary. Furthermore, suppose $\hat{A}$ and $\hat{B}$ are self-adjoint. Then by (1.2.81) the Poisson bracket of two operator summaries of two self-adjoint operators is itself an operator summary of a self-adjoint operator.

Now suppose we have some symmetries of interest whose generators together form a Lie algebra. That is, we have a basis of generators $\tau_a$ whose Poisson brackets, for some constant ${f_{ab}}^c$ satisfy:

$$\{\tau_a, \tau_b\} = {f_{ab}}^c\tau_c. \tag{1.4.31}$$

It follows from theorem 1.4 that we must have self-adjoint operators $\hat{\tau}_a$ so that $\tau_a = \langle\hat{\tau}_a\rangle$. Then, by (1.4.30):

$$\{\langle\tau_a\rangle, \langle\tau_b\rangle\} = {f_{ab}}^c\langle\tau_c\rangle \implies \langle\frac{1}{i\hbar}[\hat{\tau}_a, \hat{\tau}_b]\rangle = {f_{ab}}^c\langle\tau_c\rangle \implies \langle\psi|\left([\hat{\tau}_a, \hat{\tau}_b] - i\hbar{f_{ab}}^c\hat{\tau}_c\right)|\psi\rangle = 0, \quad \forall|\psi\rangle \tag{1.4.32}$$

$$\implies [\hat{\tau}_a, \hat{\tau}_b] = i\hbar{f_{ab}}^c\hat{\tau}_c. \tag{1.4.33}$$

Hence, the Lie bracket structure of the generators reproduces itself on the generating operators with $\frac{1}{i\hbar}[\cdot,\cdot]$ replacing the Poisson bracket.

Suppose we have a wavefunction with internal degrees of freedom. Under many transformations of interest, the way that the transformation acts on $|\psi\rangle$ will be to transform the spatial dependence in a particular way and to transform the internal components of $|\psi\rangle$ in a completely independent way. That is, the continuous transformation $\hat{U}$ corresponding to the overall transformation often factors into a piece $\hat{U}_{\mathbb{S}}$ which acts on the spatial dependence of $|\psi\rangle$ and $\hat{U}_{\text{int}}$ which acts on the internal space for $|\psi\rangle$:

$$|\psi'\rangle = \hat{U}|\psi\rangle = \hat{U}_{\mathbb{S}}\hat{U}_{\text{int}}|\psi\rangle. \tag{1.4.34}$$

This is the form taken for a broad class of transformations we will be interested in. In order for $\hat{U}$ to be unitary, it immediately follows that $\hat{U}_{\mathbb{S}}$ and $\hat{U}_{\text{int}}$ must individually be unitary on their respective spaces. As

well, they commute with each other because they are operators on different vector spaces:

$$\hat{U}^{-1} = \hat{U}^\dagger \implies \hat{U}_{\mathbb{S}}^{-1}\hat{U}_{\text{int}}^{-1} = \hat{U}_{\mathbb{S}}^\dagger \hat{U}_{\text{int}}^\dagger \implies \hat{U}_{\mathbb{S}}^{-1} = \hat{U}_{\mathbb{S}}^\dagger, \qquad \hat{U}_{\text{int}}^{-1} = \hat{U}_{\text{int}}^\dagger. \tag{1.4.35}$$

Then, differentiating (1.4.34) with respect to the transformation parameter about the identity:

$$\frac{d|\psi\rangle}{d\varepsilon} = \left(\left(\mathscr{D}\hat{U}_{\mathbb{S}}\right)\mathbb{1} + \mathbb{1}\left(\mathscr{D}\hat{U}_{\text{int}}\right)\right)|\psi\rangle. \tag{1.4.36}$$

Thus, if we define:

$$\hat{G} = i\hbar(\mathscr{D}\hat{U}), \qquad \hat{G}_{\mathbb{S}} = i\hbar(\mathscr{D}\hat{U}_{\mathbb{S}}), \qquad \hat{G}_{\text{int}} = i\hbar(\mathscr{D}\hat{U}_{\text{int}}) \tag{1.4.37}$$

we can see that the overall generating operator $\hat{G}$ can be written as a sum of a generator on the spatial dependence of the wavefunction and a generator on the internal space of the wavefunction:

$$\hat{G}|\psi\rangle = \left(\hat{G}_{\mathbb{S}} + \hat{G}_{\text{int}}\right)|\psi\rangle, \qquad \forall|\psi\rangle \implies \hat{G} = \hat{G}_{\mathbb{S}} + \hat{G}_{\text{int}}. \tag{1.4.38}$$

If we have a group of transformations which all factorize in the above way, then let $\hat{\tau}_a$ be the generating operators for those transformations as identified in (1.4.33). Then, we must have:

$$\hat{\tau}_a = \hat{\tau}_a^{\mathbb{S}} + \hat{\tau}_a^{\text{int}}, \qquad [\hat{\tau}_a, \hat{\tau}_b] = i\hbar f_{ab}{}^c\hat{\tau}_c \tag{1.4.39}$$

$$\implies [\hat{\tau}_a^{\mathbb{S}}, \hat{\tau}_b^{\mathbb{S}}] + [\hat{\tau}_a^{\text{int}}, \hat{\tau}_b^{\text{int}}] = i\hbar f_{ab}{}^c\hat{\tau}_c^{\mathbb{S}} + i\hbar f_{ab}{}^c\hat{\tau}_c^{\text{int}} \tag{1.4.40}$$

where this last line follows because the internal and spatial generators necessarily commute with each other. Because the internal and spatial generators act on different spaces, these pieces must cancel independently and so we have:

$$[\hat{\tau}_a^{\mathbb{S}}, \hat{\tau}_b^{\mathbb{S}}] = i\hbar f_{ab}{}^c\hat{\tau}_c^{\mathbb{S}}, \qquad [\hat{\tau}_a^{\text{int}}, \hat{\tau}_b^{\text{int}}] = i\hbar f_{ab}{}^c\hat{\tau}_c^{\text{int}} \tag{1.4.41}$$

and so the Lie algebra structure reproduces itself independently on the spatial and internal spaces[5]. Together this gives us:

**Theorem 1.5: Quantum Poisson Brackets Theorem**

The Poisson bracket of two operator summaries is itself an operator summary given by:

$$\{\langle\hat{A}\rangle, \langle\hat{B}\rangle\} = \langle\frac{1}{i\hbar}[\hat{A}, \hat{B}]\rangle. \tag{1.4.42}$$

If $\hat{A}$ and $\hat{B}$ are self-adjoint then the Poisson bracket of their summaries is itself the operator summary of a self-adjoint operator.

In quantum mechanics, any Lie algebra of generators $\tau_a$ with structure constants $f_{ab}{}^c$ can be written in the form $\tau_a = \langle\hat{\tau}_a\rangle$ for some operators $\hat{\tau}_a$ which satisfy:

$$[\hat{\tau}_a, \hat{\tau}_b] = i\hbar f_{ab}{}^c\hat{\tau}_c. \tag{1.4.43}$$

If those generators correspond to transformations which factorize in how they act on the spatial and internal degrees of freedom of the wavefunction then the generating operators can be written as $\hat{\tau}_a = \hat{\tau}_a^{\mathbb{S}} + \hat{\tau}_a^{\text{int}}$ where the $\hat{\tau}_a^{\mathbb{S}}$ act only on the spatial dependence of the wavefunction and the $\hat{\tau}_a^{\text{int}}$ act only on the internal degrees of freedom. As well, the spatial and internal generating operators independently reproduce the Lie algebra as:

$$[\hat{\tau}_a^{\mathbb{S}}, \hat{\tau}_b^{\mathbb{S}}] = i\hbar f_{ab}{}^c\hat{\tau}_c^{\mathbb{S}}, \qquad [\hat{\tau}_a^{\text{int}}, \hat{\tau}_b^{\text{int}}] = i\hbar f_{ab}{}^c\hat{\tau}_c^{\text{int}} \tag{1.4.44}$$

From the Schrödinger equation, if we consider two wavefunctions which are related by multiplying by a constant phase $|\psi'\rangle = e^{i\varphi}|\psi\rangle$, we can see that they both satisfy the same equation of motion:

$$i\hbar\frac{d|\psi'\rangle}{dt} - \hat{H}|\psi'\rangle = e^{i\varphi}\left(i\hbar\frac{d|\psi\rangle}{dt} - \hat{H}|\psi\rangle\right) = 0 \tag{1.4.45}$$

Thus, this fixed phase difference is preserved for all time. As well, any operator summary is completely

[5]There is a subtlety that one could add a term proportional to the identity to one of these if one subtracts the same term from the other. However, this effect would cancel in the combined transformation and so can essentially be reabsorbed into the definition of $\hat{U}_{\mathbb{S}}$ and $\hat{U}_{\text{int}}$ so that no such term appears.

unaffected by this phase:

$$\langle\psi'|\hat{G}|\psi'\rangle = \langle\psi|e^{-i\varphi}\hat{G}e^{i\varphi}|\psi\rangle = \langle\psi|\hat{G}|\psi\rangle. \tag{1.4.46}$$

So, because *all* generators in quantum mechanics and *all* terms appearing in the Lagrangian are of this form, this overall phase drops out of the results of *everything* which could couple the wavefunction to other variables (such as electromagnetism) or any of its mechanical quantities such as energy, momentum, or angular momentum. Because it is through such quantities that we make measurements, this overall phase $\varphi$ has no experimental consequences. This is to say, **the wavefunction can be modified by an arbitrary overall constant phase with no physical effects**. This is unsurprising as such a phase could have been absorbed into our original definition of the components of the wavefunction in (1.2.3) without changing any of the arguments which followed. Because the wavefunction's normalization is set to 1 and the phase is arbitrary, any vector $|\phi\rangle \in \mathbb{H}$ which is proportional to $|\psi\rangle$ by a complex number: $|\phi\rangle = z|\psi\rangle$ produces all of the same physical predictions as $|\psi\rangle$, independent of the value of $z$. A family of vectors which are all mutually proportional to each other by a complex number is called a ray. Because of this you will often hear people who are being careful say "the state of a quantum system is represented by a ray in Hilbert space" as the ray corresponding to $|\psi\rangle$ is unique, not $|\psi\rangle$ itself. Later on we will generalize our approach to quantum mechanics in terms of a quantity known as the density matrix, in which meaningless junk like this overall phase doesn't appear. It is essential to understand that only the single overall phase of $|\psi\rangle$ is meaningless, the relative phases of the $\psi_n$ when $|\psi\rangle$ is written in a basis are extremely important.

It turns out further that $|\psi\rangle$ can be determined up to an overall phase (so the physical state, the ray to which $|\psi\rangle$ corresponds, can be determined) from the operator summaries of self-adjoint operators. To see how, consider a set of orthonormal basis vectors $|n\rangle$ and the operators $\hat{S}_{mn}$ and $\hat{A}_{mn}$ defined by:

$$\hat{S}_{mn} = \frac{|m\rangle\langle n| + |n\rangle\langle m|}{2} \implies \hat{S}_{mn}^\dagger = \frac{|n\rangle\langle m| + |m\rangle\langle n|}{2} = \hat{S}_{mn} \tag{1.4.47}$$

$$\hat{A}_{mn} = \frac{|m\rangle\langle n| - |n\rangle\langle m|}{2i} \implies \hat{A}_{mn}^\dagger = \frac{|n\rangle\langle m| - |m\rangle\langle n|}{-2i} = \hat{A}_{mn}. \tag{1.4.48}$$

Then, computing the operator summaries of these operators in this basis:

$$\langle\hat{S}_{mn}\rangle = \frac{\psi_m^*\psi_n + \psi_n^*\psi_m}{2}, \qquad \langle\hat{A}_{mn}\rangle = \frac{\psi_m^*\psi_n - \psi_n^*\psi_m}{2i} \implies \langle\hat{S}_{mn}\rangle + i\langle\hat{A}_{mn}\rangle = \psi_m^*\psi_n. \tag{1.4.49}$$

Take a particular value of $n$, $n_0$, such that $\langle\hat{S}_{n_0n_0}\rangle \neq 0$. Such an $n_0$ must exist because the diagonal elements of $\langle\hat{S}_{mn}\rangle$ are automatically real, nonnegative, and sum to 1 by definition:

$$\langle\hat{S}_{nn}\rangle = \psi_n^*\psi_n \implies \sum_n\langle\hat{S}_{nn}\rangle = \sum_n\psi_n^*\psi_n = \sum_n\langle\psi|n\rangle\langle n|\psi\rangle = \langle\psi|\psi\rangle = 1. \tag{1.4.50}$$

Choose the phase of $\psi$ so that $\psi_{n_0}$ is real and positive. Then,

$$\langle\hat{S}_{n_0n_0}\rangle = |\psi_{n_0}|^2 \implies \psi_{n_0} = \sqrt{\langle\hat{S}_{n_0n_0}\rangle} \tag{1.4.51}$$

$$\langle\hat{S}_{n_0n}\rangle + i\langle\hat{A}_{n_0n}\rangle = \psi_{n_0}^*\psi_n = \sqrt{\langle\hat{S}_{n_0n_0}\rangle}\psi_n. \tag{1.4.52}$$

So the wavefunction can be fully reconstructed from these summaries by:

$$|\psi\rangle = \sum_n \psi_n|n\rangle = \sum_n \left(\frac{\langle\hat{S}_{n_0n}\rangle + i\langle\hat{A}_{n_0n}\rangle}{\sqrt{\langle\hat{S}_{n_0n_0}\rangle}}\right)|n\rangle. \tag{1.4.53}$$

A complete set of these generators is enough to reconstruct $\psi$ and thus determine the physical state! If a complete set of generators can be measured experimentally (which is typically approximately true, in the sense that it is possible to measure an arbitrarily large number and approximate the full series), it follows that **the wavefunction (viewed as a ray in Hilbert space) is observable.** Reconstructing the quantum state of a system systematically through operator summaries is known as **quantum tomography.** A very careful reader will notice that only a subset of the operator summaries $\langle\hat{S}_{mn}\rangle$ and $\langle\hat{A}_{mn}\rangle$ were sufficient to determine $|\psi\rangle$ and may worry about the possibility that if a different subset had been chosen a different $|\psi\rangle$ would have been found. Later we will realize that interacting quantum systems necessitate representing the state of the system by a more general object known as the density matrix. It is possible that a given set of operator summaries may be incompatible with the existence of a $|\psi\rangle$, and in that case a density matrix is necessary. However, for describing a single matter particle at a time with a definite known state as we are committing to in this chapter, a wavefunction description is sufficient. Typical quantum measurements

are typically highly destructive, in the sense of majorly affecting the state of the system of interest while measuring it, due to the necessary strength of the interaction coupling. Consequently, state reconstruction is typically a statistical process. I.e. typically one must prepare a system in (approximately) the same initial state many times, measure it many times, to progressively determine the values of the operator summaries necessary for quantum tomography.

### Exercises

**Problem 1.10: ⋆ Ehrenfest's Theorem and the Quantum Noether Theorem.** Let $G = \langle\psi|\hat{G}(t)|\psi\rangle$ for some self-adjoint operator $\hat{G}$ be a generator of interest.

(a) Using the Schrödinger equation, show that:

$$\frac{d\langle\hat{G}\rangle}{dt} = \langle\frac{\partial\hat{G}}{\partial t}\rangle + \langle\frac{1}{i\hbar}[\hat{G},\hat{H}]\rangle.$$

This result is known as Ehrenfest's Theorem. You should be able to see that it is a special case of Orig. (1.4.22). Because of this result, it is sometimes a useful shorthand to introduce the convective operative derivative $\frac{D}{Dt}$ as:

$$\frac{D\hat{G}}{Dt} = \frac{\partial\hat{G}}{\partial t} + \frac{1}{i\hbar}[\hat{G},\hat{H}].$$

(b) Consider the continuous transformation generated by $G$ with parameter $\varepsilon$. Show using Noether's theorem that regardless of whether this transformation is a symmetry:

$$\left.\frac{\partial\hat{H}'(\varepsilon,t')}{\partial\varepsilon}\right|_{\varepsilon=0} = \frac{D\hat{G}}{Dt}.$$

Show that the transformation is a symmetry if and only if:

$$\frac{D\hat{G}}{Dt} = 0.$$

This is the quantum mechanical special case of Noether's theorem.

**Problem 1.11: ⋆ Active vs. Passive Picture.** Just like with statistical mechanics and continuum mechanics, it is possible and sometimes useful to introduce a notion of an active picture and a passive picture. The passive picture (also known as the Schrödinger picture) is the one we have been using in which the operators are specified in how they act on Hilbert space and the wavefunction evolves in time according to the Schrödinger equation. Using the time evolution operator, one can see:

$$\langle\psi(t)|\hat{G}(t)|\psi(t)\rangle = \langle\psi(t_0)|\hat{U}^\dagger(t,t_0)\hat{G}(t)\hat{U}(t,t_0)|\psi(t_0)\rangle.$$

The active picture operator $\hat{G}_A(t,t_0)$ relative to the time $t_0$ is then defined so that:

$$\hat{G}_A(t,t_0) = \hat{U}^\dagger(t,t_0)\hat{G}(t)\hat{U}(t,t_0).$$

Because operator summaries are the quantities of interest in quantum mechanics and these active picture operators can be used to compute all operator summaries using the constant state vector $|\psi(t_0)\rangle$ (for a fixed $t_0$), the evolution of these operators contains all of the time evolution of quantum mechanics.

(a) Show that the active picture operators satisfy:

$$\frac{d\hat{G}_A}{dt} = \hat{U}^\dagger\frac{D\hat{G}}{Dt}\hat{U}.$$

(b) Suppose that $\hat{G}$ does not explicitly depend on time. Show that:

$$\frac{d\hat{G}_A}{dt} = \frac{1}{i\hbar}[\hat{G}_A,\hat{H}_A].$$

This is then the natural "operator equation of motion". Comment on its similarity to Orig. (1.4.14). The activity picture is unfortunately commonly also named after a previously mentioned Nazi scientist.

## 1.5 Summary and Spread Coordinates

The argument of this section is based on [44, 45]. It is time to try to relate the Schrödinger Hamiltonian back to our previous point particle theories. When we model matter by point particles, our fundamental

degrees of freedom are the point's position $\vec{x}(t)$ and its conjugate linear momentum $\vec{p}(t)$. Now that we are considering a wave theory, we ought to expect that point-like behavior would be approximately valid whenever the that wave is contained in a region of space small enough to treat as point-like for our experimental purposes and when that whole wave packet is all moving together with itself at about the same overall velocity so that it remains essentially localized relative to our experimental resolution over time. In such an approximation, the point around which the wavefunction is essentially localized defines the classical position $\vec{x}_{\text{cl}}$. To translate that position defined by the center of the whole wave, one must then translate the whole wave. Thus, its conjugate linear momentum is the generator of translations of the entire wave which is by definition its total linear momentum. That is, the conjugate momentum $\vec{p}_{\text{cl}}$ to the classical position should be the total momentum carried by the wave packet.

We can directly compute $\vec{p}_{\text{cl}}$ using Noether's theorem. By definition, under a translation by $\vec{\varepsilon}$, a Dirac delta with peak at $\vec{x}$ becomes a Dirac delta with peak at $\vec{x}+\vec{\varepsilon}$. Let $\hat{T}(\vec{\varepsilon})$ to be the unitary operator corresponding to a translation by $\vec{\varepsilon}$:

$$\hat{T}|\vec{x}\rangle \stackrel{\text{def}}{=} |\vec{x}+\vec{\varepsilon}\rangle \implies \hat{T} = \int_{\mathbb{S}} |\vec{x}+\vec{\varepsilon}\rangle\langle\vec{x}| d^d\vec{x} \tag{1.5.1}$$

using the completeness relation. A short calculation confirms that $\hat{T}$ is unitary. Under $\hat{T}$, the translated wavefunction $\psi_\varepsilon(t,\vec{x})$ becomes:

$$\psi_\varepsilon(t,\vec{x}) = \langle\vec{x}|\psi_\varepsilon\rangle = \langle\vec{x}|\hat{T}|\psi\rangle = \langle\vec{x}|\hat{T}\int_{\mathbb{S}}|\vec{x}'\rangle\langle\vec{x}'|d^d\vec{x}'|\psi\rangle = \int_{\mathbb{S}}\langle\vec{x}|\vec{x}'+\vec{\varepsilon}\rangle\langle\vec{x}'|\psi\rangle d^d\vec{x}' \tag{1.5.2}$$

$$= \int_{\mathbb{S}}\delta(\vec{x}'+\vec{\varepsilon}-\vec{x})\psi(t,\vec{x}')d^d\vec{x}' = \psi(t,\vec{x}-\vec{\varepsilon}). \tag{1.5.3}$$

So, a translation simply shifts the argument of the wavefunction. Then, by theorem 1.4 and (1.4.5), the total linear momentum of the wave $\vec{p}_{\text{cl}}$ is given by the generator of translations of that wave as:

$$\vec{p}_{\text{cl}} = i\hbar\langle\psi|\left(\vec{\mathscr{D}}\hat{T}\right)|\psi\rangle = i\hbar\int_{\mathbb{S}}\langle\psi|\vec{x}\rangle\langle\vec{x}|\left(\vec{\mathscr{D}}\hat{T}\right)|\psi\rangle d^d\vec{x} \tag{1.5.4}$$

$$= i\hbar\int_{\mathbb{S}}\psi^\dagger(t,\vec{x})\left.\frac{d\psi_\varepsilon(\vec{x})}{d\vec{\varepsilon}}\right|_{\vec{\varepsilon}=0} d^d\vec{x} = i\hbar\int_{\mathbb{S}}\psi^\dagger\left.\frac{d\psi(t,\vec{x}-\vec{\varepsilon})}{d\vec{\varepsilon}}\right|_{\vec{\varepsilon}=0} d^d\vec{x} \tag{1.5.5}$$

$$= i\hbar\int_{\mathbb{S}}\psi^\dagger(-\nabla\psi)d^d\vec{x} = \langle\psi|\left(-i\hbar\int|\vec{x}\rangle\nabla\langle\vec{x}|d^d\vec{x}\right)|\psi\rangle. \tag{1.5.6}$$

Thus the total linear momentum of the wavefunction can be written as the operator summary of "the momentum operator" $\hat{\vec{p}}$ defined by:

$$\hat{\vec{p}} \stackrel{\text{def}}{=} -i\hbar\int_{\mathbb{S}}|\vec{x}\rangle\nabla\langle\vec{x}|d^d\vec{x} \implies \vec{p}_{\text{cl}} = \langle\psi|\hat{\vec{p}}|\psi\rangle. \tag{1.5.7}$$

We can see that this momentum operator is self-adjoint, as we know it should be from theorem 1.4, by considering its inner product between two functions $\phi$ and $\psi$:

$$\langle\phi|\hat{\vec{p}}^\dagger|\psi\rangle = \langle\phi|\left(-i\hbar\int_{\mathbb{S}}|\vec{x}\rangle\nabla\langle\vec{x}|d^d\vec{x}\right)^\dagger|\psi\rangle = \langle\phi|i\hbar\int_{\mathbb{S}}(\nabla|\vec{x}\rangle)\langle\vec{x}|d^d\vec{x}|\psi\rangle \tag{1.5.8}$$

$$= i\hbar\int_{\mathbb{S}}(\nabla\phi(\vec{x})^*)\psi(\vec{x})d^d\vec{x} = i\hbar\int_{\mathbb{S}}\left(\nabla(\phi^*\psi)-\phi^*\nabla\psi\right)d^d\vec{x}. \tag{1.5.9}$$

Functions $\phi$ and $\psi$ which are square-integrable (meaning $\langle\psi|\psi\rangle$ and $\langle\phi|\phi\rangle$ are finite, precisely the class of wavefunctions for which our theory is defined) must decay sufficiently quickly for $|\vec{x}|\to\infty$ for their normalization integrals to converge. As such, the total derivative term in the above integral becomes a vanishing surface term for large $|\vec{x}|$. So:

$$\langle\phi|\hat{\vec{p}}^\dagger|\psi\rangle = 0 - i\hbar\int_{\mathbb{S}}\phi^*\nabla\psi d^d\vec{x} = \langle\phi|\left(-i\hbar\int_{\mathbb{S}}|\vec{x}\rangle\nabla\langle\vec{x}|d^d\vec{x}\right)\psi\rangle = \langle\phi|\hat{\vec{p}}|\psi\rangle. \tag{1.5.10}$$

Because this holds for all such $\phi$ and $\psi$, we have:

$$\hat{\vec{p}}^\dagger = \hat{\vec{p}}. \tag{1.5.11}$$

With this momentum operator by theorem 1.4 we can express the translation operator as:

$$\hat{T} = e^{-\frac{i}{\hbar}\vec{\varepsilon}\cdot\hat{\vec{p}}}. \tag{1.5.12}$$

In this form we can see quite explicitly how the translation operator actually just reproduces Taylor's theorem for analytic functions in the position basis. Using the exponential form:

$$\psi_\varepsilon(t,\vec{x}) = \langle\vec{x}|\psi_\varepsilon\rangle = \langle\vec{x}|\hat{T}|\psi\rangle = \langle\vec{x}|e^{-\frac{i}{\hbar}\vec{\varepsilon}\cdot\hat{\vec{p}}}|\psi\rangle \tag{1.5.13}$$

$$= \langle\vec{x}|\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{-i}{\hbar}\right)^2(\vec{\varepsilon}\cdot\hat{\vec{p}})^n|\psi\rangle = \sum_{n=0}^{\infty}\frac{(-i)^n}{n!\hbar^n}\langle\vec{x}|(\vec{\varepsilon}\cdot\hat{\vec{p}})^n|\psi\rangle \tag{1.5.14}$$

Using[6] :

$$\langle\vec{x}|\hat{\vec{p}}|\psi\rangle = \langle\vec{x}|\left(-i\hbar\int_{\mathbb{S}}|\vec{x}'\rangle\nabla'\langle\vec{x}'|d^d\vec{x}'\right)|\psi\rangle = -i\hbar\int_{\mathbb{S}}\langle\vec{x}|\vec{x}'\rangle\nabla'\langle\vec{x}'|\psi\rangle d^d\vec{x}' \tag{1.5.15}$$

$$= -i\hbar\int_{\mathbb{S}}\delta(\vec{x}-\vec{x}')\nabla'\psi(t,\vec{x}')d^d\vec{x}' = -i\hbar\nabla\psi(t,\vec{x}) \tag{1.5.16}$$

we can evaluate the inner product above by applying this formula $n$ times:

$$\langle\vec{x}|(\vec{\varepsilon}\cdot\hat{\vec{p}})^n|\psi\rangle = \langle\vec{x}|(\vec{\varepsilon}\cdot\hat{\vec{p}})\left((\vec{\varepsilon}\cdot\hat{\vec{p}})^{n-1}|\psi\rangle\right) - i\hbar(\vec{\varepsilon}\cdot\nabla)(\langle\vec{x}|(\vec{\varepsilon}\cdot\hat{\vec{p}})^{n-1}|\psi\rangle) = ... = (-i\hbar)^n(\vec{\varepsilon}\cdot\nabla)^n\psi(t,\vec{x}). \tag{1.5.17}$$

Then, (1.5.14) becomes:

$$\psi_\varepsilon(t,\vec{x}) = \sum_{n=0}^{\infty}\frac{(-i)^n}{n!\hbar^n}(-i\hbar)^n(\vec{\varepsilon}\cdot\nabla)^n\psi(t,\vec{x}) = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}(\vec{\varepsilon}\cdot\nabla)^n\psi(t,\vec{x}). \tag{1.5.18}$$

This last expression is the formula for the Taylor expansion of $\psi(t,\vec{x}-\vec{\varepsilon})$ in powers of $\vec{\varepsilon}$ around $\vec{\varepsilon}=0$ and so from this we recover:

$$\psi_\varepsilon(t,\vec{x}) = \psi(t,\vec{x}-\vec{\varepsilon}) \tag{1.5.19}$$

exactly as expected.

There is another way we can arrive at the expression for the momentum operator. From our study of the Legendre transformation, we know that if we compute the right hand side of $P=\frac{\partial\mathcal{L}}{\partial\dot{Q}}$ with $Q$ as the amount we have translated $\psi$, then the conjugate momentum $P$ to that $Q$ will be the generator of translations of $\psi$. Explicitly in the position basis, the Schrödinger Lagrangian is:

$$\mathcal{L} = i\hbar\int_{\mathbb{S}}\psi^\dagger(t,\vec{x})\dot{\psi}(t,\vec{x})d^d\vec{x} - \int_{\mathbb{S}}\int_{\mathbb{S}}\psi^\dagger(t,\vec{x})\langle\vec{x}|\hat{H}|\vec{x}'\rangle\psi(t,\vec{x}')d^d\vec{x}d^d\vec{x}'. \tag{1.5.20}$$

Notice that this field theory may be non-local depending on the coefficients $\langle\vec{x}|\hat{H}|\vec{x}'\rangle$. We don't know at this point what to choose for $\hat{H}$ (we will return to that question soon) but for now the choice does not matter. If we make a time dependent translation of an untranslated wavefunction $\psi_0$:

$$\psi(t,\vec{x}) = \psi_0(t,\vec{x}-\vec{\varepsilon}(t)) \implies \dot{\psi}(t,\vec{x}) = \frac{\partial\psi_0}{\partial t}(t,\vec{x}-\vec{\varepsilon}(t)) - \dot{\vec{\varepsilon}}\cdot\nabla\psi_0(t,\vec{x}-\vec{\varepsilon}(t)) = \frac{\partial\psi}{\partial t} - \dot{\vec{\varepsilon}}\cdot\nabla\psi. \tag{1.5.21}$$

Here the partial time derivative of $\psi$ indicates differentiating the explicit time dependence but not $\vec{\varepsilon}$. Then, applying a time dependent translation to the Lagrangian gives:

$$\mathcal{L} = i\hbar\int_{\mathbb{S}}\psi^\dagger\dot{\psi}d^d\vec{x} - \int_{\mathbb{S}}\int_{\mathbb{S}}\psi^\dagger(t,\vec{x})\langle\vec{x}|\hat{H}|\vec{x}'\rangle\psi(t,\vec{x}')d^d\vec{x}d^d\vec{x}' \tag{1.5.22}$$

$$= i\hbar\int_{\mathbb{S}}\psi^\dagger\left(\frac{\partial\psi}{\partial t} - \dot{\vec{\varepsilon}}\cdot\nabla\psi\right)d^d\vec{x} - \int_{\mathbb{S}}\int_{\mathbb{S}}\psi^\dagger(t,\vec{x})\langle\vec{x}|\hat{H}|\vec{x}'\rangle\psi(t,\vec{x}')d^d\vec{x}d^d\vec{x}' \tag{1.5.23}$$

$$= -i\hbar\dot{\vec{\varepsilon}}\cdot\int_{\mathbb{S}}\psi^\dagger\nabla\psi d^d\vec{x} + i\hbar\int_{\mathbb{S}}\psi^\dagger\frac{\partial\psi}{\partial t}d^d\vec{x} - \int_{\mathbb{S}}\int_{\mathbb{S}}\psi^\dagger(t,\vec{x})\langle\vec{x}|\hat{H}|\vec{x}'\rangle\psi(t,\vec{x}')d^d\vec{x}d^d\vec{x}'. \tag{1.5.24}$$

The conjugate momentum to a translation is given by the derivative of the Lagrangian with respect to the velocity of translation (as always and in accordance with Orig. (1.7.12)):

$$\vec{p}_{\text{cl}} = \frac{\partial\mathcal{L}}{\partial\dot{\vec{\varepsilon}}} = -i\hbar\int_{\mathbb{S}}\psi^\dagger\nabla\psi d^d\vec{x} = \langle\psi|\hat{\vec{p}}|\psi\rangle \tag{1.5.25}$$

exactly as we had before!

Let us now diagonalize the momentum operator, since it is self-adjoint. That is, consider an eigenvector

[6] In many beginning quantum books one is likely to see the equation "$\hat{\vec{p}} = -i\hbar\nabla$", which is horrible notation. The left hand side is supposed to be an operator in the Dirac notation while the right hand side represents how that operator acts as a differential operator in the position basis. The correct version of this statement is $\langle\vec{x}|\hat{\vec{p}}|\psi\rangle = -i\hbar\nabla\psi$.

$|\vec{p}\rangle$ of the momentum operator $\hat{\vec{p}}$ which is labeled by a vector $\vec{p}$ intended to indicate its eigenvalue:

$$\boxed{\hat{\vec{p}}|\vec{p}\rangle = \vec{p}|\vec{p}\rangle.} \tag{1.5.26}$$

This notation is always a little hard to parse the first time you see it. Taking the inner product with $\langle\vec{x}|$ we have:

$$\langle\vec{x}|\hat{\vec{p}}|\vec{p}\rangle = \vec{p}\langle\vec{x}|\vec{p}\rangle \implies -i\hbar\nabla\langle\vec{x}|\vec{p}\rangle = \vec{p}\langle\vec{x}|\vec{p}\rangle \implies \nabla\ln(\langle\vec{x}|\vec{p}\rangle) = \frac{i}{\hbar}\vec{p} \implies \langle\vec{x}|\vec{p}\rangle = Ne^{\frac{i}{\hbar}\vec{x}\cdot\vec{p}} \tag{1.5.27}$$

where here $N$ is an arbitrary constant. Thus, the eigenvectors of the momentum operator are just complex exponentials! We can choose the normalization factor $N$ so that the $|\vec{p}\rangle$ vectors satisfy the same orthonormality condition as the $|\vec{x}\rangle$ vectors, using the fundamental Fourier integral:

$$\langle\vec{p}|\vec{p}'\rangle = \langle\vec{p}|\int_{\mathbb{S}}|\vec{x}\rangle\langle\vec{x}|d^d\vec{x}|\vec{p}'\rangle = \int_{\mathbb{S}}\langle\vec{p}|\vec{x}\rangle\langle\vec{x}|\vec{p}'\rangle d^d\vec{x} \tag{1.5.28}$$

$$= \int_{\mathbb{S}}\left(Ne^{\frac{i}{\hbar}\vec{x}\cdot\vec{p}}\right)^* Ne^{\frac{i}{\hbar}\vec{x}\cdot\vec{p}'}d^d\vec{x} = |N|^2\int_{\mathbb{S}}e^{\frac{i}{\hbar}\vec{x}\cdot(\vec{p}'-\vec{p})}d^d\vec{x} \tag{1.5.29}$$

$$= |N|^2(2\pi)^d\delta\left(\frac{\vec{p}-\vec{p}'}{\hbar}\right) = (2\pi\hbar)^d|N|^2\delta(\vec{p}-\vec{p}'). \tag{1.5.30}$$

By convention we choose the normalization $N = \frac{1}{(2\pi\hbar)^{\frac{d}{2}}}$ so that:

$$\boxed{\langle\vec{x}|\vec{p}\rangle = \frac{e^{\frac{i}{\hbar}\vec{x}\cdot\vec{p}}}{(2\pi\hbar)^{\frac{d}{2}}}, \qquad \langle\vec{p}|\vec{p}'\rangle = \delta(\vec{p}-\vec{p}').} \tag{1.5.31}$$

Consequently, the momentum eigenstates also satisfy the same completeness relation as the position eigenstates and the momentum operator admits a nice expression in terms of its eigenstates (letting $\mathbb{K}$ be the $\mathbb{R}^d$ vector space of all values of $\vec{p}$):

$$\int_{\mathbb{K}}|\vec{p}\rangle\langle\vec{p}|d^d\vec{p} = \int_{\mathbb{S}}\int_{\mathbb{K}}\int_{\mathbb{S}}|\vec{x}\rangle\langle\vec{x}|\vec{p}\rangle\langle\vec{p}|\vec{x}'\rangle\langle\vec{x}'|d^d\vec{x}d^d\vec{p}d^d\vec{x}' = \int_{\mathbb{S}}\int_{\mathbb{K}}\int_{\mathbb{S}}|\vec{x}\rangle\frac{e^{\frac{i}{\hbar}\vec{x}\cdot\vec{p}}}{(2\pi\hbar)^{\frac{d}{2}}}\frac{e^{-\frac{i}{\hbar}\vec{p}\cdot\vec{x}'}}{(2\pi\hbar)^{\frac{d}{2}}}\langle\vec{x}'|d^d\vec{x}d^d\vec{p}d^d\vec{x}' \tag{1.5.32}$$

$$= \int_{\mathbb{S}}\int_{\mathbb{S}}|\vec{x}\rangle\delta(\vec{x}-\vec{x}')\langle\vec{x}'|d^d\vec{x}d^d\vec{x}' = \mathbb{1}. \tag{1.5.33}$$

$$\hat{\vec{p}} = -i\hbar\int_{\mathbb{S}}|\vec{x}\rangle\nabla\langle\vec{x}|d^d\vec{x} = -i\hbar\int_{\mathbb{K}}\int_{\mathbb{S}}|\vec{x}\rangle\frac{\partial\langle\vec{x}|\vec{p}\rangle}{\partial\vec{x}}\langle\vec{p}|d^d\vec{x}d^d\vec{p} \tag{1.5.34}$$

$$= -i\hbar\int_{\mathbb{K}}\int_{\mathbb{S}}|\vec{x}\rangle\frac{i}{\hbar}\vec{p}\langle\vec{x}|\vec{p}\rangle\langle\vec{p}|d^d\vec{x}d^d\vec{p} = \int_{\mathbb{K}}\vec{p}|\vec{p}\rangle\langle\vec{p}|d^d\vec{p}. \tag{1.5.35}$$

Thus, there is a momentum eigenvector for each real vector $\vec{p}$ and these vectors are orthonormal and complete. This allows us to make a few nice identifications. For example, we can define the "momentum space wavefunction" $\widetilde{\psi}(t,\vec{p})$ analogously to our "position space wavefunction" $\psi(t,\vec{x})$ by:

$$\widetilde{\psi}(t,\vec{p}) = \langle\vec{p}|\psi(t)\rangle. \tag{1.5.36}$$

Using the completeness relation, this function just gives the expansion coefficients for $|\psi\rangle$ when written as a linear combination of the elements of the momentum basis:

$$|\psi\rangle = \int_{\mathbb{K}}|\vec{p}\rangle\langle\vec{p}|d^d\vec{p}|\psi\rangle = \int_{\mathbb{K}}\widetilde{\psi}(t,\vec{p})|\vec{p}\rangle d^d\vec{p}. \tag{1.5.37}$$

Let us see how the position and momentum space wavefunctions relate:

$$\widetilde{\psi}(t,\vec{p}) = \int_{\mathbb{S}}\langle\vec{p}|\vec{x}\rangle\langle\vec{x}|\psi\rangle d^d\vec{x} = \int_{\mathbb{S}}\frac{e^{-\frac{i}{\hbar}\vec{x}\cdot\vec{p}}}{(2\pi\hbar)^{\frac{d}{2}}}\psi(t,\vec{x})d^d\vec{x}. \tag{1.5.38}$$

The momentum space wavefunction is simply the Fourier transform of the position space wavefunction with the wavenumber $\vec{k} = \frac{\vec{p}}{\hbar}$! This is actually quite straightforward to understand. We found that in the position basis, the momentum operator acts as (or proportionally to) the spatial derivative and so its eigenvectors are exponential functions, the eigenfunctions of the derivative. The inverse Fourier transform exactly expresses a spatial function as a linear combination of complex exponentials and the Fourier transform gives the coefficients of linear combination. That is all that has happened by identifying the momentum basis; we have simply put the Fourier transform into Dirac notation!

Now that we understand the momentum basis a bit better, we can naturally define "translations" of the

momentum space in exactly the same way as we did for position. That is, define the operator $\hat{S}(\vec{\varepsilon})$ as the unitary operator which acts as:

$$\hat{S}(\vec{\varepsilon})|\vec{p}\rangle \overset{\text{def}}{=} |\vec{p}+\vec{\varepsilon}\rangle \implies \hat{S} = \int_{\mathbb{K}} |\vec{p}+\vec{\varepsilon}\rangle\langle\vec{p}|d^d\vec{p}. \tag{1.5.39}$$

Then, under a momentum translation the wavefunction $\psi_\varepsilon(t,\vec{x})$ becomes:

$$\psi_\varepsilon(t,\vec{x}) = \langle\vec{x}|\psi_\varepsilon(t)\rangle = \langle\vec{x}|\hat{S}|\psi(t)\rangle = \int_{\mathbb{K}}\int_{\mathbb{S}}\langle\vec{x}|\vec{p}+\vec{\varepsilon}\rangle\langle\vec{p}|\vec{x}'\rangle\langle\vec{x}'|\psi(t)\rangle d^d\vec{x}'\,d^d\vec{p} \tag{1.5.40}$$

$$= \int_{\mathbb{S}}\int_{\mathbb{K}} \frac{e^{\frac{i}{\hbar}\vec{x}\cdot(\vec{p}+\vec{\varepsilon})}}{(2\pi\hbar)^{\frac{d}{2}}}\frac{e^{-\frac{i}{\hbar}\vec{x}'\cdot\vec{p}}}{(2\pi\hbar)^{\frac{d}{2}}}\psi(t,\vec{x}')d^d\vec{p}\,d^d\vec{x}' = e^{\frac{i}{\hbar}\vec{\varepsilon}\cdot\vec{x}}\int_{\mathbb{S}}\delta(\vec{x}-\vec{x}')\psi(t,\vec{x}')d^d\vec{x}' = e^{\frac{i}{\hbar}\vec{\varepsilon}\cdot\vec{x}}\psi(t,\vec{x}). \tag{1.5.41}$$

Hence, a translation in momentum simply multiplies the wavefunction by a phase factor which is linear in position. We can see by direct calculation that this does in fact add momentum $\vec{\varepsilon}$:

$$\vec{p}'_{\text{cl}} = \langle\psi_\varepsilon|\hat{\vec{p}}|\psi_\varepsilon\rangle = \langle\psi_\varepsilon|\left(-i\hbar\int_{\mathbb{S}}|\vec{x}\rangle\nabla\langle\vec{x}|d^d\vec{x}|\right)\psi_\varepsilon\rangle = -i\hbar\int_{\mathbb{S}}\psi_\varepsilon^\dagger\nabla\psi_\varepsilon d^d\vec{x} \tag{1.5.42}$$

$$= -i\hbar\int_{\mathbb{S}}\psi^\dagger e^{-\frac{i}{\hbar}\vec{\varepsilon}\cdot\vec{x}}\nabla\left(e^{\frac{i}{\hbar}\vec{\varepsilon}\cdot\vec{x}}\psi\right)d^d\vec{x} = -i\hbar\int_{\mathbb{S}}\psi^\dagger\nabla\psi d^d\vec{x} + \vec{\varepsilon}\int_{\mathbb{S}}\psi^\dagger\psi d^d\vec{x} \tag{1.5.43}$$

$$= \langle\psi|\hat{\vec{p}}|\psi\rangle + \varepsilon\langle\psi|\psi\rangle = \vec{p}_{\text{cl}} + \vec{\varepsilon}. \tag{1.5.44}$$

Next we wonder, how should we extract the classical position? Whatever $\vec{x}_{\text{cl}}$ (the associated classical position to $|\psi\rangle$ should be) is, it must make a canonically conjugate pair with $\vec{p}_{\text{cl}}$. Notice that in general $Q'=P,\quad P'=-Q,\quad t'=t,\quad \mathcal{H}'=\mathcal{H}$ is a canonical transformation. That is to say the conjugate momentum to $P$ is actually $-Q$. Applying this, one could consider the Lagrangian $\mathcal{L}'(t,Q',\dot{Q}')$ gotten by Legendre transforming in these primed coordinates:

$$\mathcal{L}'(t,Q',\dot{Q}') = \dot{Q}'P' - \mathcal{H} \implies P' = \frac{\partial\mathcal{L}}{\partial\dot{Q}'}. \tag{1.5.45}$$

In the unprimed coordinates, this reads:

$$\mathcal{L}'(t,P,\dot{P}) = -\dot{P}Q - \mathcal{H} \implies -Q = \frac{\partial\mathcal{L}'}{\partial\dot{P}}. \tag{1.5.46}$$

So if we write the Lagrangian in terms of $P$ and $\dot{P}$, we can actually recover $Q$ from (1.5.46) in the same way we would usually get $P$ from a Lagrangian[7] written in terms of $Q$ and $\dot{Q}$. Before we derived the expression for $\vec{p}_{\text{cl}}$ by differentiating the Lagrangian with respect to a time dependent translation of position. This line of reasoning allows us to derive $-\vec{x}_{\text{cl}}$ by differentiating the Lagrangian with respect to a time dependent translation of momentum. Applying such a translation to a generic $\psi_0$:

$$\psi(t,\vec{x}) = e^{\frac{i}{\hbar}\vec{\varepsilon}\cdot\vec{x}}\psi_0(t,\vec{x}) \implies \dot{\psi} = \frac{i}{\hbar}\dot{\vec{\varepsilon}}\cdot\vec{x}e^{\frac{i}{\hbar}\vec{\varepsilon}\cdot\vec{x}}\psi_0(t,\vec{x}) + e^{\frac{i}{\hbar}\vec{\varepsilon}\cdot\vec{x}}\frac{\partial\psi_0}{\partial t}(t,\vec{x}) = \frac{i}{\hbar}\dot{\vec{\varepsilon}}\cdot\vec{x}\psi + \frac{\partial\psi}{\partial t}. \tag{1.5.47}$$

Then, applying this to the Lagrangian:

$$\mathcal{L} = i\hbar\int_{\mathbb{S}}\psi^\dagger\dot{\psi}d^d\vec{x} - \int_{\mathbb{S}}\int_{\mathbb{S}}\psi^\dagger(t,\vec{x})\langle\vec{x}|\hat{H}|\vec{x}'\rangle\psi(t,\vec{x}')d^d\vec{x}d^d\vec{x}' \tag{1.5.48}$$

$$= i\hbar\int_{\mathbb{S}}\psi^\dagger\left(\frac{i}{\hbar}\dot{\vec{\varepsilon}}\cdot\vec{x}\psi + \frac{\partial\psi}{\partial t}\right)d^d\vec{x} - \int_{\mathbb{S}}\int_{\mathbb{S}}\psi^\dagger(t,\vec{x})\langle\vec{x}|\hat{H}|\vec{x}'\rangle\psi(t,\vec{x}')d^d\vec{x}d^d\vec{x}' \tag{1.5.49}$$

$$= -\dot{\vec{\varepsilon}}\cdot\int_{\mathbb{S}}\psi^\dagger\vec{x}\psi d^d\vec{x} + i\hbar\int_{\mathbb{S}}\psi^\dagger\frac{\partial\psi}{\partial t}d^d\vec{x} - \int_{\mathbb{S}}\int_{\mathbb{S}}\psi^\dagger(t,\vec{x})\langle\vec{x}|\hat{H}|\vec{x}'\rangle\psi(t,\vec{x}')d^d\vec{x}d^d\vec{x}'. \tag{1.5.50}$$

And so, the classical position is given from (1.5.46) as:

$$\vec{x}_{\text{cl}} = -\frac{\partial\mathcal{L}}{\partial\dot{\vec{\varepsilon}}} = \int_{\mathbb{S}}\psi^\dagger\vec{x}\psi d^d\vec{x} = \int_{\mathbb{S}}\langle\psi|\vec{x}\rangle\vec{x}\langle\vec{x}|\psi\rangle d^d\vec{x} = \langle\psi|\int_{\mathbb{S}}\vec{x}|\vec{x}\rangle\langle\vec{x}|d^d\vec{x}|\psi\rangle. \tag{1.5.51}$$

Defining then the position operator $\hat{\vec{x}}$:

$$\hat{\vec{x}} = \int_{\mathbb{S}}\vec{x}|\vec{x}\rangle\langle\vec{x}|d^d\vec{x} \implies \vec{x}_{\text{cl}} = \langle\psi|\hat{\vec{x}}|\psi\rangle. \tag{1.5.52}$$

[7] One may be concerned about calling both of these quantities the Lagrangian. Notice that this new "Lagrangian" in terms of $P$ and $\dot{P}$ is actually related to the typical expression for $\mathcal{L}$ by adding a total time derivative and so there is nothing to worry about.

The position operator is immediately self-adjoint, because position values are real valued:

$$\hat{\vec{x}}^\dagger = \int_{\mathbb{S}} \vec{x}^* |\vec{x}\rangle\langle\vec{x}| d^d\vec{x} = \hat{\vec{x}}. \tag{1.5.53}$$

As well, we can see that the form in (1.5.52) is analogous to (1.5.35) for the momentum operator. Equation (1.5.52) is quite a reasonable expression for what we mean by the position operator as from (1.5.51):

$$\vec{x}_{\text{cl}} = \langle\psi|\hat{\vec{x}}|\psi\rangle = \int_{\mathbb{S}} \vec{x}|\psi|^2 d^d\vec{x} \tag{1.5.54}$$

and so the result is a weighted average position where the weight is $|\psi|^2$. Thus, when the wavefunction is essentially all located around a particular position, $|\psi|^2$ will only make significant contributions to this integral near that position and $\vec{x}_{\text{cl}}$ will coincide with that position, exactly as we intended with the meaning of classical position. Because $-\vec{x}_{\text{cl}}$ is the generator of momentum translations, by theorem 1.4 we can express the momentum translation operator as:

$$\hat{S}(\vec{\varepsilon}) = e^{\frac{i}{\hbar}\vec{\varepsilon}\cdot\hat{\vec{x}}}. \tag{1.5.55}$$

Now that we have identified the classical position and momentum within the wavefunction, we ought to check that they are in fact canonically conjugate in the way we intended. By the canonical brackets in Orig. (1.4.7) then we would like to check:

$$\{\vec{x}_{\text{cl}}, \vec{x}_{\text{cl}}\} \stackrel{?}{=} 0, \qquad \{\vec{x}_{\text{cl}}, \vec{p}_{\text{cl}}\} \stackrel{?}{=} \mathbb{1}, \qquad \{\vec{p}_{\text{cl}}, \vec{p}_{\text{cl}}\} \stackrel{?}{=} 0. \tag{1.5.56}$$

From theorem 1.5, these are equivalent to:

$$[\hat{\vec{x}}, \hat{\vec{x}}] \stackrel{?}{=} 0, \qquad [\hat{\vec{x}}, \hat{\vec{p}}] \stackrel{?}{=} i\hbar\mathbb{1}, \qquad [\hat{\vec{p}}, \hat{\vec{p}}] \stackrel{?}{=} 0. \tag{1.5.57}$$

So, let us compute these commutators. To do so, it is easiest to work in components and to consider these operators acting on an arbitrary vector $|\psi\rangle$:

$$[\hat{x}^a, \hat{x}^b]|\psi\rangle = (\hat{x}^a\hat{x}^b - \hat{x}^b\hat{x}^a)|\psi\rangle = \int_{\mathbb{S}} |\vec{x}\rangle\langle\vec{x}| d^d\vec{x} \left(\hat{x}^a\hat{x}^b - \hat{x}^b\hat{x}^a\right)|\psi\rangle \tag{1.5.58}$$

$$= \int_{\mathbb{S}} |\vec{x}\rangle \left(x^a x^b - x^b x^a\right) \langle\vec{x}|\psi\rangle d^d\vec{x} = \int_{\mathbb{S}} |\vec{x}\rangle 0 \langle\vec{x}|\psi\rangle d^d\vec{x} = 0. \quad \checkmark \tag{1.5.59}$$

$$[\hat{x}^a, \hat{p}_b]|\psi\rangle = (\hat{x}^a\hat{p}_b - \hat{p}_b\hat{x}^a)|\psi\rangle = \int |\vec{x}\rangle\langle\vec{x}| d^d\vec{x} \left(\hat{x}^a\hat{p}_b - \hat{p}_b\hat{x}^a\right)|\psi\rangle \tag{1.5.60}$$

$$= \int_{\mathbb{S}} |\vec{x}\rangle \left(x^a\langle\vec{x}|\hat{p}_b|\psi\rangle - \langle\vec{x}|\hat{p}_b\hat{x}^a|\psi\rangle\right) d^d\vec{x} = \int_{\mathbb{S}} |\vec{x}\rangle \left(-i\hbar x^a \frac{\partial\psi}{\partial x^b} + i\hbar \frac{\partial}{\partial x^b}\left(\langle\vec{x}|\hat{x}^a|\psi\rangle\right)\right) d^d\vec{x} \tag{1.5.61}$$

$$= \int_{\mathbb{S}} |\vec{x}\rangle \left(-i\hbar x^a \frac{\partial\psi}{\partial x^b} + i\hbar \frac{\partial}{\partial x^b}\left(x^a\psi\right)\right) d^d\vec{x} = \int_{\mathbb{S}} |\vec{x}\rangle \left(-i\hbar x^a \frac{\partial\psi}{\partial x^b} + i\hbar\delta^a_b\psi + i\hbar x^a \frac{\partial\psi}{\partial x^b}\right) d^d\vec{x} \tag{1.5.62}$$

$$= i\hbar\delta^a_b \int_{\mathbb{S}} |\vec{x}\rangle\langle\vec{x}|\psi\rangle d^d\vec{x} = i\hbar\delta^a_b|\psi\rangle. \quad \checkmark \tag{1.5.63}$$

$$[\hat{p}_a, \hat{p}_b]|\psi\rangle = (\hat{p}_a\hat{p}_b - \hat{p}_b\hat{p}_a)|\psi\rangle = \int_{\mathbb{K}} |\vec{p}\rangle\langle\vec{p}| d^d\vec{p} (\hat{p}_a\hat{p}_b - \hat{p}_b\hat{p}_a)|\psi\rangle \tag{1.5.64}$$

$$= \int_{\mathbb{K}} |\vec{p}\rangle (p_a p_b - p_b p_a)\langle\vec{p}|\psi\rangle d^d\vec{p} = \int_{\mathbb{K}} |\vec{p}\rangle 0 \langle\vec{p}|\psi\rangle d^d\vec{p} = 0. \quad \checkmark. \tag{1.5.65}$$

These computations should remind you of Orig. section 1.3. The commutation relations are exactly as suspected in (1.5.57) and thus the Poisson brackets are exactly as suspected in (1.5.56).

**Theorem 1.6: Classical Variables in Quantum Mechanics**

The position and momentum operators are self-adjoint operators defined by:

$$\hat{\vec{x}} = \int_{\mathbb{S}} \vec{x}|\vec{x}\rangle\langle\vec{x}| d^d\vec{x}, \qquad \hat{\vec{p}} = -i\hbar \int_{\mathbb{S}} |\vec{x}\rangle\nabla\langle\vec{x}| d^d\vec{x}. \tag{1.5.66}$$

They satisfy the canonical commutation relations:

$$[\hat{\vec{x}}, \hat{\vec{x}}] = 0, \qquad [\hat{\vec{x}}, \hat{\vec{p}}] = i\hbar\mathbb{1}, \qquad [\hat{\vec{p}}, \hat{\vec{p}}] = 0. \tag{1.5.67}$$

The quantities $\vec{x}_{\text{cl}} = \langle\hat{\vec{x}}\rangle$ and $\vec{p}_{\text{cl}} = \langle\hat{\vec{p}}\rangle$ always correspond to the spatial-average position and total linear momentum of the wavefunction and when the wavefunction is approximately localized they correspond to the position and momentum present in the classical theory of point matter. They always satisfy the canonical brackets:

$$\{\vec{x}_{\text{cl}}, \vec{x}_{\text{cl}}\} = 0, \qquad \{\vec{x}_{\text{cl}}, \vec{p}_{\text{cl}}\} = \mathbb{1}, \qquad \{\vec{p}_{\text{cl}}, \vec{p}_{\text{cl}}\} = 0. \tag{1.5.68}$$

### 1.5.1 Operator Spreads

We say that $\vec{x}_{\text{cl}}$ and $\vec{p}_{\text{cl}}$ correspond to classical particle-like position and momentum, but under what precise circumstances should we recover classical mechanics from our wave theory of matter? As we said, this wave should produce essentially classical dynamics for $\vec{x}_{\text{cl}}$ and $\vec{p}_{\text{cl}}$ whenever that wave is essentially localized in its position and the whole wave moves with an essentially well-defined velocity. Because the Legendre transformation acts as a way to trade in position and velocity dependence for position and momentum dependence, having an essentially well-defined position and velocity implies having an essentially well-defined position and momentum. Thus, we expect classical mechanics to be recovered for $\vec{x}_{\text{cl}}$ and $\vec{p}_{\text{cl}}$ whenever the wavefunction is not very spread out over space and simultaneously also not very spread out in momentum space, exactly like in Orig. problem 4.48. Let us study how to quantify such conditions.

Consider an operator $\hat{\mathcal{O}}$ and a normalized eigenstate $|\psi\rangle$ of that operator with eigenvalue $\lambda$. Then:

$$\hat{\mathcal{O}}|\psi\rangle = \lambda|\psi\rangle \implies \langle\psi|\hat{\mathcal{O}}|\psi\rangle = \lambda\langle\psi|\psi\rangle \implies \lambda = \langle\hat{\mathcal{O}}\rangle. \tag{1.5.69}$$

Further consider the nonlinear operator $\delta\hat{\mathcal{O}}$ defined by:

$$\delta\hat{\mathcal{O}} = \hat{\mathcal{O}} - \langle\psi|\hat{\mathcal{O}}|\psi\rangle\mathbb{1} \implies \delta\hat{\mathcal{O}}|\psi\rangle = \hat{\mathcal{O}}|\psi\rangle - \langle\hat{\mathcal{O}}\rangle|\psi\rangle = \lambda|\psi\rangle - \lambda|\psi\rangle = 0. \tag{1.5.70}$$

And so, $\delta\hat{\mathcal{O}}$ annihilates any normalized eigenstate of $\hat{\mathcal{O}}$. In fact, the converse is true as suppose instead that for a vector $|\phi\rangle$ the operator $\delta\hat{\mathcal{O}}$ annihilates that vector:

$$\delta\hat{\mathcal{O}}|\phi\rangle = 0 \implies \hat{\mathcal{O}}|\phi\rangle - \langle\phi|\mathcal{O}|\phi\rangle|\phi\rangle = 0 \implies \hat{\mathcal{O}}|\phi\rangle = \langle\phi|\hat{\mathcal{O}}|\phi\rangle|\phi\rangle. \tag{1.5.71}$$

Hence, $|\phi\rangle$ must be an eigenvector of $\hat{\mathcal{O}}$ with eigenvalue $\langle\phi|\hat{\mathcal{O}}|\phi\rangle$. It further follows by taking the inner product of both sides with $\langle\phi|$ that $\langle\phi|\phi\rangle = 1$[8]. And so, $\delta\hat{\mathcal{O}}$ annihilates a vector if and only if that vector is a normalized eigenstate of $\hat{\mathcal{O}}$. A vector in a normed vector space has norm 0 if and only if it is the 0 vector:

$$\||u\rangle\|^2 = \langle u|u\rangle = 0 \iff |u\rangle = 0 \tag{1.5.72}$$

and so the fact that $\delta\hat{\mathcal{O}}$ annihilates a vector if and only if that vector $|\psi\rangle$ is a normalized eigenstate of $\hat{\mathcal{O}}$ is equivalent in both directions to the condition that $\delta\hat{\mathcal{O}}|\psi\rangle$ has norm 0:

$$|\psi\rangle \text{ is a normalized eigenstate of } \hat{\mathcal{O}} \iff \delta\hat{\mathcal{O}}|\psi\rangle = 0 \iff \left\|\delta\hat{\mathcal{O}}|\psi\rangle\right\|^2 = \langle\psi|\delta\hat{\mathcal{O}}^\dagger\delta\hat{\mathcal{O}}|\psi\rangle = 0. \tag{1.5.73}$$

Thus, the quantity $\delta\mathcal{O}^2$ defined by:

$$\delta\mathcal{O}^2 = \langle\psi|\delta\hat{\mathcal{O}}^\dagger\delta\hat{\mathcal{O}}|\psi\rangle \tag{1.5.74}$$

for a general state $|\psi\rangle$ is 0 if and only if $|\psi\rangle$ is a normalized eigenstate of $\hat{\mathcal{O}}$. This expression is called the operator spread of $\hat{\mathcal{O}}$ in the state $|\psi\rangle$ and is generally written as:

$$\text{Spr}(\hat{\mathcal{O}}) = \delta\mathcal{O}^2 = \langle\psi|(\hat{\mathcal{O}}^\dagger - \langle\hat{\mathcal{O}}\rangle^*\mathbb{1})(\hat{\mathcal{O}} - \langle\hat{\mathcal{O}}\rangle\mathbb{1})|\psi\rangle. \tag{1.5.75}$$

Because the operator spread is expressible as the norm of a vector, it is non-negative. Further, if $\hat{\mathcal{O}}$ is normal then using the spectral expansion of the operator $\hat{\mathcal{O}}$:

$$\text{Spr}(\hat{\mathcal{O}}) = \sum_n |\lambda_n - \langle\hat{\mathcal{O}}\rangle|^2|\psi_n|^2. \tag{1.5.76}$$

When all of the non-zero $\psi_n$ correspond to eigenvalues $\lambda_n$ which are near each other, $\langle\hat{\mathcal{O}}\rangle$ will be near those $\lambda_n$ and all of the $|\lambda_n - \langle\hat{\mathcal{O}}\rangle|^2$ will be small, leading to a small spread. So, when $|\psi\rangle$ is nearly an eigenstate of $\hat{\mathcal{O}}$ the operator spread of $\hat{\mathcal{O}}$ will be nearly 0. The operator spread of $\hat{\mathcal{O}}$ essentially measures the extent to which $|\psi\rangle$ fails to be an eigenstate of $\hat{\mathcal{O}}$. When we want to consider a wavefunction which is localized

[8]There is a slight exception in the case that $\langle\phi|\hat{\mathcal{O}}|\phi\rangle = 0$ in which case no condition on normalization can be derived. The rest of the steps of reasoning presented here still follow, just with the word normalized removed. Ultimately, this case will not affect our final conclusion.

in position, we will concretely mean a wavefunction for which the operator spread of the position for that wavefunction is negligible (for the purposes at hand; how large of a spread will count as negligible will of course depend on experimental resolution). As well, a wavefunction which is localized in momentum will mean that it has negligible momentum operator spread.

An occasionally useful way of interpreting the operator spread comes from what it tells us about how far the operator summary is from the nearest eigenvalue of $\hat{\mathcal{O}}$. For a normal operator $\hat{\mathcal{O}}$ and fixed normalized $|\psi\rangle$ which is not necessarily an eigenvector of $\hat{\mathcal{O}}$, define the operator $\hat{V}_\psi$ by:

$$\hat{V}_\psi = (\hat{\mathcal{O}}^\dagger - \langle\psi|\hat{\mathcal{O}}|\psi\rangle^* \mathbb{1})(\hat{\mathcal{O}} - \langle\psi|\hat{\mathcal{O}}|\psi\rangle \mathbb{1}) - \mathrm{Spr}_\psi(\hat{\mathcal{O}})\mathbb{1} \tag{1.5.77}$$

where the operator spread given above is intended as being taken with respect to the vector $|\psi\rangle$. It immediately follows by computation that $[\hat{V}_\psi, \hat{V}_\psi^\dagger] = 0$ and so $\hat{V}_\psi$ is normal as well. Further, any eigenvector $|n\rangle$ of $\hat{\mathcal{O}}$ with eigenvalue $\lambda_n$ is also an eigenvector of $\hat{V}_\psi$ with eigenvalue $v_n$ given as:

$$\hat{V}_\psi|n\rangle = v_n|n\rangle \implies v_n = |\lambda_n - \langle\hat{\mathcal{O}}\rangle|^2 - \mathrm{Spr}_\psi(\hat{\mathcal{O}}). \tag{1.5.78}$$

So, all of the eigenvalues of $\hat{V}_\psi$ are real. Let $v_*$ be the least eigenvalue of $\hat{V}_\psi$ (or, more rigorously, the greatest real number less than or equal to $v_n$ for all $n$, i.e. the infimum of the spectrum of $\hat{V}_\psi$). Let $\lambda_*$ be the associated eigenvalue of $\hat{\mathcal{O}}$ corresponding to $v_*$:

$$v_* = |\lambda_* - \langle\hat{\mathcal{O}}\rangle|^2 - \mathrm{Spr}_\psi(\hat{\mathcal{O}}) \tag{1.5.79}$$

Then, we have:

$$\langle\psi|\hat{V}_\psi|\psi\rangle = \langle\psi|(v_*\mathbb{1} + \hat{V}_\psi - v_*\mathbb{1})|\psi\rangle = v_* + \sum_n (v_n - v_*)|\psi_n|^2. \tag{1.5.80}$$

Each term in the sum in the final equality is nonnegative, and so we have:

$$\langle\psi|\hat{V}_\psi|\psi\rangle \geq v_*. \tag{1.5.81}$$

Using the definitions of $\hat{V}_\psi$ and $\lambda_*$:

$$\langle\psi|\left((\hat{\mathcal{O}}^\dagger - \langle\psi|\hat{\mathcal{O}}|\psi\rangle^*\mathbb{1})(\hat{\mathcal{O}} - \langle\psi|\hat{\mathcal{O}}|\psi\rangle\mathbb{1}) - \mathrm{Spr}_\psi(\hat{\mathcal{O}})\mathbb{1}\right)|\psi\rangle \geq v_* \implies v_* \leq 0 \tag{1.5.82}$$

$$\implies |\lambda_* - \langle\hat{\mathcal{O}}\rangle| \leq \sqrt{\mathrm{Spr}(\hat{\mathcal{O}})}. \tag{1.5.83}$$

Hence, for any state $|\psi\rangle$ and normal operator $\hat{\mathcal{O}}$, there always exists an eigenvalue of $\hat{\mathcal{O}}$ within one square-root of the spread of $\hat{\mathcal{O}}$ from the summary of $\hat{\mathcal{O}}$. The operator spread tells you at most how far the operator summary is from an eigenvalue of the operator.

Another quantity which is very closely related to the operator spread but generalized to two operators is the operator correlation. For two operators $\hat{A}$ and $\hat{B}$, their correlation is defined by:

$$\mathrm{Cor}(\hat{A}, \hat{B}) = \langle\psi|\frac{(\hat{A}^\dagger - \langle\hat{A}\rangle^*\mathbb{1})(\hat{B} - \langle\hat{B}\rangle\mathbb{1}) + (\hat{B} - \langle\hat{B}\rangle\mathbb{1})(\hat{A}^\dagger - \langle\hat{A}\rangle^*\mathbb{1})}{2}|\psi\rangle. \tag{1.5.84}$$

One can see immediately that the correlation satisfies: $\mathrm{Cor}(\hat{A}, \hat{B}) = \mathrm{Cor}(\hat{B}, \hat{A})^*$. The operator spread of a normal operator is just a special case of the correlation:

$$\mathrm{Cor}(\hat{\mathcal{O}}, \hat{\mathcal{O}}) = \langle\psi|\frac{(\hat{\mathcal{O}}^\dagger - \langle\hat{\mathcal{O}}\rangle^*\mathbb{1})(\hat{\mathcal{O}} - \langle\hat{\mathcal{O}}\rangle\mathbb{1}) + (\hat{\mathcal{O}} - \langle\hat{\mathcal{O}}\rangle\mathbb{1})(\hat{\mathcal{O}}^\dagger - \langle\hat{\mathcal{O}}\rangle^*\mathbb{1})}{2}|\psi\rangle \tag{1.5.85}$$

$$= \langle\psi|(\hat{\mathcal{O}}^\dagger - \langle\hat{\mathcal{O}}\rangle^*\mathbb{1})(\hat{\mathcal{O}} - \langle\hat{\mathcal{O}}\rangle\mathbb{1})|\psi\rangle = \mathrm{Spr}(\hat{\mathcal{O}}). \tag{1.5.86}$$

All of this gives us:

**Theorem 1.7: Spectral Spread Theorem**

The operator correlation in a state $|\psi\rangle$ between two operators $\hat{A}$ and $\hat{B}$ is defined by:

$$\mathrm{Cor}(\hat{A}, \hat{B}) = \langle\psi|\frac{(\hat{A}^\dagger - \langle\hat{A}\rangle^*\mathbb{1})(\hat{B} - \langle\hat{B}\rangle\mathbb{1}) + (\hat{B} - \langle\hat{B}\rangle\mathbb{1})(\hat{A}^\dagger - \langle\hat{A}\rangle^*\mathbb{1})}{2}|\psi\rangle \tag{1.5.87}$$

and satisfies:

$$\mathrm{Cor}(\hat{B}, \hat{A}) = \mathrm{Cor}(\hat{A}, \hat{B})^*. \tag{1.5.88}$$

The operator spread in a state $|\psi\rangle$ of an operator $\hat{\mathcal{O}}$ is defined by:

$$\text{Spr}(\hat{\mathcal{O}}) = \text{Cor}(\hat{\mathcal{O}}, \hat{\mathcal{O}}) = \langle (\hat{\mathcal{O}}^\dagger - \langle \hat{\mathcal{O}} \rangle^* \mathbb{1})(\hat{\mathcal{O}} - \langle \hat{\mathcal{O}} \rangle \mathbb{1}) \rangle. \tag{1.5.89}$$

The operator spread is 0 if and only if $|\psi\rangle$ is a normalized eigenstate of $\hat{\mathcal{O}}$. The spread is non-negative and acts as a measure of the extent to which $|\psi\rangle$ fails to be an eigenstate of $\hat{\mathcal{O}}$. For any $|\psi\rangle$ and normal operator $\hat{\mathcal{O}}$, there is always an eigenvalue of $\hat{\mathcal{O}}$ within a distance of $\sqrt{\text{Spr}(\hat{\mathcal{O}})}$ from $\langle \hat{\mathcal{O}} \rangle$.

The summary position can be written in a way which makes it look quite like the definition of the center of mass:

$$\langle \hat{\vec{x}} \rangle = \langle \psi | \hat{\vec{x}} | \psi \rangle = \int_{\mathbb{S}} \vec{x} |\psi|^2 d^d \vec{x}. \tag{1.5.90}$$

This precisely would be the center of mass (see Orig. (2.8.28)) if we interpret $dm = m|\psi|^2 d^d\vec{x}$ and thus interpret the wavefunction as having mass density $\rho_m = m|\psi|^2$ where $m$ is its total mass. As well, the correlations of the components of the position are quite closely related to the moment of inertia tensor $I^{ab}$:

$$\text{Cor}(\hat{x}^a, \hat{x}^b) = \int_{\mathbb{S}} (x^a - \langle \hat{x}^a \rangle)(x^b - \langle \hat{x}^b \rangle) |\psi|^2 d^d\vec{x} \implies I^{ab} = m \left( \text{Cor}(\hat{x}^c, \hat{x}_c)\delta^{ab} - \text{Cor}(\hat{x}^a, \hat{x}^b) \right). \tag{1.5.91}$$

We will find that this identification for the mass density is more than just an analogy in the coming sections. The fact that we are not dealing with a rigid body is expressed in $I^{ab}$ having generic time dependence, rather than simply rotating its indices (as in Orig. section 2.8).

Consider two vectors $|u\rangle$ and $|v\rangle$ and assume $|v\rangle \neq 0$. Define a third vector $|w\rangle$ by:

$$|w\rangle = |u\rangle + \alpha |v\rangle \tag{1.5.92}$$

for some constant $\alpha$. Let us choose $\alpha$ so that $|w\rangle$ and $|v\rangle$ are orthogonal:

$$\langle v|w\rangle = \langle v|u\rangle + \alpha \langle v|v\rangle \stackrel{!}{=} 0 \implies \alpha = -\frac{\langle v|u\rangle}{\langle v|v\rangle}. \implies |w\rangle = |u\rangle - \frac{\langle v|u\rangle}{\langle v|v\rangle}|v\rangle. \tag{1.5.93}$$

By definition of the norm, $\langle w|w\rangle \geq 0$ with equality if and only if $|w\rangle = 0$. Therefore:

$$\langle w|w\rangle = \left( \langle u| - \frac{\langle u|v\rangle}{\langle v|v\rangle}\langle v| \right) \left( |u\rangle - \frac{\langle v|u\rangle}{\langle v|v\rangle}|v\rangle \right) = \langle u|u\rangle - \frac{|\langle u|v\rangle|^2}{\langle v|v\rangle} \geq 0 \implies \langle u|u\rangle\langle v|v\rangle \geq |\langle u|v\rangle|^2. \tag{1.5.94}$$

Precise equality can only hold in the event $|w\rangle$ is 0. When $|w\rangle = 0$, $|u\rangle$ and $|v\rangle$ are then colinear and in fact one can see immediately from (1.5.93) that the converse holds as well. To derive this inequality we had to assume $|v\rangle \neq 0$. However, if $|v\rangle = 0$ the inequality holds trivially. Thus, we have the **Cauchy-Schwarz inequality**, that for all vectors $|u\rangle$ and $|v\rangle$:

$$\boxed{\langle u|u\rangle\langle v|v\rangle \geq |\langle u|v\rangle|^2.} \tag{1.5.95}$$

with equality holding if and only if the two vectors are colinear. Now let us consider two operators $\hat{A}$ and $\hat{B}$ and a state $|\psi\rangle$ and define:

$$|u\rangle = (\hat{A} - \langle \hat{A} \rangle \mathbb{1})|\psi\rangle, \qquad |v\rangle = (\hat{B} - \langle \hat{B} \rangle \mathbb{1})|\psi\rangle. \tag{1.5.96}$$

Then:

$$\langle u|u\rangle = \langle \psi | (\hat{A}^\dagger - \langle \hat{A} \rangle^* \mathbb{1})(\hat{A} - \langle \hat{A} \rangle \mathbb{1}) |\psi\rangle = \text{Spr}(\hat{A}) \tag{1.5.97}$$

$$\langle v|v\rangle = \langle \psi | (\hat{B}^\dagger - \langle \hat{B} \rangle^* \mathbb{1})(\hat{B} - \langle \hat{B} \rangle \mathbb{1}) |\psi\rangle = \text{Spr}(\hat{B}) \tag{1.5.98}$$

$$\langle u|v\rangle = \langle \psi | (\hat{A}^\dagger - \langle \hat{A} \rangle^* \mathbb{1})(\hat{B} - \langle \hat{B} \rangle \mathbb{1}) |\psi\rangle \tag{1.5.99}$$

$$\begin{aligned} &= \langle \psi | \left( \frac{(\hat{A}^\dagger - \langle \hat{A} \rangle^* \mathbb{1})(\hat{B} - \langle \hat{B} \rangle \mathbb{1}) - (\hat{B} - \langle \hat{B} \rangle \mathbb{1})(\hat{A}^\dagger - \langle \hat{A} \rangle^* \mathbb{1})}{2} \right. \\ &\qquad \left. + \frac{(\hat{A}^\dagger - \langle \hat{A} \rangle^* \mathbb{1})(\hat{B} - \langle \hat{B} \rangle \mathbb{1}) + (\hat{B} - \langle \hat{B} \rangle \mathbb{1})(\hat{A}^\dagger - \langle \hat{A} \rangle^* \mathbb{1})}{2} \right) |\psi\rangle \end{aligned} \tag{1.5.100}$$

$$= \langle \psi | \frac{1}{2} [\hat{A}^\dagger - \langle \hat{A} \rangle^* \mathbb{1}, \hat{B} - \langle \hat{B} \rangle \mathbb{1}] |\psi\rangle + \text{Cor}(\hat{A}, \hat{B}) \tag{1.5.101}$$

$$= \frac{1}{2}\langle[\hat{A}^\dagger, \hat{B}]\rangle + \text{Cor}(\hat{A}, \hat{B}). \tag{1.5.102}$$

Thus, by the Cauchy-Schwarz inequality:

$$\text{Spr}(\hat{A})\text{Spr}(\hat{B}) \geq \left|\frac{1}{2}\langle[\hat{A}^\dagger, \hat{B}]\rangle + \text{Cor}(\hat{A}, \hat{B})\right|^2. \tag{1.5.103}$$

For the special case of $\hat{A}$ and $\hat{B}$ being self-adjoint, we have:

$$\langle v|u\rangle = \frac{1}{2}\langle[\hat{A}, \hat{B}]\rangle^* + \text{Cor}(\hat{A}, \hat{B})^* = -\frac{1}{2}\langle[\hat{A}, \hat{B}]\rangle + \text{Cor}(\hat{A}, \hat{B}) \tag{1.5.104}$$

$$|\langle u|v\rangle|^2 = \left(\frac{1}{2}\langle[\hat{A}, \hat{B}]\rangle + \text{Cor}(\hat{A}, \hat{B})\right)\left(-\frac{1}{2}\langle[\hat{A}, \hat{B}]\rangle + \text{Cor}(\hat{A}, \hat{B})\right) \tag{1.5.105}$$

$$= \text{Cor}(\hat{A}, \hat{B})^2 - \frac{1}{4}\langle[\hat{A}, \hat{B}]\rangle = \text{Cor}(\hat{A}, \hat{B})^2 + \left(\frac{1}{2i}\langle[\hat{A}, \hat{B}]\rangle\right)^2. \tag{1.5.106}$$

And so, (1.5.103) becomes:

$$\text{Spr}(\hat{A})\text{Spr}(\hat{B}) \geq \text{Cor}(\hat{A}, \hat{B})^2 + \left(\frac{1}{2i}\langle[\hat{A}, \hat{B}]\rangle\right)^2. \tag{1.5.107}$$

Together this gives us the uncertainty theorem:

**Theorem 1.8: Uncertainty Theorem**

For any two operators $\hat{A}$ and $\hat{B}$:

$$\text{Spr}(\hat{A})\text{Spr}(\hat{B}) \geq \left|\frac{1}{2}\langle[\hat{A}^\dagger, \hat{B}]\rangle + \text{Cor}(\hat{A}, \hat{B})\right|^2. \tag{1.5.108}$$

For the special case of two self-adjoint operators:

$$\text{Spr}(\hat{A})\text{Spr}(\hat{B}) \geq \text{Cor}(\hat{A}, \hat{B})^2 + \left(\frac{1}{2i}\langle[\hat{A}, \hat{B}]\rangle\right)^2. \tag{1.5.109}$$

### 1.5.2 Classical Limit

The correlation and commutator of operators $\hat{A}$ and $\hat{B}$ place limits on how small the spreads of $\hat{A}$ and $\hat{B}$ can be made simultaneously. The most basic special case of interest is when the operators are the position and momentum operators in the same direction $\vec{n}$. That is, define $\hat{x} = \vec{n} \cdot \hat{\vec{x}}$ and $\hat{p} = \vec{n} \cdot \hat{\vec{p}}$ where $\vec{n}$ is a constant unit vector. Then:

$$[\hat{x}, \hat{p}] = [\vec{n} \cdot \hat{\vec{x}}, \vec{n} \cdot \hat{\vec{p}}] = \vec{n} \cdot [\hat{\vec{x}}, \hat{\vec{p}}] \cdot \vec{n} = i\hbar\mathbb{1}. \tag{1.5.110}$$

Consequently:

$$\text{Spr}(\hat{x})\text{Spr}(\hat{p}) \geq \text{Cor}(\hat{x}, \hat{p})^2 + \left(\frac{1}{2i}\langle[\hat{x}, \hat{p}]\rangle\right)^2 \geq \left(\frac{1}{2i}\langle i\hbar\mathbb{1}\rangle\right)^2 = \frac{\hbar^2}{4}. \tag{1.5.111}$$

Thus, there is a limit to how small the spreads of position and momentum can be made simultaneously. This is not surprising. As a state's position spread gets smaller, it comes closer to being an eigenstate of the position operator and thus gets closer to being a Dirac delta. As a state's momentum spread gets smaller, it comes closer to being an eigenstate of the momentum operator and thus gets closer to being a complex exponential. A Dirac delta is entirely localized to a point whereas a complex exponential does not decay in magnitude at all with distance. So, there is a limit to how well a function can simultaneously approximate these two very different behaviors. **This is a general fact about the Fourier transform, is purely mathematical, and is not derived from any physics**.

Suppose we have the experimental sensitivity to physics only on length scales of at least $w_x$ and momenta up to a resolution $w_p$. The classical regime is whenever the wavefunction spreads are essentially negligible compared to these resolutions. So, with $\sigma_x = \sqrt{\text{Spr}(\hat{x})}$ and $\sigma_p = \sqrt{\text{Spr}(\hat{p})}$, classicality requires:

$$\sigma_x \ll w_x, \qquad \sigma_p \ll w_p. \tag{1.5.112}$$

$\sigma_x$ and $\sigma_p$ are called the position and momentum widths of the wavefunction. Using (1.5.111):

$$w_p^2 \gg \text{Spr}(\hat{p}) \geq \frac{\hbar^2}{4\text{Spr}(\hat{x})} \implies \text{Spr}(\hat{x}) \gg \frac{\hbar^2}{4w_p^2}, \qquad w_x^2 \gg \text{Spr}(\hat{x}) \geq \frac{\hbar^2}{4\text{Spr}(\hat{p})} \implies \text{Spr}(\hat{p}) \gg \frac{\hbar^2}{4w_x^2} \tag{1.5.113}$$

So, we have that in order for a wavefunction to be in the "classical regime" for given experimental resolutions we must have the position and momentum widths sit in the ranges:

$$\frac{\hbar}{2w_p} \ll \sigma_x \ll w_x, \qquad \frac{\hbar}{2w_x} \ll \sigma_p \ll w_p. \tag{1.5.114}$$

As well, simply multiplying the two pieces of (1.5.112) using the uncertainty theorem gives:

$$w_x w_p \gg \frac{\hbar}{2}. \tag{1.5.115}$$

That is, the classical limit is only achievable in an experiment with position and momentum resolutions satisfying (1.5.115) and for wavefunctions satisfying (1.5.114). If your experimental resolution is too good, so that $w_x w_p$ is too small, then you will intrinsically have to deal with the full wavefunction and cannot expect any sort of simplification to classical mechanics. Because much of the world is well described by classical mechanics, the constant $\hbar$ in (1.5.115) must have a very small numerical value *when expressed in the SI unit system* (as the scales in SI units – the meter, second, and kilogram – that comprise the units of $\hbar$ are all of orders of magnitude in which classical mechanics is typically a very good approximation). This suspicion is confirmed by experiment (see the homework of next section).

The requirement that under classical conditions the Schrödinger equation should reproduce classical mechanics places some restrictions on what $\hat{H}$ can be. The quantum Hamiltonian is exactly $\langle\psi|\hat{H}|\psi\rangle$ and so the exact dynamics of the operator summaries of the position and momenta are given by:

$$\frac{d\vec{x}_{\text{cl}}}{dt} = \{\vec{x}_{\text{cl}}, \mathcal{H}\} = \langle\frac{1}{i\hbar}[\hat{\vec{x}}, \hat{H}]\rangle, \qquad \frac{d\vec{p}_{\text{cl}}}{dt} = \{\vec{p}_{\text{cl}}, \mathcal{H}\} = \langle\frac{1}{i\hbar}[\hat{\vec{p}}, \hat{H}]\rangle. \tag{1.5.116}$$

There are just special cases of Ehrenfest's theorem as proven in problem 2.10. Under approximately classical conditions, these should reproduce Hamilton's equations for the corresponding classical system. Therefore, for approximately classical wavefunctions, we must have:

$$\langle\frac{1}{i\hbar}[\hat{\vec{x}}, \hat{H}]\rangle \approx \frac{\partial\mathcal{H}_{\text{cl}}}{\partial\vec{p}_{\text{cl}}}, \qquad \langle\frac{1}{i\hbar}[\hat{\vec{p}}, \hat{H}]\rangle \approx -\frac{\partial\mathcal{H}_{\text{cl}}}{\partial\vec{x}_{\text{cl}}} \tag{1.5.117}$$

where $\mathcal{H}_{\text{cl}}(t, \vec{x}_{\text{cl}}, \vec{p}_{\text{cl}})$ is the Hamiltonian of the corresponding classical system that our quantum system is supposed to reproduce. Generally, these conditions will **NOT** be enough to determine the Hamiltonian operator $\hat{H}$. However, requiring that the quantum system should have some symmetries of interest and reproduce the classical equations of motion in the classical limit is sometimes (but still not always) enough to determine what the Hamiltonian operator $\hat{H}$ should be for the system. A quantum system is not defined by its classical limit! The less fundamental theory (classical mechanics) is defined by the (limit of the) more fundamental theory (quantum mechanics). A quantum system is just *constrained* by its classical limit.

We can see more precisely what classical physics tells us about the Hamiltonian operator by using the results of Orig. problem 4.48 for wave packet solutions to a generic quadratic field theory. The Schrödinger Hamiltonian can be written explicitly in the position basis:

$$\mathcal{H} = \langle\psi|\hat{H}|\psi\rangle = \int_{\mathbb{S}}\int_{\mathbb{S}} \psi^\dagger(t, \vec{x})\langle\vec{x}|\hat{H}(t)|\vec{x}'\rangle\psi(t, \vec{x}')d^3\vec{x}'d^3\vec{x}. \tag{1.5.118}$$

Using the Taylor expansion of $\psi(t, \vec{x})$:

$$\mathcal{H} = \int_{\mathbb{S}}\int_{\mathbb{S}} \psi^\dagger(t, \vec{x})\langle\vec{x}|\hat{H}(t)|\vec{x}'\rangle \sum_{n=0}^{\infty} \frac{1}{n!}(x'^{a_1} - x^{a_1})...(x'^{a_n} - x^{a_n})\partial^n_{a_1...a_n}\psi(t, \vec{x})d^3\vec{x}'d^3\vec{x}. \tag{1.5.119}$$

Then, we can define the differential operator $H(t, \vec{x}, -i\hbar\nabla)$ by:

$$H(t, \vec{x}, -i\hbar\nabla) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\int_{\mathbb{S}}\langle\hat{\vec{x}}|\hat{H}(t)|\vec{x}'\rangle(x'^{a_1} - x^{a_1})...(x'^{a_n} - x^{a_n})d^3\vec{x}'\right)\partial^n_{a_1...a_n}. \tag{1.5.120}$$

With this, the Schrödinger Hamiltonian becomes:

$$\mathcal{H} = \int_{\mathbb{S}} \psi^\dagger(t, \vec{x})H(t, \vec{x}, -i\hbar\nabla)\psi(t, \vec{x})d^3\vec{x} \tag{1.5.121}$$

precisely in the form found in Orig. problem 4.48.

If our wavefunction is approximately localized in momentum space then $|\widetilde{\psi}(t,\vec{p})|^2$ (where $\widetilde{\psi}(t,\vec{p}) = \langle\vec{p}|\psi(t)\rangle$) will be small outside of a small neighborhood surrounding the classical momentum $\vec{p}_{\text{cl}}(t)$. The momentum space wavefunction $\widetilde{\psi}$ is just the Fourier transform of $\psi(t,\vec{x})$ up to multiplying by $\frac{1}{\hbar^{\frac{3}{2}}}$ and replacing the wavenumber $\vec{k}$ with $\frac{\vec{p}}{\hbar}$. Therefore, if the wavefunction is approximately localized in momentum space then it is also approximately localized in Fourier wavenumber space around a wavenumber $\vec{K}(t)$ which is related to $\vec{p}_{\text{cl}}(t)$ by:

$$\vec{p}_{\text{cl}} = \hbar\vec{K}. \tag{1.5.122}$$

This is known as the **de Broglie relation**. If the wavefunction is approximately localized in momentum space then the $H(t,\vec{x},-i\hbar\nabla)$ operator immediately acts on $\psi$ approximately by:

$$H(t,\vec{x},-i\hbar\nabla)\psi(t,\vec{x}) \approx H(t,\vec{x},\vec{p}_{\text{cl}})\psi(t,\vec{x}) \tag{1.5.123}$$

If the wavefunction is also approximately localized in position space, $|\psi(t,\vec{x})|^2$ will be small outside of a small neighborhood surrounding the classical position $\vec{x}_{\text{cl}}(t)$ allowing us to complete the position integral in (1.5.121) to find:

$$\mathcal{H} \approx H(t,\vec{x}_{\text{cl}},\vec{p}_{\text{cl}}). \tag{1.5.124}$$

In order for this to produce classical physics then, $H$ must approximately be the classical Hamiltonian $\mathcal{H}_{\text{cl}}(t,\vec{x}_{\text{cl}},\vec{p}_{\text{cl}})$. This does not fully fix $H$ (one can't simply just say $H = \mathcal{H}_{\text{cl}}$ as functions) because no wavefunction is exactly localized in both position and momentum space and so there is no exactly classical wavefunction. Instead, we can only require that $H = \mathcal{H}_{\text{cl}}+$ corrections which are small when the wavefunction is approximately classical. We can be more formal and exact about how what extra terms can be in $H$ that are missing from $\mathcal{H}_{\text{cl}}$, but we will save that for chapter 2.

Under the classical limit, we can also use (1.5.123) in the Schrödinger equation to see:

$$i\hbar\dot{\psi}(t,\vec{x}) \approx H(t,\vec{x},\vec{p}_{\text{cl}})\psi(t,\vec{x}). \tag{1.5.125}$$

If $\vec{x}$ is near $\vec{x}_{\text{cl}}$ then we can approximately replace $H(t,\vec{x},\vec{p}_{\text{cl}})$ with $H(t,\vec{x}_{\text{cl}},\vec{p}_{\text{cl}})$. If $\vec{x}$ is far from $\vec{x}_{\text{cl}}$ and $\psi$ is approximately classical then $\psi(t,\vec{x}) \approx 0$. In either case:

$$H(t,\vec{x},\vec{p}_{\text{cl}})\psi(t,\vec{x}) \approx H(t,\vec{x}_{\text{cl}},\vec{p}_{\text{cl}})\psi(t,\vec{x}) \tag{1.5.126}$$

and so we can see that an approximately classical wavefunction has approximately well defined angular frequency $\Omega$ given by:

$$H(t,\vec{x}_{\text{cl}},\vec{p}_{\text{cl}}) = \hbar\Omega. \tag{1.5.127}$$

This last condition is known as the **Planck relation**. The approximate equations of motion derived for $\vec{x}_{\text{cl}}$ and $\vec{K}$ in Orig. problem 4.48 were:

$$\frac{d\vec{x}_{\text{cl}}}{dt} \approx \frac{\partial\Omega}{\partial\vec{K}}, \qquad \frac{d\vec{K}}{dt} \approx -\frac{\partial\Omega}{\partial\vec{x}_{\text{cl}}}. \tag{1.5.128}$$

Using the Planck and de Broglie relations we have:

$$\frac{d\vec{x}_{\text{cl}}}{dt} \approx \frac{\partial H}{\partial\vec{p}_{\text{cl}}}, \qquad \frac{d\vec{p}_{\text{cl}}}{dt} \approx -\frac{\partial H}{\partial\vec{x}_{\text{cl}}} \tag{1.5.129}$$

which are just Hamilton's equations (when we use that $H \approx \mathcal{H}_{\text{cl}}$). Therefore, **classical physics emerges as the wave packet approximation of quantum mechanics.**

### 1.5.3 Meaning of $\hbar$

In Orig. (5.2.61) and (5.2.63) we found that the whole of analytical mechanics is agnostic to the choice of the units of the action and that rescaling our choice of units of action is equivalent to rescaling our choice of units of mass. Therefore, when we chose the kilogram as a reference unit of mass it implicitly (when combined with our choices of units for length and time) determined a numerical value of $\hbar$ in that unit system. This is just like how when one chooses the meter as a reference unit of length and the second as a reference unit of time it implicitly determines a numerical value of $c$. However, we know that $c$ can be eliminated from all equations and set to a numerical value of 1 by choosing our unit of length to be the distance traveled by light in our unit of time. This choice is uniquely special because $c$ is the only invariant speed (due to Lorentz symmetry). A very similar choice can be made for $\hbar$. Due to the Planck and de Broglie relations, the momentum 4-vector $(\text{E},\vec{p})$ of a normalized wave with well defined angular frequency and wave number is

related to the wavenumber 4-vector $(\omega, \vec{k})$ by:

$$p^\mu = \hbar k^\mu. \tag{1.5.130}$$

But, the units of energy and momentum depend on our choice of the units of action and so can be rescaled one time only by an arbitrary factor. So, **if we define our units of the action so that the energy of a normalized Schrödinger wave of well-defined angular frequency is equal to that angular frequency then $\hbar = 1$ identically**. Thus, just like $c$, **$\hbar$ is not a fundamental constant.**

In units in which $c \neq 1$, one needs to perform an experiment in order to measure $c$. This can be accomplished through any experiment which can measure how far light travels in a measurable time. Alternatively, in units in which $c = 1$, one needs to perform the same experiment in order to calibrate what numerical value to assign to the distance traveled by the light. That is, if one lets light travel for 1 nanosecond and measures how far it goes, marking the start and end points of its travel, the distance between those two points defines the distance of a light-nanosecond. Light identically travels one light-nanosecond per nanosecond and so has speed 1. Similarly, in units in which $\hbar \neq 1$, one needs to perform an experiment in order to measure $\hbar$ (see the homework next section). Alternatively, in units in which $\hbar = 1$, one needs to perform the same experiment in order to calibrate what numerical value to assign to the mass used in that experiment. No more or less experimental information is necessary no matter what units you use.

There is further analogy between $c$ and $\hbar$. How velocities compare to $c$ controls the nonrelativistic limit. So, formally taking $c \to \infty$ (as it is the equivalent of $v/c \to 0$) turns relativistic equations into their nonrelativistic counterparts. Similarly, the classicality conditions (1.5.114) and (1.5.115) show that the extent to which it is impossible to have a wavefunction which is simultaneously completely localized in position space and in momentum space is controlled by $\hbar$. Therefore, $\hbar \to 0^+$ formally gives the classical limit. The reason I am saying these are *formal* limits is that $c$ and $\hbar$ in reality always have fixed numerical values (and are 1 in my favorite units). In every experiment they are always fixed; there is no experiment in which $c \to \infty$ or $\hbar = 0$. Properly, the nonrelativistic or classical limit is approximately achieved whenever the genuinely dynamical relevant experimental parameters (velocities or wavepacket spreads) are small.

For the rest of this and the next chapter we will keep around factors of $\hbar$ and not choose its value to be 1. However, beginning in chapter 3 we will set it to 1 just as we have with $c$ and $\epsilon_0$ and we will keep it that way unless stated otherwise. We found in (1.2.22) that $\hbar$ has units of the action and in Orig. (5.2.62) that the action has units $[S] = \frac{ML^2}{T}$. In modern SI units the kilogram is defined so that $\hbar$ takes the exact value [15]:

$$\hbar = \frac{6.62607015}{2\pi} \times 10^{-34} \frac{\text{kg m}^2}{\text{s}}. \tag{1.5.131}$$

In order to establish this value experimentally (meaning to take an object and measure its mass in kilograms as defined by $\hbar$ having this value) requires an experiment that can compare macroscopically measurable quantities to $\hbar$. We describe a few ways to do this in the homework throughout this chapter in problems 1.49, 1.50, 1.51. Once we have defined our unit of mass so that $\hbar = 1$, we will have an induced unit of mass $M = \frac{T}{L^2}$ so that the action is unitless. The unit system in which $\epsilon_0 = 1$ and $\hbar = 1$ we call **Schrödinger Natural Units (SNU)**. This leaves time and length as the only independent units. We have not set $c = 1$ because for the bulk of this chapter we consider nonrelativistic theories, so $c$ does not appear anywhere. If we use meters and seconds as our units of length and time, then in SNU the unit of mass is $\frac{\text{s}}{\text{m}^2}$.

Quantum mechanics also makes it possible to define the second at a level of precision that was previously impossible. In SI units, the second is defined exactly as the duration of 9192631770 periods of the electromagnetic radiation emitted by the transition between the hyperfine levels of the unperturbed ground state of the $^{133}$Cs atom. To unpack this phrase, when modeled as a quantum system, the Cesium-133 atom has a lowest energy state (the ground state) and due to the interactions of the electromagnetic fields generated by the atomic nucleus with the electrons in the atom, there is a characteristic resonant angular frequency $\omega$ of electromagnetic waves emitted in the transition of this atom between its lowest energy configurations (the ground state and its nearest neighbor eigenstate of the interacting Hamiltonian operator). That angular frequency of electromagnetic radiation then defines a period $\tau = \frac{2\pi}{\omega}$ and the second is defined to be exactly $9192631770\tau$. This gives the second a universally reproducible definition and it for all but the most precise of purposes coincides with the more colloquial understanding of a second as $1/86400^{\text{th}}$ of a solar day on Earth.

Looking ahead to when we do incorporate relativity and so use $c = 1$ units, $\hbar = c = \epsilon_0 = k_B = 1$ units leave only one single remaining independent unit. A unit of time can be converted to length using $c$, or to a mass using $\hbar$ and $c$, or to a charge using $\hbar$, $c$, and $\epsilon_0$, or to a temperature (in separate temperature units rather than energy units, gross) using $\hbar$, $c$, and $k_B$. This is our final unit system, and it is called either

**Natural Units (NU)** or **High Energy Units (HEU)**. It is the standard system used in particle physics. For the last remaining unit, if you want to correspond as closely to SI as possible, you should take the free unit to be time in seconds (with the second defined as above in terms of Cesium). Alternatively, to correspond to particle physics practice as closely as possible, you should take the free unit to be energy in electronvolts eV. The electronvolt is defined exactly as the energy gained by an electron when it is accelerated through an electrostatic potential difference of 1 Volt (1 Volt $= 1\frac{\text{kg m}^2}{\text{s}^2\text{C}}$ in SI units where C is the SI unit of charge the Coulomb defined so that the electron charge is precisely $1.602176634 \times 10^{-19}$C). Using the SI definition of the second gives that in HEU the electronvolt and the second are exactly related by:

$$1\text{eV} = 2\pi\frac{1.602176634}{6.62607015}10^{15}\frac{1}{\text{s}} \approx 1.51926744787862620125 \times 10^{15}\frac{1}{\text{s}}. \tag{1.5.132}$$

So, for HEU we take either the eV or the second as our one defining unit and use the number presented here as a conversion factor between those two choices. Alternatively, using the Cesium definition, the electronvolt is defined so that the energy difference due to hyperfine splitting of the unperturbed ground state of Cesium-133 is exactly:

$$\Delta\text{E} = \frac{6.62607015 \times 0.919263177}{1.602176634} \times 10^{-5}\text{ eV} \approx 38.0176702609049942617\ \mu\text{eV} \tag{1.5.133}$$

in terms of microelectronvolts ($10^{-6}$eV). The electronvolt is an extremely convenient unit for atomic, nuclear, and particle physics. Unfortunately, it is a bit gross for every day quantities. Using $\hbar = c = \epsilon_0 = k_B = 1$ (and for the case of the Coulomb, the experimental value of the fine-structure constant $\alpha$, see problem 1.52), all of the more familiar SI units of the meter, second, kilogram, Coulomb, and Kelvin can be expressed in terms of electronvolts. In particular:

| SI Unit | eV Equivalent |
|---|---|
| 1s | $\approx 1.5192674479 \times 10^{15}\ \text{eV}^{-1}$ |
| 1m | $\approx 5.0677307162 \times 10^{6}\ \text{eV}^{-1}$ |
| 1kg | $\approx 5.6095886038 \times 10^{35}$ eV |
| 1C | $\approx 1.8900670154 \times 10^{18}$ |
| 1K | $\approx 8.6173332621 \times 10^{-5}$ eV. |

(1.5.134)

The huge range of exponents and unobvious way to verbally pronounce $\text{eV}^{-1}$ make it clear why we don't use high energy units in daily social life. The range of human reasonable temperatures isn't so bad in electronvolts, with the range $-50°$C to $50°$C corresponding to 19.2meV to 27.8meV and a comfortable room temperature of $20°$C corresponding to 25.2meV.

### 1.5.4 Rigged Hilbert Space

The position and momentum operators are surprisingly mathematically complicated objects even though in the position basis they simply multiply by a real number or take a derivative. We found that their eigenstates satisfy:

$$\langle\vec{x}|\vec{x}'\rangle = \delta(\vec{x}-\vec{x}'), \qquad \langle\vec{p}|\vec{p}'\rangle = \delta(\vec{p}-\vec{p}'). \tag{1.5.135}$$

Automatically it follows that if we set $|\phi\rangle = |\vec{x}\rangle$ for a particular value of $\vec{x}$, then $|\phi\rangle$ does not have finite norm:

$$\langle\phi|\phi\rangle = \delta(\vec{x}-\vec{x}) = \delta(0). \tag{1.5.136}$$

$\delta(0)$ is not a number, because the Dirac delta is not a proper function. It effectively acts like $+\infty$. To understand it in any meaningful sense you can consider a sequence of functions $|\phi_n\rangle$ which limit to $|\vec{x}\rangle$ and see that the norm of $|\phi_n\rangle$ limits to $\infty$ as $n \to \infty$. For example, if you take $|\phi_n\rangle$ to be a Gaussian in position space centered at $\vec{x}$ and with width proportional to $\frac{1}{n}$ then you can find explicitly what the normalization of that Gaussian has to be in order for $|\phi_n\rangle$ to behave (when inner producted with other functions) as $|\vec{x}\rangle$ in the $n \to \infty$ limit. You can then see explicitly that that normalization scales like $\sqrt{n}$, and so blows up for large $n$. The Gaussian is a particular choice where the calculation can be done completely explicitly but the result is independent of how $|\vec{x}\rangle$ is approached as a limit. Because $|\vec{x}\rangle$ has infinite norm, it isn't mathematically a proper element in the Hilbert space $\mathbb{H}$. The identical argument holds for $|\vec{p}\rangle$.

Neither $|\vec{x}\rangle$ for any $\vec{x}$ nor $|\vec{p}\rangle$ for any $\vec{p}$ are in $\mathbb{H}$. This isn't actually a crazy thought. The Hilbert space is supposed to be the space of physically reasonable states of our system (finite momentum, finite energy, finite generators in general, and so finite norm) and both $|\vec{x}\rangle$ and $|\vec{p}\rangle$ are quite pathological states if taken too seriously. $|\vec{x}\rangle$ is entirely localized in position space, and so effectively represents a point particle, which we

know from the Orig. is fundamentally an unphysical idea. $|\vec{p}\rangle$ is not localized at all in position space as it is a plane wave with infinite spatial extent. Even though it carries a finite energy density, it carries that energy density uniformly throughout all space and so carries infinite total energy. The fact that neither $|\vec{x}\rangle$ nor $|\vec{p}\rangle$ is in our space of exactly allowed physical states is thus a good thing; they are quite unphysical!

However, mathematically $|\vec{x}\rangle$ and $|\vec{p}\rangle$ are useful tools for doing calculations. It is important to in some way formalize what they are and to what space they belong. After all, it doesn't make sense in a purely rigorous sense to say that the $|\vec{x}\rangle$ form a "basis" of the Hilbert space if they aren't even in the Hilbert space to begin with. The appropriate mathematical fix is called a Rigged Hilbert Space, which is a set of three vectors spaces, the first and smallest we call $\Phi$, the second is our usual Hilbert space $\mathbb{H}$, and the third and largest we call $\widetilde{\Phi}$. $\Phi$ is called the space of test functions. It is the subset of $\mathbb{H}$ that is the space of smooth (infinitely differentiable) functions of compact support (only nonzero in a closed finite region). The Hilbert space $\mathbb{H}$ of functions of finite norm is a subset of $\widetilde{\Phi}$. $\widetilde{\Phi}$ is the space of distributions. A distribution (like the Dirac delta) is an object that is not generally a well defined mathematical function but has well defined behavior when integrated against a test function. In other words, the inner product of a test function with a test function, Hilbert space element, or distribution is well defined and finite; the inner produce of a Hilbert space element with a test function or Hilbert space element is well defined and finite but not generally so with a distribution; and, the inner produce of a distribution is generally only well defined and finite when with a test function. Both $|\vec{x}\rangle$ and $|\vec{p}\rangle$ are distributions in $\widetilde{\Phi}$, but not vectors in $\mathbb{H}$.

The spectral theorem in infinite dimensions guarantees that a normal operator has an orthogonal spanning set of generalized eigenvectors. Generalized eigenvector here means a "vector" which is only guaranteed to be an element of $\widetilde{\Phi}$, not necessarily of $\mathbb{H}$. These generalized eigenvectors are orthogonal in the sense that different ones have vanishing inner products with each other. They cannot generally be normalized to a finite value but a normalization condition can still be imposed (such as fixing $\langle\vec{p}|\vec{p}'\rangle = \delta(\vec{p}-\vec{p}')$ with no prefactor on the Dirac deltas). They are a spanning set in the sense that each element of $\mathbb{H}$ can be written as a (potentially infinite) linear combination of them. So, we informally say that the operator has an orthonormal basis of eigenvectors, in analogy with the finite dimensional linear algebra. Physicists tend to be pretty relaxed with this language as opposed to mathematicians that study functional analysis who are very careful. The takeaway from this subsection is that we will maintain only about the typical physics level of rigor in our discussions of Hilbert spaces but it should be known that an entirely rigorous treatment can be done and that while our proofs given here could be improved in term of rigor, our essential results are durable when treated with higher level analysis (and phrased in appropriately rigorized language).

### 1.5.5 Phenomenological Born Rule

An individual quantum particle, such as an electron, is extremely small and not directly detectable by human senses. In order to make meaningful measurements of quantum systems, we have to couple that microscopic system to a macroscopic measuring apparatus (we call it a detector). Beyond our control, that microscopic system will also inevitably be coupled (however weakly) to an enormous environment of variables (think of all the degrees of freedom in the air all around you) which we for the purposes of the description include as part of the detector (in actual practice, the environmental degrees of freedom are infinite, due to fields). The distinctions microscopic and macroscopic used here are precisely those from statistical mechanics. A macroscopic apparatus necessarily consists of an extremely large number of microscopic degrees of freedom, as physical objects are comprised of many many microscopic constituents. The part of the detector we can read is called the meter (some particular macroscopic variable of the detector) and all of the many microscopic parts of the detector we cannot read are called the environment. We might wish that it would be possible to build a detector so that the reading of the meter was completely determined by the system we wish to measure, however the fact that the environment is necessarily large makes this overly optimistic. In order to make the meter sensitive to the microscopic system we want to measure, we have to make it able to be perturbed macroscopically by small microscopic fluctuations. This sensitivity inevitably also makes it at least partly sensitive to the environment, which has unaccounted for microscopic dynamics. Because we have to treat the environment as having many averaged out degrees of freedom, **a proper treatment of the measurement of quantum systems requires using statistical mechanics**.

Such a detailed proper treatment will have to wait for chapter 2, once we have built up more technical machinery. The main results of that analysis are two approximate effective principles known as the Born rule and the collapse rule. While we won't go deep on them yet, for making connection to experimental results it is useful to at least describe them now. An **ideal measurement** is the optimal measurement scheme

possible consistent with quantum mechanics and statistical mechanics, in the sense of introducing the least error. Suppose we have a microscopic quantum system with state $|\psi\rangle$ for which we are going to measure a single property by an ideal detector. In an ideal measurement, it is possible to divide the Hilbert space $\mathbb{H}$ of our microscopic system into orthogonal subspaces $\mathbb{H}_1$, $\mathbb{H}_2$, ... which span the full Hilbert space and so that if $|\psi\rangle \in \mathbb{H}_n$ then the meter is guaranteed to read off a particular value $M_n$ which is distinct for each different value of $n$. We define the projector operator $\hat{\pi}_n$ to project vectors in $\mathbb{H}$ into the subspace $\mathbb{H}_n$, i.e. is $\mathbb{1}$ on $\mathbb{H}_n$ and 0 for vectors orthogonal to $\mathbb{H}_n$. Such a projector is automatically self-adjoint:

$$\hat{\pi}_n = \hat{\pi}_n^\dagger. \tag{1.5.137}$$

The requirement that the subspaces are orthogonal is given by:

$$\hat{\pi}_m \hat{\pi}_n = \delta_{mn} \hat{\pi}_n. \tag{1.5.138}$$

The requirement that the subspaces span the full Hilbert space is given by:

$$\sum_n \hat{\pi}_n = \mathbb{1}. \tag{1.5.139}$$

One may wish that for each $|\psi\rangle \in \mathbb{H}$ there would be a guaranteed corresponding meter reading, however due to the statistical mechanics of the system-meter-environment interaction, this is too much to ask. The large number of interacting degrees of freedom makes the full system-meter-environment system chaotic, and consequently a probabilistic description of the effective results becomes necessary due to our imperfect knowledge of the exact microscopic initial state of the meter-environment system (the detector) and our inability to exactly calculate the dynamics of the interaction system. These subspaces $\mathbb{H}_n$ are stable, in the sense that under the full system interactions $|\psi\rangle$ will stay in one if it is already in one. But, if $|\psi\rangle$ is composed of a linear combination of subspaces, it will dynamically and chaotically (and so effectively randomly from a macroscopic perspective) decay into one of the subspaces and thus select a meter value. Which subspace is selected depends sensitively on the initial conditions of the full system, but by calibrating the detector to be unbiased on average, one can derive the probability of selection for each subspace. Doing so gives **the Born rule**: the probability that the meter $M$ will read the value $M_n$ is given by:

$$\text{Prob}(M = M_n) = \langle\psi|\hat{\pi}_n|\psi\rangle. \tag{1.5.140}$$

This use of projectors to generate probabilities is called a **projection valued measure (PVM)**.

After the interaction has occurred, the wavefunction will have been driven into the corresponding subspace $\mathbb{H}_n$ so that we also have **the collapse rule**: if the meter takes the value $M_n$, the system wavefunction $|\psi_n\rangle$ after the interaction is given by:

$$|\psi_n\rangle = \frac{\hat{\pi}_n|\psi\rangle}{\|\hat{\pi}_n|\psi\rangle\|}. \tag{1.5.141}$$

Sometimes it is useful to express the $\hat{\pi}_n$ explicitly in terms of an eigenbasis. Let $|n, a\rangle$ (with $a$ a label) be a set of orthonormal vectors which span $\mathbb{H}_n$. The range of the label $a$ is the dimension of $\mathbb{H}_n$. Then:

$$\hat{\pi}_m|n, a\rangle = \delta_{mn}|n, a\rangle, \qquad \hat{\pi}_n = \sum_a |n, a\rangle\langle a, n|. \tag{1.5.142}$$

So, the resulting probability for $M = M_n$ in terms of this basis is:

$$\text{Prob}(M = M_n) = \langle\psi|\hat{\pi}_n|\psi\rangle = \sum_a |\langle n, a|\psi\rangle|^2. \tag{1.5.143}$$

In the special case that for a particular $n = n_*$ the corresponding subspace $\mathbb{H}_*$ is 1 dimensional, $a$ only takes the value 1 and we can let $|\phi\rangle = |n_*, 1\rangle$. Under that circumstance, these formulas simplify to:

$$\hat{\pi}_* = |\phi\rangle\langle\phi|, \qquad \text{Prob}(M = M_*) = |\langle\phi|\psi\rangle|^2. \tag{1.5.144}$$

Because of this expression, most sources will call the quantity $|\langle\phi|\psi\rangle|^2$ the probability that when $|\psi\rangle$ is measured, it will be found to become $|\phi\rangle$. Importantly, this is only valid for the special case of a 1 dimensional subspace and under the circumstances that a real measurement may be approximated by an ideal one. Further, it is only valid in the physical context of an actual measurement in which $|\phi\rangle\langle\phi|$ is one of the stable states produced by that particular measurement apparatus because measurement is a dynamical interactive process. $|\langle\psi|\psi\rangle|^2$ has no meaning as a probability outside of a specific measurement context. Conceptualizing it as a probability is not a fundamental or exact principle, which is why I call it an intensity or projective intensity and avoid overuse/misuse of the term probability.

If the meter is going to tell us anything about our system of interest, we want the subspaces $\mathbb{H}_n$ on

which it takes predictable values to be physically meaningful subspaces for that system. Consequently, in an ideal measurement we further require that these subspaces correspond to the eigenspaces of some operator of interest $\hat{G}$ in the quantum system. That way, there is a correspondence between the eigenvalues $G_n$ of $\hat{G}$ and the meter readings $M_n$ so that if $|\psi\rangle$ is an eigenstate of $\hat{G}$ with eigenvalue $G_n$, the meter predictably reads the value $M_n$. This structure lets us (in an ideal measurement) correlate as much as possible the reading of the meter with the value of some operator summary of interest in the system we are measuring. Using the properties of the $\hat{\pi}_n$, this condition gives a simple formula for any function $f$ of $\hat{G}$:

$$f(\hat{G})|\psi\rangle = \sum_n \|\hat{\pi}_n|\psi\rangle\| f(\hat{G})|\psi_n\rangle = \sum_n f(G_n)\hat{\pi}_n|\psi\rangle, \quad \forall\ |\psi\rangle \implies f(\hat{G}) = \sum_n f(G_n)\hat{\pi}_n. \tag{1.5.145}$$

To understand why in an ideal measurement a correspondence between the stable subspaces $\mathbb{H}_n$ and the eigenspaces of an operator of interest $\hat{G}$ is natural, let us consider an example. Suppose we have a detector screen with pixels sensitive to incoming charges. The width of a pixel is $w$. If the incoming wave-packet is entirely localized to a single pixel, then an ideal screen would only have that corresponding pixel light up. We can define the subspace $\mathbb{H}_n$ as the subspace of the Hilbert space which has wavefunctions that are only nonzero on the $n^{\text{th}}$ pixel. Then, we can define the coarse-grained position of that pixel by the position value $x_n$ at that pixel's center and the coarse-grained position operator by:

$$\hat{x}_w = \sum_n x_n \hat{\pi}_n. \tag{1.5.146}$$

Each eigenvalue $x_n$ of $\hat{x}_w$ corresponds to the meter value, in this case $n$, of which pixel lights up when the charge is incident.

Many sources will say that an ideal measurement acts as an uncertain measurement of $\hat{G}$, which is very misleading (not maliciously, just a poor choice of words). Most pedantically, the meter is what is observed, not anything related to $\hat{G}$. Of course, we do want to interpret that meter result in terms of something about the system, so that element of the leap is forgivable. However, $\hat{G}$ isn't a variable, it is just an operator; the state of the system is $|\psi\rangle$ and any measurement reveals information about $|\psi\rangle$, not $\hat{G}$, which doesn't even depend on the system's state. Saying that a measurement measures an operator and gives some probability distribution makes $\hat{G}$ sound like a random variable, which it isn't. The state is $|\psi\rangle$. The operators are just (possibly infinite dimensional) matrices of state-independent numbers. The operator $\hat{\vec{p}}$ is no more or less of a random variable than the matrix $\begin{pmatrix} 2 & 3 \\ -1 & 0 \end{pmatrix}$ (which is to say, not at all).

Instead, it is much more accurate to say that the repeated ideal measurement described above acts as an uncertain measurement of $\langle\hat{G}\rangle$ (which actually does depend on the state $|\psi\rangle$). Because an ideal measurement of a microscopic quantum system by a macroscopic detector gives probabilistic results, in order to learn specific facts about the state $|\psi\rangle$ one needs to prepare the system many times and perform many measurements. Suppose we prepare and measure the same state $|\psi\rangle$ a total of $N$ times and that in observation $k$ we observe the meter reading $M_{\sigma(k)}$ where $\sigma(k)$ is the function that maps which observation is considered to the specific relevant value of $n$. Let $K_n$ be the number of times we observe the value $M_n$, or in other words the number of values of $k$ for which $\sigma(k) = n$. Then, because we have made $N$ observations:

$$\sum_n K_n = N. \tag{1.5.147}$$

We can define the empirical frequency $q_n$ of outcome $M_n$ as:

$$q_n = \frac{K_n}{N}. \tag{1.5.148}$$

In the large $N$ limit, $q_n \to \text{Prob}(M = M_n)$ due to the definition of probability. In an ideal measurement, if we read a value $M_n$ on the meter, we infer that $|\psi\rangle$ has collapsed to $|\psi_n\rangle \in \mathbb{H}_n$ which is then an eigenstate of $\hat{G}$ with eigenvalue $G_n$. Correspondingly, in that state:

$$\langle\psi_n|\hat{G}|\psi_n\rangle = G_n. \tag{1.5.149}$$

Thus, we identify that our uncertain measurement of $\langle\hat{G}\rangle$ has given the value $G_n$ a fraction $q_n$ of the time. Consequently, our sample mean $\bar{G}$ of the empirical values is:

$$\bar{G} = \sum_n q_n \langle\psi_n|\hat{G}|\psi_n\rangle = \sum_n q_n G_n. \tag{1.5.150}$$

In the large $N$ limit then we have:

$$\lim_{N\to\infty}\bar{G}=\sum_n G_n\text{Prob}(M=M_n)=\sum_n G_n\langle\psi|\hat{\pi}_n|\psi\rangle=\langle\psi|\sum_n G_n\hat{\pi}_n|\psi\rangle=\langle\psi|\hat{G}|\psi\rangle=\langle\hat{G}\rangle. \tag{1.5.151}$$

That is, $\bar{G}$ (which is fully determined by the meter readings) in the limit of large data converges to the true value $\langle\hat{G}\rangle$, so that repeated ideal measurements act as a measurement of $\langle\hat{G}\rangle$. This correspondence between the statistical quantity $\bar{G}$ and the operator summary $\langle\hat{G}\rangle$ in the large data limit is why most sources call $\langle\hat{G}\rangle$ an expectation value or operator expectation. However, I view this as a derived correspondence emergent in the ideal measurement model, not a fundamental physical fact, and so avoid this language. Importantly, in this understand $\langle\hat{G}\rangle$ is a real physical property of the system taking the same value in every single experimental instance before the measurement happens. It is in the dynamics of the measurement process that each value $G_{\sigma(k)}$ is realized and from the distribution of these resultant values that the true value $\langle\hat{G}\rangle$ is reconstructed.

In life we always have a finite amount of data so let us determine the behavior of $\bar{G}$ at finite $N$. Let $p_n=\text{Prob}(M=M_n)$. We can alternatively write $\bar{G}$ as:

$$\bar{G}=\frac{1}{N}\sum_{k=1}^N G_{\sigma(k)}. \tag{1.5.152}$$

Because $\sigma(k)$ takes the value $n$ with probability $p_n$, $G_{\sigma(k)}$ takes the value $G_n$ with probability $p_n$. As well, the $\sigma(k)$ and hence $G_{\sigma(k)}$ are identically independently distributed for different $k$. Thus:

$$\mathbb{E}\bar{G}=\frac{1}{N}\sum_{k=1}^N\mathbb{E}[G_{\sigma(k)}]=\frac{1}{N}\sum_{k=1}^N\sum_n p_nG_n=\frac{1}{N}N\langle\hat{G}\rangle=\langle\hat{G}\rangle. \tag{1.5.153}$$

As well:

$$\text{Var}[\bar{G}]=\mathbb{E}[(\bar{G})^2]-(\mathbb{E}\bar{G})^2=\frac{1}{N^2}\sum_{k=1}^N\sum_{k'=1}^N\mathbb{E}[G_{\sigma(k)}G_{\sigma(k')}]-\langle\hat{G}\rangle^2 \tag{1.5.154}$$

$$=\frac{1}{N^2}\sum_{k=1}^N\mathbb{E}[G^2_{\sigma(k)}]+\frac{1}{N^2}\sum_{k=1}^N\sum_{\substack{k'=1\\k'\neq k}}^N\mathbb{E}[G_{\sigma(k)}G_{\sigma(k')}]-\langle\hat{G}\rangle^2 \tag{1.5.155}$$

$$=\frac{1}{N^2}N\sum_n p_nG_n^2+\frac{1}{N^2}\sum_{k=1}^N\sum_{\substack{k'=1\\k'\neq k}}^N\langle\hat{G}\rangle^2-\langle\hat{G}\rangle^2=\frac{\langle\hat{G}^2\rangle-\langle\hat{G}\rangle^2}{N}=\frac{\text{Spr}(\hat{G})}{N}. \tag{1.5.156}$$

Therefore, we can view $\bar{G}$ as giving a statistical measurement of $\langle\hat{G}\rangle$ with a quantifiable variance that decays with the number of trials. We can also identify the operator spread $\text{Spr}(\hat{G})$ as the variance of a single ideal measurement of $\langle\hat{G}\rangle$. Recalling the idea of quantum tomography, we now see that by preparing a system many times with the same state $|\psi\rangle$ and measuring many different generators many different times each, we can develop increasingly precise knowledge of the value of each generator, and approximately reconstruct $|\psi\rangle$ from these values. Thus again, the state $|\psi\rangle$ as a ray in Hilbert space is entirely observable in the limit of large data. While for individual measurements we have statistical results due to the Born rule, for large numbers of measurements we have predictable statistical behavior. Consequently, well-prepared beams of many particles produce reliable intensities.

The similarity between the Born rule and the adiabatic invariants of the Schrödinger Lagrangian is not accidental. In an ideal measurement, the interaction between the system-meter-environment will cause a distinct response of the meter on each subspace $\mathbb{H}_n$. Consequently, the term added to the Hamiltonian to account for the interaction between the system and the detector will be of the form:

$$\mathcal{H}_{\text{int}}=\sum_n\langle\psi|\hat{\pi}_n|\psi\rangle h_n \tag{1.5.157}$$

where $h_n$ is a function of the meter-environment variables distinct for each $n$. By each $h_n$ being distinct, the meter will have a distinct response on each subspace. During the interaction, $\mathcal{H}_{\text{int}}$ must completely dominate over the free Hamiltonians in order for the system and the detector to become properly correlated and so for the short time that the interaction happens, $\mathcal{H}_{\text{int}}$ is effectively the total Hamiltonian. Eigenstates of the Hamiltonian operator $\sum_n\hat{\pi}_nh_n$ are precisely the eigenstates $|n,a\rangle$ of the $\hat{\pi}_n$ and the adiabatic invariants of

this Hamiltonian $|\langle n, a|\psi\rangle|^2$ are precisely the ingredients of the Born rule. In fact, it is because these Born rule intensities are adiabatically invariant i.e. stable enough on experimental time scales that they are the only possible quantities to manifest as effective macroscopic probabilities in the Born rule.

You might wonder to what extent it is possible to simultaneously (in one experimental trial) measure two properties of a single instance of a quantum system. Because the projectors $\hat{\pi}_n$ have to be selected based on the associated predictable response of the meter and the property $G = \langle \hat{G} \rangle$ of the quantum system which is measured is built after this selection in terms of those projectors, it is only self-consistent to infer values of another property $G' = \langle \hat{G}' \rangle$ as well if it can also be built out of those projectors in the same measurement process. Therefore, $\hat{G}$ and $\hat{G}'$ must commute (which is automatically true if they are both built out of the same $\hat{\pi}_n$). Conversely, if $\hat{G}$ and $\hat{G}'$ commute, it is possible to simultaneously diagonalize them and thus build a simultaneous set of eigenspace projectors around which an appropriate detector could be designed. That is, two properties are simultaneously measurable in a single ideal experimental instance if and only if the corresponding generating operators to those properties commute.

The inability in an ideal measurement to simultaneously measure properties corresponding to noncommuting operators emphasizes an important point: **probabilities in quantum mechanics are contextual**. What does this mean? Suppose that the probabilities of quantum measurement actually existed regardless of measurement or experimental context. Then, it would be possible to define a single probability space in which all of the probabilities of quantum mechanics live. It would be possible to answer the question of what is the probability of simultaneously reading off any two particular meter values corresponding to different properties. However, if the two properties do not commute there is no shared eigenspace, in contradiction with the conditions under which a meter value is read. Noncommuting operators indicate properties that cannot coexist in a single probability space. This makes perfect sense as it is the chaotic dynamical measurement process from which probabilities emerge. Measurements of different properties require physically different detector setups. We will see rigorously when we study Bell inequalities next chapter that generically for a quantum system it is mathematically impossible to define a probability space that can simultaneously describe all of the results of different possible measurements of that system when considering measurements that cannot be performed simultaneously. Attempting to put the properties corresponding to noncommuting operators into a shared probability space leads to genuine mathematical contradictions.

## Exercises

**Problem 1.12: ⋆ Commutators becoming derivatives.**

(a) Show by repeating and generalizing the arguments between equations (1.5.118) and (1.5.121) that a generic operator $\hat{F}$ on $\mathbb{H}_{\mathbb{S}}$ may be written as a function $f(t, \hat{\vec{x}}, \hat{\vec{p}})$ (replace $\hat{H}$ in those equations by $\hat{F}$ and define $f$ the way $H$ is defined, then complete the argument).

(b) Show that if $\hat{F}$ commutes with $\hat{\vec{x}}$ then it is not a function of $\hat{\vec{p}}$.

(c) Show that if $\hat{F}$ commutes with $\hat{\vec{p}}$ then it is not a function of $\hat{\vec{x}}$.

(d) Show that if $\hat{F}$ is not a function of $\hat{\vec{x}}$ then:

$$[\hat{\vec{x}}, \hat{F}] = i\hbar \frac{\partial \hat{F}}{\partial \hat{\vec{p}}}.$$

(e) Show that if $\hat{F}$ is not a function of $\hat{\vec{p}}$ then:

$$[\hat{\vec{p}}, \hat{F}] = -i\hbar \frac{\partial \hat{F}}{\partial \hat{\vec{x}}}.$$

If the operator derivative is defined appropriately to take care of commutation difficulties, the results of this and the previous part actual hold for $\hat{F}$ which depend on both $\hat{\vec{x}}$ and $\hat{\vec{p}}$.

**Problem 1.13: ⋆ Restoring factors of $\hbar$.** Occasionally, you have to compare equations with someone who prefers to use unit systems in which $\hbar$ is not set to 1 by definition. Suppose you have a quantity of interest $F$ whose units are known in Schrödinger Natural Units (in which $\hbar = \epsilon_0 = 1$) to be $[F]_{\text{SNU}} = T^\alpha L^\beta$ where $L$ indicates units of length and $T$ indicates units of time and known in your friend's unit system (for example SI units) to be $[F]_{\text{Friend}} = M^\gamma T^\lambda L^\mu Q^\nu$ where $L$ indicates units of length and $Q$ indicates units of charge. Show that:

$$\gamma + \lambda - \frac{\nu}{2} = \alpha, \qquad -2\gamma + \mu + \frac{\nu}{2} = \beta$$

and that $F$ in your two unit systems are related by:

$$F_{\text{Friend}} = \hbar^{\gamma+\frac{\nu}{2}} \epsilon_0^{\frac{\nu}{2}} F_{SNU}$$

with $\hbar$ and $\epsilon_0$ being their respective values in your friend's unit system. Explain then that applying this to every quantity of interest in an equation allows you to restore powers of $\hbar$ and $\epsilon_0$ in equations whenever you would like to and thus you have lost nothing by choosing units which eliminate them.

**Problem 1.14: Understanding Natural Units.** From the defined values of $\hbar = c = \epsilon_0 = k_B = 1$ in natural units and the defined numerical values of $\hbar, c, k_B, q_e$ in SI units, together with the measured value of the fine-structure constant $\alpha$ (defined in problem 1.52), reproduce the values in table (1.5.134).

**Problem 1.15: Dilation Operator.** Consider a system in 1 spatial dimension so that $\mathbb{H}$ is the space of square-integrable complex functions of a single real variable. For constant $\bar{x}$ and $\bar{p}$, define the dilation generator $\hat{G}$ and dilation operator $\hat{D}(\varepsilon)$ by:

$$\hat{G} = \frac{(\hat{x} - \mathbb{1}\bar{x})(\hat{p} - \mathbb{1}\bar{p}) + (\hat{p} - \mathbb{1}\bar{p})(\hat{x} - \mathbb{1}\bar{x})}{2}, \qquad \hat{D}(\varepsilon) = e^{-\frac{i}{\hbar}\varepsilon\hat{G}}.$$

(a) Define $|\psi(\varepsilon)\rangle = \hat{D}(\varepsilon)|\psi_0\rangle$ for a fixed normalized wavefunction $|\psi_0\rangle$. Show that:

$$i\hbar\frac{d|\psi(\varepsilon)\rangle}{d\varepsilon} = \hat{G}|\psi\rangle.$$

(b) Show that the above in position space becomes:

$$\frac{\partial\psi}{\partial\varepsilon} + (x - \bar{x})\frac{\partial\psi}{\partial x} - \frac{i}{\hbar}\bar{p}(x - \bar{x})\psi + \frac{\psi}{2} = 0$$

and show that this can be rewritten as:

$$\frac{\partial}{\partial\varepsilon}\left(e^{\frac{\varepsilon}{2} - \frac{i}{\hbar}\bar{p}x}\psi\right) + (x - \bar{x})\frac{\partial}{\partial x}\left(e^{\frac{\varepsilon}{2} - \frac{i}{\hbar}\bar{p}x}\psi\right) = 0.$$

Show that this is solved by:

$$\psi(\varepsilon, x) = e^{-\frac{\varepsilon}{2} + \frac{i}{\hbar}(1 - e^{-\varepsilon})\bar{p}(x - \bar{x})}\psi(0, \bar{x} + e^{-\varepsilon}(x - \bar{x})).$$

(c) Suppose we choose the constants $\bar{x}$ and $\bar{p}$ so that they equal $\langle\hat{x}\rangle$ and $\langle\hat{p}\rangle$ when $|\psi_0\rangle$ is used to compute these operator summaries. Show that for all $\varepsilon$:

$$\langle\psi(\varepsilon)|\hat{x}|\psi(\varepsilon)\rangle = \bar{x}, \qquad \langle\psi(\varepsilon)|\hat{p}|\psi(\varepsilon)\rangle = \bar{p}.$$

(d) Show that:

$$\frac{d}{d\varepsilon}(\hat{D}^\dagger(\hat{x} - \mathbb{1}\bar{x})\hat{D}) = \hat{D}^\dagger(\hat{x} - \mathbb{1}\bar{x})\hat{D}, \qquad \frac{d}{d\varepsilon}(\hat{D}^\dagger(\hat{p} - \mathbb{1}\bar{p})\hat{D}) = -\hat{D}^\dagger(\hat{p} - \mathbb{1}\bar{p})\hat{D}$$

and that these are solved by:

$$\hat{D}^\dagger(\hat{x} - \mathbb{1}\bar{x})\hat{D} = e^{\varepsilon}(\hat{x} - \mathbb{1}\bar{x}), \qquad \hat{D}^\dagger(\hat{p} - \mathbb{1}\bar{p})\hat{D} = e^{-\varepsilon}(\hat{p} - \mathbb{1}\bar{p}).$$

(e) Now define the central moments $M_{mn}$ by:

$$M_{mn}(\varepsilon) = \langle\psi(\varepsilon)|\text{Sym}((\hat{x} - \bar{x}\mathbb{1}); m, (\hat{p} - \bar{p}\mathbb{1}); n)|\psi(\varepsilon)\rangle.$$

Here in the symmetrized product the $; m$ indicates that they are $m$ factors of $\hat{x} - \bar{x}\mathbb{1}$ and $; n$ indicates there are $n$ factors of $\hat{p} - \bar{p}\mathbb{1}$. Show that:

$$M_{mn}(\varepsilon) = e^{\varepsilon(m-n)}M_{mn}(0).$$

Use this result to explain how the dilation operator stretches the wavefunction out uniformly in position space and squeezes it in uniformly in momentum space. Explain further why if a transformation does one of these then it must do the other as well.

**Problem 1.16: Nondegenerate Time Independent Perturbation Theory.** As usual, most systems are too hard to solve exactly. Suppose we have a time independent Hamiltonian operator $\hat{H}_0$ whose eigenvector problem we can solve exactly with orthonormal eigenvectors $|n_0\rangle$ and eigenvalues $E_n^{(0)}$. Assume for this question that none of the eigenvalues of $\hat{H}_0$ are repeated. Suppose that the actual Hamiltonian operator $\hat{H}$ for our system is time independent and related to $\hat{H}_0$ by:

$$\hat{H} = \hat{H}_0 + \varepsilon\hat{H}_1$$

for a small parameter $\varepsilon$ and $\hat{H}_1$ some operator. The goal of this question is to find the first few terms in powers of $\varepsilon$ for the solution of the eigenvectors $|n\rangle$ and eigenvalues $E_n$ of $\hat{H}$. These series will take the form:

$$E_n = E_n^{(0)} + \varepsilon E_n^{(1)} + \varepsilon^2 E_n^{(2)} + \mathcal{O}(\varepsilon^3), \qquad |n\rangle = |n_0\rangle + \varepsilon|n_1\rangle + \varepsilon^2|n_2\rangle + \mathcal{O}(\varepsilon^3).$$

(a) Expand the eigenvector equation for $\hat{H}$ to order $\varepsilon^2$ to show that:

$$\hat{H}_0|n_1\rangle + \hat{H}_1|n_0\rangle = E_n^{(0)}|n_1\rangle + E_n^{(1)}|n_0\rangle, \qquad \hat{H}_0|n_2\rangle + \hat{H}_1|n_1\rangle = E_n^{(0)}|n_2\rangle + E_n^{(1)}|n_1\rangle + E_n^{(2)}|n_0\rangle.$$

(b) The phase of eigenvectors can in general be chosen arbitrarily. Show that we can choose the phase of $|n\rangle$ so that:

$$\langle n_0|n_1\rangle = 0.$$

(c) Assume we have made the above phase choice. Show that the first two energy corrections and first state correction are given by:

$$E_n^{(1)} = \langle n_0|\hat{H}_1|n_0\rangle, \qquad E_n^{(2)} = \sum_{m\neq n} \frac{|\langle n_0|\hat{H}_1|m_0\rangle|^2}{E_n^{(0)} - E_m^{(0)}}, \qquad |n_1\rangle = \sum_{m\neq n} \frac{\langle m_0|\hat{H}_1|n_0\rangle}{E_n^{(0)} - E_m^{(0)}}|m_0\rangle.$$

**Problem 1.17: Degenerate Time Independent Perturbation Theory.** Consider the same setup as the previous problem except now there may be repeats among the eigenvalues of $\hat{H}_0$. Because of this, the eigenvectors of $\hat{H}_0$ are not uniquely defined (as linear combinations of eigenvectors with the same eigenvalue remain an eigenvector). But, only certain choices of eigenvectors allow the $\varepsilon$ series is exist and precisely which choice is necessary is dependent upon $\hat{H}_1$. Assume that we choose the phase of $|n_0\rangle$ so that $\langle n_0|n_1\rangle = 0$.

(a) Show that:

$$E_n^{(1)} = \langle n_0|\hat{H}_1|n_0\rangle.$$

(b) Show that for $m \neq n$ and $E_m^{(0)} \neq E_n^{(0)}$:

$$\langle m_0|n_1\rangle = \frac{\langle m_0|\hat{H}_1|n_0\rangle}{E_n^{(0)} - E_m^{(0)}}.$$

(c) Show that for $m \neq n$ and $E_m^{(0)} = E_n^{(0)}$:

$$\langle m_0|\hat{H}_1|n_0\rangle = 0.$$

Explain that this implies that the unperturbed eigenvectors must be selected so that the perturbation $\hat{H}_1$ is diagonalized on each eigenspace of $\hat{H}_0$. Often this is enough to fully specify the proper choice of unperturbed eigenvectors. However, if there exist $n, m$ so that $E_n^{(0)} = E_m^{(0)}$ and $E_n^{(1)} = E_m^{(1)}$ then the degeneracy is said not to be lifted at first order and the proper choice of unperturbed eigenvectors is still not determined. In that case one must keep going to higher orders until enough conditions are placed to fully specify the unperturbed eigenvectors.

(d) Show that:

$$|n_1\rangle = \sum_{E_m^{(0)}\neq E_n^{(0)}} \frac{\langle m_0|\hat{H}_1|n_0\rangle}{E_n^{(0)} - E_m^{(0)}}|m_0\rangle + \sum_{\substack{m\neq n \\ E_m^{(0)}=E_n^{(0)}}} \sum_{E_k^{(0)}\neq E_n^{(0)}} \frac{\langle m_0|\hat{H}_1|k_0\rangle\langle k_0|\hat{H}_1|n_0\rangle}{\langle n_0|\hat{H}_1|n_0\rangle(E_n^{(0)} - E_k^{(0)})}|m_0\rangle$$

$$E_n^{(2)} = \sum_{E_m^{(0)}\neq E_n^{(0)}} \frac{|\langle n_0|\hat{H}_1|m_0\rangle|^2}{E_n^{(0)} - E_m^{(0)}}.$$

**Problem 1.18: Time Dependent Perturbation Theory.** Consider a system with a Hamiltonian operator of the form $\hat{H}(t) = \hat{H}_0 + \varepsilon\hat{H}_1(t)$ for a small parameter $\varepsilon$ and $\hat{H}_0$ a time independent operator whose eigenvector problem we can solve exactly with orthonormal eigenvectors $|n\rangle$ and eigenvalues $E_n$. Importantly, note that in this question $|n\rangle$ and $E_n$ are the eigenvectors and eigenvalues of $\hat{H}_0$ and we do not care about the eigenvectors or eigenvalues of $\hat{H}$. Define the coefficients $a_n(t)$ by:

$$a_n(t) = e^{\frac{i}{\hbar}E_n t}\langle n|\psi(t)\rangle.$$

Show that $|\psi(t)\rangle$ can be written as:

$$|\psi(t)\rangle = \sum_n e^{-\frac{i}{\hbar}E_n t} a_n(t)|n\rangle.$$

Further, show from the Schrödinger equation that $a_n(t)$ satisfies:

$$i\hbar\frac{da_n(t)}{dt} = \varepsilon\sum_m e^{\frac{i}{\hbar}(E_n-E_m)t}\langle n|\hat{H}_1(t)|m\rangle a_m(t).$$

Explain how this equation can be used to develop a series solution for $a_n(t)$ in powers of $\varepsilon$ and thus can be used to approximately solve the original Schrödinger equation.

**Problem 1.19: Fermi's Golden Rule (Approximate Form).** Consider a perturbation $\hat{H}_1$ which "turns on" at time $t = 0$ and which is constant until a time $t = \tau$ at which point it is "turned off" ($\hat{H}_1$ is a non-zero constant operator for $0 < t < \tau$ and 0 otherwise) so that the total Hamiltonian is $\hat{H} = \hat{H}_0 + \varepsilon\hat{H}_1$. Suppose that the system begins at $t = 0$ in an eigenstate $|\iota\rangle$ of $\hat{H}_0$ with energy eigenvalue $E_\iota$.

(a) Show that the $a_n$ coefficients of time dependent perturbation theory can be written as:

$$a_n(t) = \delta_{n\iota} + \varepsilon\delta a_n(t) + \mathcal{O}(\varepsilon^2)$$

where the $\delta a_n(t)$ coefficients satisfy:

$$\frac{d\delta a_n(t)}{dt} = -\frac{i}{\hbar}e^{\frac{i}{\hbar}(E_n-E_\iota)t}\langle n|\hat{H}_1|\iota\rangle$$

(b) Show that the $\delta a_n(t)$ are given by:

$$\delta a_n(t) = \frac{1-e^{\frac{i}{\hbar}(E_n-E_\iota)t}}{E_n-E_\iota}\langle n|\hat{H}_1|\iota\rangle.$$

(As written, this formula assumes $E_n \neq E_\iota$. Also find the correct formula for $E_n = E_\iota$ and show that the above formula correctly limits to that one in the $E_n \to E_\iota$ limit.)

(c) Let $P_{\iota\to f}$ be the "transition intensity" for beginning in the eigenstate $|\iota\rangle$ of $\hat{H}_0$ at time $t = 0$ and ending in the eigenstate $|f\rangle$ of $\hat{H}_0\rangle$ at time $t = \tau$ with $f \neq \iota$. That is, $P_{\iota\to f}$ is defined by:

$$P_{\iota\to f} = |\langle f|\psi(\tau)\rangle|^2.$$

Explain mathematically how $P_{\iota\to f}$ measures how close $|\psi(\tau)\rangle$ is to being $|f\rangle$, with a value of 1 indicating that they are identical and a value of 0 indicating that they are orthogonal. Show that to lowest order in perturbation theory:

$$P_{\iota\to f} = \frac{4\varepsilon^2|\langle f|\hat{H}_1|\iota\rangle|^2}{(E_f-E_\iota)^2}\sin^2\left(\frac{(E_f-E_\iota)\tau}{2\hbar}\right) + \mathcal{O}(\varepsilon^3)$$

The finite time transition rate $\gamma_{\iota\to f}$ is defined by the time rate of change of the transition intensity as we leave the perturbation on for slightly longer. Show that:

$$\gamma_{\iota\to f} \stackrel{\text{def}}{=} \frac{dP_{\iota\to f}}{d\tau} = \frac{2\varepsilon^2|\langle f|\hat{H}_1|\iota\rangle|^2}{\hbar(E_f-E_\iota)}\sin\left(\frac{E_f-E_\iota}{\hbar}\tau\right) + \mathcal{O}(\varepsilon^3).$$

(d) Show for an arbitrary smooth function $F(x)$ (by making the integral change of variables $y = \alpha x$) that:

$$\lim_{\alpha\to\infty}\int_{-\infty}^{\infty}\frac{\sin(\alpha x)}{x}F(x)dx = F(0)\int_{-\infty}^{\infty}\frac{\sin x}{x}dx.$$

(e) Define the function $G(\beta)$ by:

$$G(\beta) = \int_0^\infty \frac{\sin x}{x}e^{-\beta x}dx.$$

Show that:

$$G(0) = \frac{1}{2}\int_{-\infty}^{\infty}\frac{\sin x}{x}dx, \qquad \lim_{\beta\to\infty}G(\beta) = 0, \qquad \frac{dG}{d\beta} = \int_0^\infty e^{-\beta x}\sin x dx = -\frac{1}{1+\beta^2}.$$

Use these to show:

$$\int_{-\infty}^{\infty}\frac{\sin x}{x}dx = \pi.$$

(f) Use the results of the previous two parts to show:

$$\lim_{\alpha\to\infty}\frac{\sin(\alpha x)}{x} = \pi\delta(x)$$

(g) Consider the long term behavior of the finite time transition rate $\gamma_{\iota\to f}$ for $\tau \to \infty$. Show using the

result of the previous part that the long time transition rate (also simply called the transition rate) $\Gamma_{\iota\to f}$ is given by:

$$\Gamma_{\iota\to f} = \lim_{\tau\to\infty} \gamma_{\iota\to f} = \frac{2\pi}{\hbar}\varepsilon^2 |\langle f|\hat{H}_1|\iota\rangle|^2 \delta(E_f - E_\iota) + \mathcal{O}(\varepsilon^3)$$

This formula is the approximate form of Fermi's Golden Rule. We see that it gives the lowest order in perturbation theory estimate of the stabilized rate at which a state nearly approximated by $|\iota\rangle$ transitions to a state nearly approximated by $|f\rangle$ under the perturbation $\varepsilon\hat{H}_1$. Later in this chapter when we study scattering, we will derive the exact form of Fermi's Golden Rule for transition rates.

**Problem 1.20: Positive Operators.**

(a) Let $\hat{\mathcal{O}}$ be a self-adjoint operator such that for all $|\psi\rangle \in \mathbb{H}$:

$$\langle\psi|\hat{\mathcal{O}}|\psi\rangle \geq 0.$$

Such an $\hat{\mathcal{O}}$ is called nonnegative. Show that all of the eigenvalues of $\hat{\mathcal{O}}$ are nonnegative. Show that if the above $\geq$ can be replaced by $>$ then all of the eigenvalues of $\hat{\mathcal{O}}$ are positive (such an operator is called positive).

(b) Show that if $\geq$ above cannot be replaced with $>$ then $\hat{\mathcal{O}}$ must have 0 as an eigenvalue and that 0 is its least eigenvalue. Thus, show that for a self-adjoint operator $\hat{V}$ if $v$ is the greatest real number such that:

$$\langle\psi|\hat{V}|\psi\rangle \geq v$$

for all $|\psi\rangle \in \mathbb{H}$ then $v$ is an eigenvalue of $\hat{V}$ and $v$ is the lowest eigenvalue of $\hat{V}$. An operator $\hat{V}$ for which such a $v$ exists is said to be bounded below.

**Problem 1.21: ⋆ Baker–Campbell–Hausdorff formula.** For operators $\hat{A}$ and $\hat{B}$ which do not commute:

$$e^{\hat{A}}e^{\hat{B}} = e^{\hat{C}} \neq e^{\hat{A}+\hat{B}}$$

for some $\hat{C}$. The goal of this question is to find the first few terms of a formula for $\hat{C}$ (known as the BCH formula).

(a) Consider the function $\hat{C}(\varepsilon)$ defined by:

$$e^{\hat{C}(\varepsilon)} = e^{\varepsilon\hat{A}}e^{\varepsilon\hat{B}}.$$

Taylor expand both sides in powers of $\varepsilon$ to show:

$$\hat{C}(\varepsilon) = \varepsilon\hat{A} + \varepsilon\hat{B} + \frac{\varepsilon^2}{2}[\hat{A},\hat{B}] + \frac{\varepsilon^3}{12}[\hat{A}-\hat{B},[\hat{A},\hat{B}]] + \mathcal{O}(\varepsilon^4).$$

The all orders BCH formula is quite complicated.

(b) The above series does not terminate for general operators and the exact expression for $\hat{C}$ is very complicated. However, assume the special case that:

$$[\hat{A},[\hat{A},\hat{B}]] = [\hat{B},[\hat{A},\hat{B}]] = 0.$$

Show that in this case, we have the exact "simple" BCH formula:

$$e^{\hat{A}}e^{\hat{B}} = e^{\hat{A}+\hat{B}+\frac{1}{2}[\hat{A},\hat{B}]}.$$

**Problem 1.22: General Gaussian Integrals.** Recall the result of the single variable Gaussian integral from Orig. problem 3.7.(m) that, for complex $a, b, c$ satisfying $\mathrm{Re}(a) > 0$ or if $\mathrm{Re}(b) = 0$ then also for $\mathrm{Re}(a) = 0$:

$$\int_{-\infty}^{\infty} e^{-\frac{a}{2}x^2+bx+c}dx = \sqrt{\frac{2\pi}{a}}e^{\frac{b^2}{2a}+c}.$$

(a) Show that:

$$\int_{\mathbb{R}^n} e^{-\frac{1}{2}A_{ab}x^a x^b + B_a x^a + C} d^n\vec{x} = \frac{(2\pi)^{\frac{n}{2}}}{\sqrt{\det A}} e^{\frac{1}{2}A^{-ab}B_a B_b + C}$$

for a real symmetric positive definite matrix $A$ with components $A_{ab}$, complex vector $B$ with components $B_a$, and a complex number $C$. Do so by diagonalizing $A_{ab}$ and changing coordinates in the integral to show that the integral becomes a product of $n$ independent single variable Gaussian integrals, which you know how to perform.

(b) Use analogous techniques to show:

$$\int_{\mathbb{R}^n} e^{-\frac{i}{2}A_{ab}x^a x^b + iB_a x^a + iC} d^n\vec{x} = \frac{(2\pi)^{\frac{n}{2}}}{i^{\frac{n}{2}}\sqrt{\det A}} e^{\frac{i}{2}A^{-ab}B_a B_b + iC}$$

for a real symmetric nondegenerate matrix with components $A_{ab}$, a real vector with components $B_a$, and complex number $C$.

**Problem 1.23: Wavefunction of Minimum Uncertainty is a Gaussian.** Consider a quantum system in 1 spatial dimension. Due to the uncertainty theorem we know the minimum value possible for $\sigma_x\sigma_p$ among all wavefunctions in $\mathbb{H}$ is $\frac{\hbar}{2}$. In this question we will find all wavefunctions which achieve this minimum value.

(a) For a given normalized wavefunction $|\psi\rangle$, let $\mathfrak{q} = \langle\hat{x}\rangle$ and $\mathfrak{p} = \langle\hat{p}\rangle$. With $|u\rangle = (\hat{x} - \mathfrak{q}\mathbb{1})|\psi\rangle$ and $|v\rangle = (\hat{p} - \mathfrak{p}\mathbb{1})|\psi\rangle$, use the Cauchy-Schwarz inequality to show that:

$$\sigma_x\sigma_p = \frac{\hbar}{2} \iff \mathrm{Cor}(\hat{x},\hat{p}) = 0 \text{ and } \exists\ \ell \in \mathbb{C}\ :\ |v\rangle = \frac{i\hbar}{2\ell^2}|u\rangle.$$

(b) Show that the final condition of the previous part, when put into position space components becomes the differential equation:

$$\frac{\partial\psi}{\partial x} - \frac{i}{\hbar}\mathfrak{p}\psi + \frac{x-\mathfrak{q}}{2\ell^2}\psi = 0.$$

Solve this differential equation to show that for some real constants $N, \varphi$:

$$\psi(x) = Ne^{-\frac{(x-\mathfrak{q})^2}{4\ell^2} + \frac{i}{\hbar}\mathfrak{p}x + i\varphi}.$$

(c) Show that $\psi(x)$ is only of finite norm if $\ell \in \mathbb{R}$ (and take, without loss of generality, $\ell > 0$). Show that we can choose $N$ and $\varphi$ so that $\langle\psi|\psi\rangle = 1$ and so that $\psi(x)$ becomes:

$$\psi(x) = \frac{1}{(2\pi\ell^2)^{\frac{1}{4}}} e^{-\frac{(x-\mathfrak{q})^2}{4\ell^2} + \frac{i}{\hbar}\mathfrak{p}x - \frac{i\mathfrak{q}\mathfrak{p}}{2\hbar}}.$$

Confirm that $\langle\hat{x}\rangle = \mathfrak{q}$, $\langle\hat{p}\rangle = \mathfrak{p}$, $\mathrm{Cor}(\hat{x},\hat{p}) = 0$, and $\sigma_x\sigma_p = \frac{\hbar}{2}$. Show as well that $\sigma_x = \ell$, giving the interpretation of $\ell$. Therefore, a wavefunction is of minimal uncertainty if and only if it takes the above form, known as a Gaussian or coherent state.

**Problem 1.24: ⋆ First Look at Coherent States.** Let $|\mathfrak{q},\mathfrak{p}\rangle$ be the coherent state with $\langle\hat{x}\rangle = \mathfrak{q}$, $\langle\hat{p}\rangle = \mathfrak{p}$ and $\sigma_x = \ell$ (with $\ell$ fixed throughout this problem). That is, in position space:

$$\langle x|\mathfrak{q},\mathfrak{p}\rangle = \frac{1}{(2\pi\ell^2)^{\frac{1}{4}}} e^{-\frac{(x-\mathfrak{q})^2}{4\ell^2} + \frac{i}{\hbar}\mathfrak{p}x - \frac{i\mathfrak{q}\mathfrak{p}}{2\hbar}}.$$

Because coherent states are maximally localized in momentum space for a given amount of localization in position space, they are "maximally classical".

(a) Show that the coherent states at fixed $\ell$ satisfy:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\mathfrak{q},\mathfrak{p}\rangle\langle\mathfrak{q},\mathfrak{p}| \frac{d\mathfrak{q}d\mathfrak{p}}{2\pi\hbar} = \mathbb{1}.$$

This looks like a completeness relation but it should properly be called an over-completeness relation because the coherent states are **NOT** a basis. They are over-complete in the sense that they are not a linearly independent set. Explain how it is possible to fully determine an operator $\hat{\mathcal{O}}$ from its coherent state matrix elements $\langle\mathfrak{q},\mathfrak{p}|\hat{\mathcal{O}}|\mathfrak{q}'\mathfrak{p}'\rangle$. In fact, it is actually possible to completely recover the operator from its diagonal coherent state elements $\langle\mathfrak{q},\mathfrak{p}|\hat{\mathcal{O}}|\mathfrak{q},\mathfrak{p}\rangle$, although doing so is more difficult.

(b) Show that the coherent states have the inner products:

$$\langle\mathfrak{q},\mathfrak{p}|\mathfrak{q}',\mathfrak{p}'\rangle = e^{-\frac{(\mathfrak{q}-\mathfrak{q}')^2}{8\ell^2} - \frac{\ell^2(\mathfrak{p}-\mathfrak{p}')^2}{2\hbar^2} + \frac{i}{2\hbar}(\mathfrak{q}\mathfrak{p}' - \mathfrak{q}'\mathfrak{p})}.$$

Thus, the coherent states are approximately orthogonal for significantly different $(\mathfrak{q},\mathfrak{p})$ and $(\mathfrak{q}',\mathfrak{p}')$ (where significantly different is determined by $\ell,\hbar$).

## 1.6 Scalar Theory - Galilean Schrödinger Equation

We are finally ready to look at our first explicit theory of matter without any point particles. Ultimately, we know we will have to consider quantum mechanics with Lorentz symmetry, because Galilean boost

symmetry is not fundamental, but it is very fruitful to learn about Galilean Quantum Mechanics first. Since all we are changing from Galilean Classical Mechanics in the Orig. is that we are taking a wave theory of matter, we take as our assumptions exactly those of GCM with a single replacement:

**Assumption :: GQMI - Matter is wave-like**

A unit of matter is described by a matter wave field called the wavefunction $\psi(t, \vec{x})$.

**Assumption :: GQMII - Galilean boost symmetry**

Isolated systems have symmetry under shifting the position of all objects at the same fixed rate (boosting by a constant velocity) while keeping time unchanged.

**Assumption :: GQMIII - Space is static and Euclidean**

The background space on which waves propagate is not itself dynamical and is instead fixed without any curvature or holes, extending infinitely in all directions.

**Assumption :: GQMIV: Parity Symmetry**

Isolated systems have parity symmetry.

We first consider a scalar particle with no internal degrees of freedom so that $\mathbb{H} = \mathbb{H}_{\mathbb{S}}$. Recall from Orig. section 2.1 that the Poisson brackets for the classical Galilean theory are:

$$\{\vec{x}, \vec{G}\} = 0, \qquad \{\vec{p}, \vec{G}\} = m\mathbb{1}, \qquad \{\vec{v}, \vec{G}\} = \mathbb{1} \tag{1.6.1}$$

where $\vec{v}$ is the velocity and $\vec{G}$ is the generator of Galilean boosts at the time the boost is applied. We found that the classical position is given by $\langle\hat{\vec{x}}\rangle$ and the classical momentum by $\langle\hat{\vec{p}}\rangle$. We also know from Theorem 1.4 that as generators, whatever the effective classical $\vec{v}$ and $\vec{G}$ are they must be expressible as $\langle\hat{\vec{v}}\rangle$ and $\langle\hat{\vec{G}}\rangle$ for some self-adjoint $\hat{\vec{v}}$ and $\hat{\vec{G}}$. Then, because we know that the Poisson brackets induce the same algebra onto the operators by Theorem 1.5 as the original quantities:

$$[\hat{\vec{x}}, \hat{\vec{G}}] = 0, \qquad [\hat{\vec{p}}, \hat{\vec{G}}] = i\hbar m\mathbb{1}, \qquad [\hat{\vec{v}}, \hat{\vec{G}}] = i\hbar\mathbb{1}. \tag{1.6.2}$$

By the result of problem 1.12, $\hat{\vec{G}}$ can be written as a function $t, \hat{\vec{x}}, \hat{\vec{p}}$ and because $\hat{\vec{G}}$ commutes with $\hat{\vec{x}}$, it cannot depend on $\hat{\vec{p}}$ (which does not commute with $\hat{\vec{x}}$). Therefore:

$$[\hat{\vec{x}}, \hat{\vec{G}}] = 0 \implies \hat{\vec{G}} = \hat{\vec{G}}(t, \hat{\vec{x}}) \tag{1.6.3}$$

$$[\hat{\vec{p}}, \hat{\vec{G}}] = i\hbar m\mathbb{1} = m[\hat{\vec{x}}, \hat{\vec{p}}] \implies [\hat{\vec{p}}, \hat{\vec{G}} + m\hat{\vec{x}}] = 0. \tag{1.6.4}$$

Similarly, because $\hat{\vec{G}} + m\hat{\vec{x}}$ commutes with $\hat{\vec{p}}$, it cannot be a function of $\hat{\vec{x}}$. Thus, $\hat{\vec{G}} + m\hat{\vec{x}}$ can only be a function of time times the identity operator:

$$\hat{\vec{G}} = -m\hat{\vec{x}} + \vec{a}(t)\mathbb{1}. \tag{1.6.5}$$

$$[\hat{\vec{v}}, \hat{\vec{G}}] = i\hbar\mathbb{1} \implies [\hat{\vec{v}}, -m\hat{\vec{x}} + \vec{a}\mathbb{1}] = [\hat{\vec{x}}, \hat{\vec{p}}] \implies [\hat{\vec{x}}, \hat{\vec{p}} - m\hat{\vec{v}}] = 0. \tag{1.6.6}$$

$\hat{\vec{p}} - m\hat{\vec{v}}$ commutes with $\hat{\vec{x}}$, so it cannot be a function of $\hat{\vec{p}}$. Hence, there exists a function $\hat{\vec{K}}(t, \hat{\vec{x}})$ so that:

$$\hat{\vec{v}} = \frac{\hat{\vec{p}} - \hat{\vec{K}}(t, \hat{\vec{x}})}{m}. \tag{1.6.7}$$

Now we will find an explicit relationship between the velocity operator and the Hamiltonian. We intend $\vec{v} = \langle\hat{\vec{v}}\rangle$ to represent the best effective classical position of the wavefunction, but because $\langle\hat{\vec{x}}\rangle$ is the best

classical position, we ought to just take $\vec{v} = \frac{d\langle\hat{\vec{x}}\rangle}{dt}$. Then, using Ehrenfest's theorem:

$$\langle\hat{\vec{v}}\rangle = \frac{d\langle\hat{\vec{x}}\rangle}{dt} = \langle\frac{1}{i\hbar}[\hat{\vec{x}}, \hat{H}]\rangle \implies \hat{\vec{v}} = \frac{1}{i\hbar}[\hat{\vec{x}}, \hat{H}] \tag{1.6.8}$$

where the last step follows because the previous must hold for all $|\psi\rangle$. Together (1.6.7) and (1.6.8) give:

$$\frac{1}{i\hbar}[\hat{\vec{x}}, \hat{H}] = \frac{\hat{\vec{p}} - \hat{\vec{K}}(t, \hat{\vec{x}})}{m}. \tag{1.6.9}$$

It is not obvious how to solve this equation for $\hat{H}$ in general. But, the general solution can be constructed from a reasonably guessed particular solution. Consider $\hat{H}_0$:

$$\hat{H}_0 = \frac{(\hat{\vec{p}} - \hat{\vec{K}})^2}{2m} \tag{1.6.10}$$

$$\implies \frac{1}{i\hbar}[\hat{\vec{x}}, \hat{H}_0] = \frac{1}{i\hbar}[\hat{\vec{x}}, \frac{\hat{\vec{p}}^2 - \hat{\vec{p}}\cdot\hat{\vec{K}} - \hat{\vec{K}}\cdot\hat{\vec{p}} + \hat{\vec{K}}^2}{2m}] \tag{1.6.11}$$

$$= \frac{1}{2mi\hbar}[\hat{\vec{x}}, \hat{\vec{p}}^2] - \frac{1}{2mi\hbar}[\hat{\vec{x}}, \hat{\vec{p}}]\cdot\hat{\vec{K}} - \frac{1}{2mi\hbar}\hat{\vec{K}}\cdot[\hat{\vec{x}}, \hat{\vec{p}}] + 0 \tag{1.6.12}$$

$$= \frac{1}{2mi\hbar}\left([\hat{\vec{x}}, \hat{\vec{p}}]\cdot\hat{\vec{p}} + \hat{\vec{p}}\cdot[\hat{\vec{x}}, \hat{\vec{p}}]\right) - \frac{1}{2mi\hbar}i\hbar\mathbb{1}\cdot\hat{\vec{K}} - \frac{1}{2mi\hbar}\hat{\vec{K}}\cdot i\hbar\mathbb{1} \tag{1.6.13}$$

$$= \frac{1}{2mi\hbar}\left(i\hbar\mathbb{1}\cdot\hat{\vec{p}} + \hat{\vec{p}}\cdot i\hbar\mathbb{1}\right) - \frac{\hat{\vec{K}}}{m} \tag{1.6.14}$$

$$= \frac{\hat{\vec{p}} - \hat{\vec{K}}}{m}. \tag{1.6.15}$$

Hence $\hat{H}_0$ is a particular solution to (1.6.9). This particular solution was a reasonable guess because commuting with a function of $\hat{\vec{p}}$ with $\hat{\vec{x}}$ just takes the $\hat{\vec{p}}$ derivative of that function and this $\hat{H}_0$ is an antiderivative of $\hat{\vec{v}}$ with respect to $\hat{\vec{p}}$. Now consider the difference between $\hat{H}$, a general solution, and $\hat{H}_0$:

$$\frac{1}{i\hbar}[\hat{\vec{x}}, \hat{H} - \hat{H}_0] = \frac{1}{i\hbar}[\hat{\vec{x}}, \hat{H}] - \frac{1}{i\hbar}[\hat{\vec{x}}, \hat{H}_0] = 0 \tag{1.6.16}$$

So, $\hat{H} - \hat{H}_0$ cannot be a function of $\hat{\vec{p}}$. This leaves $\hat{H}$ determined up to adding a function $\hat{U}(t, \hat{\vec{x}})$:

$$\hat{H} = \frac{(\hat{\vec{p}} - \hat{\vec{K}}(t, \hat{\vec{x}}))^2}{2m} + \hat{U}(t, \hat{\vec{x}}). \tag{1.6.17}$$

Thus (1.6.17) is the general solution for the Hamiltonian operator consistent with our necessary symmetries. Now that we know the Hamiltonian operator, we can compute the equation of motion via (1.2.28)

$$\mathcal{L} = i\hbar\langle\psi|\dot{\psi}\rangle - \langle\psi|\hat{H}|\psi\rangle \implies i\hbar|\dot{\psi}\rangle = \hat{H}|\psi\rangle \tag{1.6.18}$$

$$\implies \boxed{i\hbar\frac{\partial\psi}{\partial t}(t, \vec{x}) = \left(-\frac{\hbar^2}{2m}\left(\nabla - \frac{i}{\hbar}\vec{K}(t, \vec{x})\right)^2 + U(t, \vec{x})\right)\psi(t, \vec{x}).} \tag{1.6.19}$$

Equation (1.6.19) is an extremely successful model, the Galilean Schrödinger equation. The Hamiltonian operator appearing here has a strikingly similar form to Orig. Theorem 2.1. However, at the moment we cannot justify associating the parameter $m$ appearing here and the functions $\hat{U}$ and $\hat{\vec{K}}$ with the classical mass and potentials. To interpret these quantities we need to look at the classical limiting behavior of this equation.

Let us compute the dynamics of $\vec{x}_{\text{cl}}$ predicted by the Galilean Schrödinger equation:

$$\frac{d\langle\hat{\vec{x}}\rangle}{dt} = \frac{1}{i\hbar}\langle[\hat{\vec{x}}, \hat{H}]\rangle = \frac{\langle\hat{\vec{p}}\rangle - \langle\hat{\vec{K}}\rangle}{m} \tag{1.6.20}$$

$$\frac{d^2\langle\hat{\vec{x}}\rangle}{dt^2} = \frac{1}{m}\frac{d\langle\hat{\vec{p}}\rangle}{dt} - \frac{1}{m}\frac{d\langle\hat{\vec{K}}\rangle}{dt} \tag{1.6.21}$$

$$\frac{d\langle\hat{\vec{p}}\rangle}{dt} = \frac{1}{i\hbar}\langle[\hat{\vec{p}}, \hat{H}]\rangle \tag{1.6.22}$$

$$= \frac{1}{2mi\hbar}\left\langle[\hat{\vec{p}}, \hat{p}_a - \hat{K}_a](\hat{p}^a - \hat{K}^a) + (\hat{p}_a - \hat{K}_a)[\hat{\vec{p}}, \hat{p}^a - K^a]\right\rangle + \frac{1}{i\hbar}\langle[\hat{\vec{p}}, \hat{U}]\rangle \tag{1.6.23}$$

$$= \frac{1}{2}\left\langle \frac{\partial \hat{K}_a}{\partial \hat{\vec{x}}}\hat{v}^a + \hat{v}^a \frac{\partial \hat{K}_a}{\partial \hat{\vec{x}}} \right\rangle - \langle \nabla \hat{U} \rangle \tag{1.6.24}$$

$$\frac{d\langle \hat{\vec{K}} \rangle}{dt} = \frac{\partial \langle \hat{\vec{K}} \rangle}{\partial t} + \frac{1}{i\hbar}\langle [\hat{\vec{K}}, \hat{H}] \rangle = \frac{\partial \langle \hat{\vec{K}} \rangle}{\partial t} + \frac{1}{2}\langle \frac{\partial \hat{\vec{K}}}{\partial \hat{x}^a}\hat{v}^a + \hat{v}^a \frac{\partial \hat{\vec{K}}}{\partial \hat{x}^a} \rangle. \tag{1.6.25}$$

Collecting these together:

$$m\frac{d^2\langle \hat{\vec{x}} \rangle}{dt^2} = \langle -\frac{\partial \hat{\vec{K}}}{\partial t} - \nabla \hat{U} \rangle - \frac{1}{2}\left\langle (\nabla \times \hat{\vec{K}}) \times \hat{\vec{v}} \right\rangle + \frac{1}{2}\left\langle \hat{\vec{v}} \times (\nabla \times \hat{\vec{K}}) \right\rangle. \tag{1.6.26}$$

This is satisfied exactly for all wavefunctions in the theory.

Importantly, the operator summary of a function of an operator is **NOT** the same thing as that function of the operator summary. And so, in general this differential equation does not close (meaning it does not contain sufficient information to solve for $\vec{x}_{\text{cl}}$ from initial data about $\vec{x}_{\text{cl}}$ only). However, we can find what happens in the classical limit by considering a wavefunction which has approximately well defined position and velocity (as always, relative to a given experimental tolerance). For a highly localized wavefunction:

$$\langle f(\hat{\vec{x}}) \rangle = \langle \psi | f(\hat{\vec{x}}) | \psi \rangle = \int_{\mathbb{S}} f(\vec{x})|\psi|^2 d^3\vec{x} \approx f(\langle \hat{\vec{x}} \rangle) \int_{\mathbb{S}} |\psi|^2 d^3\vec{x} = f(\langle \hat{\vec{x}} \rangle)\langle \psi | \psi \rangle = f(\langle \hat{\vec{x}} \rangle) \tag{1.6.27}$$

because the only significant contributions to the integral come from the region near $\langle \hat{\vec{x}} \rangle$ and if the function $f$ does not vary much over this region then we can approximate its value anywhere on this region by its value at $\langle \hat{\vec{x}} \rangle$. This allows us to approximately move $f$ outside of the integral. Making the same approximation that the velocity is approximately well defined and so can be approximately replaced by its operator summary $\frac{d\langle \hat{\vec{x}} \rangle}{dt}$, we then have that in the classical limit:

$$m\ddot{\vec{x}}_{\text{cl}} \approx -\frac{\partial \vec{K}}{\partial t}(t, \vec{x}_{\text{cl}}) - \nabla U(t, \vec{x}_{\text{cl}}) + \dot{\vec{x}}_{\text{cl}} \times (\nabla \times \vec{K}(t, \vec{x}_{\text{cl}})) \tag{1.6.28}$$

Hence the center $\vec{x}_{\text{cl}}$ of a highly localized wavefunction approximately satisfies the classical equation of motion making the identification of $m$ with the classical mass and $U$ and $\vec{K}$ with the classical potentials. Again, typically no wavefunction exactly satisfies (1.6.28), but the fact that a highly localized one approximately does explains how the Galilean equation of motion was a successful macroscopic description of particles. Together we have:

**Theorem 1.9: The Galilean Schrödinger Equation**

In general, the Hamiltonian operator for a scalar Galilean Schrödinger wave is given by:

$$\hat{H} = \frac{(\hat{\vec{p}} - \hat{\vec{K}}(t, \hat{\vec{x}}))^2}{2m} + \hat{U}(t, \hat{\vec{x}}) \tag{1.6.29}$$

where $m$ is the mass of the wave and $U(t, \vec{x})$ and $\vec{K}(t, \vec{x})$ are the scalar and vector potentials experienced by the wave.

### 1.6.1 Interaction with Electromagnetism

In classical mechanics, when the potentials appearing in the Hamiltonian are due to electromagnetic fields, we identify $\vec{K} = q\vec{A}$ and $U = q\phi$ with $q$ the electric charge of the body and $\phi$ and the $\vec{A}$ the scalar and vector potentials. This identification is still true in quantum mechanics because the potentials in the Hamiltonian operator correspond to the potentials in the classical Hamiltonian (as identified in (1.6.28)). Then, for an electrically charged wavefunction of charge $q$:

$$\hat{H} = \frac{(\hat{\vec{p}} - q\hat{\vec{A}})^2}{2m} + q\hat{\phi} \tag{1.6.30}$$

where $\hat{\phi} = \phi(t, \hat{\vec{x}})$ and $\hat{\vec{A}} = \vec{A}(t, \hat{\vec{x}})$. We also know that the potentials are only actually defined up to the gauge transformation:

$$\vec{A}' = \vec{A} + \nabla f \qquad \phi' = \phi - \frac{\partial f}{\partial t}. \tag{1.6.31}$$

If we insert these transformed potentials into the Hamiltonian operator, we arrive at the transformed Hamiltonian operator:

$$\hat{H}' = \frac{(\hat{\vec{p}} - q\hat{\vec{A}}')^2}{2m} + q\hat{\phi}' = \frac{(\hat{\vec{p}} - q\hat{\vec{A}} - q\nabla\hat{f})^2}{2m} + q\hat{\phi} - q\frac{\partial \hat{f}}{\partial t} \tag{1.6.32}$$

where $\hat{f} = f(t, \hat{\vec{x}})$. In order for us to not violate the redundancy in the potentials, we need this transformed operator to be equivalent to our untransformed Hamiltonian operator up to symmetry transformation. Because the gauge transformation does not affect time, time is not reversed and so by Wigner's theorem there must exist a unitary operator $\hat{U}$ so that $\hat{H}$ and $|\psi\rangle$ are transformed by:

$$|\psi'\rangle = \hat{U}|\psi\rangle, \qquad \hat{H}' = \hat{U}\left(\hat{H} - i\hbar\frac{\partial}{\partial t}\right)\hat{U}^\dagger \tag{1.6.33}$$

For a generic wavefunction and with a bit of algebra:

$$\langle\vec{x}|\hat{H}'|\psi'\rangle = -\frac{\hbar^2}{2m}\left(\nabla - i\frac{q}{\hbar}\vec{A} - i\frac{q}{\hbar}\nabla f\right)^2\psi' + q\phi\psi' - q\frac{\partial f}{\partial t}\psi' \tag{1.6.34}$$

$$= \langle\hat{\vec{x}}|\left(e^{i\frac{q}{\hbar}\hat{f}}\hat{H}e^{-i\frac{q}{\hbar}\hat{f}} - i\hbar e^{i\frac{q}{\hbar}\hat{f}}\frac{\partial}{\partial t}\left(e^{-i\frac{q}{\hbar}\hat{f}}\right)\right)|\psi\rangle. \tag{1.6.35}$$

Therefore:

$$\hat{U} = e^{i\frac{q}{\hbar}\hat{f}} \implies \psi'(t,\vec{x}) = e^{i\frac{q}{\hbar}f(t,\vec{x})}\psi(t,\vec{x}). \tag{1.6.36}$$

This derivation, by just running Wigner's theorem, shows that our Lagrangian is then invariant under the combined transformation:

$$\psi'(t,\vec{x}) = e^{\frac{iq}{\hbar}f}\psi \qquad \vec{A}' = \vec{A} + \nabla f, \qquad \phi' = \phi - \frac{\partial f}{\partial t} \tag{1.6.37}$$

$$\implies \mathcal{L}' = i\hbar\langle\psi'|\dot{\psi}'\rangle - \langle\psi'|\hat{H}'|\psi'\rangle = i\hbar\langle\psi|\dot{\psi}\rangle - \langle\psi|\hat{H}|\psi\rangle = \mathcal{L}. \tag{1.6.38}$$

So, the interacting theory of the wavefunction with the electromagnetic potentials still has exactly the same gauge symmetry. Under the gauge transformation, $\psi$ is multiplied by a spacetime dependent phase. We reasoned at the end of Orig. section 3.1 that electromagnetism has **local U(1) gauge symmetry**. Here we literally see those $1 \times 1$ complex unitary matrices manifest in how the wavefunction is transformed!

In classical mechanics, $\vec{p}$ is not gauge invariant but $\vec{p} - q\vec{A}$ is (and similarly so for $\mathcal{H}$ versus $\mathcal{H} - q\phi$). In quantum mechanics, $\hat{\vec{p}} - q\hat{\vec{A}}$ likewise behaves simpler under gauge transformations than $\hat{\vec{p}}$ (and $\hat{H} - q\hat{\phi}$ simpler than $\hat{H}$). In particular, using (1.6.36), we have:

$$\hat{U}(\hat{H} - q\hat{\phi})\hat{U}^\dagger = (\hat{H} - q\hat{\phi})', \qquad \hat{U}(\hat{\vec{p}} - q\hat{\vec{A}})\hat{U}^\dagger = (\hat{\vec{p}} - q\hat{\vec{A}})'. \tag{1.6.39}$$

It immediately follows from these that $(\hat{H} + q\hat{\phi})|\psi\rangle$ and $(\hat{\vec{p}} - q\hat{\vec{A}})|\psi\rangle$ transform the same way under gauge transformation that $|\psi\rangle$ does, meaning:

$$\left((\hat{H} + q\hat{\phi})|\psi\rangle\right)' = \hat{U}(\hat{H} + q\hat{\phi})|\psi\rangle, \qquad \left((\hat{\vec{p}} + q\hat{\vec{A}})|\psi\rangle\right)' = \hat{U}(\hat{\vec{p}} + q\hat{\vec{A}})|\psi\rangle. \tag{1.6.40}$$

Thus, the operators $\hat{H} - q\hat{\phi}$ and $\hat{\vec{p}} - q\hat{\vec{A}}$ are said to be **gauge covariant**, as they preserve the gauge transformation properties of quantities they act on. $\hat{H}$ and $\hat{\vec{p}}$ on their own do not have this property. In position space, the momentum operator acts (up to a scalar multiplier) as taking the spatial derivative and in the Schrödinger equation, the Hamiltonian similarly acts proportionally to a time derivative. Thus, we define the **gauge covariant derivatives** $\mathscr{D}_t$ and $\vec{\mathscr{D}}$ by:

$$\mathscr{D}_t\psi \stackrel{\text{def}}{=} \frac{\partial\psi}{\partial t} + i\frac{q}{\hbar}\phi\psi, \qquad \vec{\mathscr{D}}\psi \stackrel{\text{def}}{=} \nabla\psi - i\frac{q}{\hbar}\vec{A}\psi \tag{1.6.41}$$

so that:

$$i\hbar\mathscr{D}_t\psi = \langle\vec{x}|(\hat{H} - q\hat{\phi})|\psi\rangle, \qquad -i\hbar\vec{\mathscr{D}}\psi = \langle\vec{x}|(\hat{\vec{p}} - q\hat{\vec{A}})|\psi\rangle. \tag{1.6.42}$$

The gauge covariance of these derivatives is expressed by the facts:

$$(\mathscr{D}_t\psi)' = e^{i\frac{q}{\hbar}f}\mathscr{D}_t\psi, \qquad (\vec{\mathscr{D}}\psi)' = e^{i\frac{q}{\hbar}f}\vec{\mathscr{D}}\psi \tag{1.6.43}$$

The Galilean Schrödinger equation (1.6.19) in terms of the gauge covariant derivatives is very simply:

$$i\hbar\mathscr{D}_t\psi = -\frac{\hbar^2}{2m}\vec{\mathscr{D}}^2\psi \tag{1.6.44}$$

where $\vec{\mathscr{D}}^2 = \vec{\mathscr{D}} \cdot \vec{\mathscr{D}}$. This is just the equation of motion for a free particle (meaning with no potentials) with all the derivatives replaced by the gauge covariant derivatives! Compare all these results with Orig. problem 6.13.

We can write our Lagrangian explicitly in the position basis as:

$$\mathcal{L} = \int_{\mathbb{S}} \left( i\hbar\psi^* \dot{\psi} - \psi^* \left( \frac{(-i\hbar\nabla - q\vec{A})^2}{2m} + q\phi \right) \psi \right) d\mathbb{S} \tag{1.6.45}$$

$$= \int_{\mathbb{S}} \left( \frac{i}{2}\hbar\psi^*\dot{\psi} + \frac{i}{2}\hbar\psi^*\dot{\psi} - \psi^* \frac{-\hbar^2\nabla^2\psi + i\hbar q\nabla\cdot(\vec{A}\psi) + i\hbar q\vec{A}\cdot\nabla\psi + q^2|\vec{A}|^2\psi}{2m} - q\psi^*\psi\phi \right) d\mathbb{S} \tag{1.6.46}$$

$$= \int_{\mathbb{S}} \left( \frac{i\hbar}{2}\psi^* \left( \dot{\psi} + i\frac{q}{\hbar}\phi\psi \right) - \frac{i\hbar}{2}\left( \dot{\psi}^* - i\frac{q}{\hbar}\phi\psi^* \right)\psi - \frac{\hbar^2 \left( \nabla + i\frac{q}{\hbar}\vec{A} \right)\psi^* \cdot \left( \nabla - i\frac{q}{\hbar}\vec{A} \right)\psi}{2m} \right) d\mathbb{S} + \frac{d}{dt}(i\hbar\langle\psi|\psi\rangle) \tag{1.6.47}$$

$$\simeq \int_{\mathbb{S}} \left( -\hbar \mathrm{Im}(\psi^* \mathscr{D}_t \psi) - \frac{\hbar^2}{2m}|\vec{\mathscr{D}}\psi|^2 \right) d\mathbb{S}. \tag{1.6.48}$$

Therefore we have:

**Theorem 1.10: Galilean Schrödinger Scalar with Gauge Theory**

The Lagrangian density for a scalar Galilean Schrödinger field experiencing electromagnetism is:

$$\mathfrak{L} = -\hbar \mathrm{Im}(\psi^* \mathscr{D}_t \psi) - \frac{\hbar^2}{2m}|\vec{\mathscr{D}}\psi|^2 \tag{1.6.49}$$

where $\mathscr{D}_t\psi = \left(\partial_t + i\frac{q}{\hbar}\phi\right)\psi$ and $\vec{\mathscr{D}}\psi = \left(\nabla - i\frac{q}{\hbar}\vec{A}\right)\psi$ are the gauge covariant derivatives. This theory has local U(1) gauge symmetry:

$$\psi' = e^{\frac{iq}{\hbar}f(t,\vec{x})}\psi \qquad \phi' = \phi - \frac{\partial f}{\partial t}, \qquad \vec{A}' = \vec{A} + \nabla f. \tag{1.6.50}$$

### 1.6.2 Gauge Transformation Geometry

Local U(1) symmetry has a fundamentally geometric interpretation. For the purpose of this subsection, it will save words and symbols to use the 4 coordinate notation $t = x^0$; no actual relativity is required here, this is just a notation choice. At each point in spacetime, $\psi(x)$ is a complex number. A complex number $a + ib$ may be viewed as a vector in the plane with coordinates $(a, b)$, identifying 1 and $i$ with the unit vectors $(1, 0)$ and $(0, 1)$. Multiplying a complex number by $e^{i\theta}$ simply rotates that vector by an angle $\theta$ in the complex plane. So, our gauge transformation can be seen as us having the freedom to rotate $\psi(x)$ independently in the complex plane at each event in spacetime by an angle $\frac{q}{\hbar}f(x)$ provided we simultaneously change $\vec{A}$ and $\phi$ in the described way. Imagine a Hilbert space $\mathbb{V}$ isomorphic to the complex plane but with a copy at spacetime event so that we have $\mathbb{V}(x)$. For each $x$, $\mathbb{V}(x)$ is then a two dimensional real vector space to which we can identify two orthonormal basis vectors $\vec{\xi}_1(x)$ and $\vec{\xi}_2(x)$:

$$\vec{\xi}_1 \cdot \vec{\xi}_1 = 1, \qquad \vec{\xi}_1 \cdot \vec{\xi}_2 = 0, \qquad \vec{\xi}_2 \cdot \vec{\xi}_2 = 1 \tag{1.6.51}$$

where these inner products are taken with both vectors at the same $x$ as for different $x$ the vectors are in different vector spaces and so their inner products are not defined. We can consider the wavefunction $\Psi(x)$ as being for each point in space and time an element of the corresponding vector space $\mathbb{V}(x)$. In that space it will have two real components $\psi^1(x)$ and $\psi^2(x)$ so that:

$$\Psi(x) = \psi^1(x)\vec{\xi}_1(x) + \psi^2(x)\vec{\xi}_2(x) = \psi^a(x)\vec{\xi}_a(x) \tag{1.6.52}$$

Here we define $\psi_1$ and $\psi_2$ to be real and related to our usual complex valued wavefunction by:

$$\psi(x) = \frac{\psi^1(x) + i\psi^2(x)}{\sqrt{2}} \tag{1.6.53}$$

with the $\sqrt{2}$ just being a nice convention. Under our U(1) transformation:

$$\psi'(x) = e^{\frac{iq}{\hbar}f(x)}\psi(x) \tag{1.6.54}$$

$$\implies \frac{\psi'^1(x) + i\psi'^2(x)}{\sqrt{2}} = \left(\cos\left(\frac{qf}{\hbar}\right) + i\sin\left(\frac{qf}{\hbar}\right)\right)\frac{\psi^1(x) + i\psi^2(x)}{\sqrt{2}} \tag{1.6.55}$$

$$\implies \psi'^1 = \psi^1\cos\left(\frac{qf}{\hbar}\right) - \psi^2\sin\left(\frac{qf}{\hbar}\right), \qquad \psi'^2 = \psi^1\sin\left(\frac{qf}{\hbar}\right) + \psi^2\cos\left(\frac{qf}{\hbar}\right). \tag{1.6.56}$$

Now you can see plainly that this is a rotation. If under the U(1) transformation we also require our basis vectors $\vec{\xi}_1$ and $\vec{\xi}_2$ to transform so that the $\Psi$ is unchanged we have:

$$\Psi' \stackrel{!}{=} \Psi \implies \psi'^1\vec{\xi}'_1 + \psi'^2\vec{\xi}'_2 = \psi^1\vec{\xi}_1 + \psi^2\vec{\xi}_2. \tag{1.6.57}$$

Using (1.6.56) and that this holds regardless of the values of $\psi'_1$ and $\psi'_2$:

$$\vec{\xi}'_1 = \cos\left(\frac{qf}{\hbar}\right)\vec{\xi}_1 - \sin\left(\frac{qf}{\hbar}\right)\vec{\xi}_2, \qquad \vec{\xi}'_2 = \sin\left(\frac{qf}{\hbar}\right)\vec{\xi}_1 + \cos\left(\frac{qf}{\hbar}\right)\vec{\xi}_2. \tag{1.6.58}$$

In this view the gauge transformation is simply a change of basis of the internal space $\mathbb{V}(x)$. We transform our basis vectors $\vec{\xi}_1$ and $\vec{\xi}_2$ and components $\psi^1$ and $\psi^2$ together in such a way that the actual geometric object they contribute to, $\Psi$, is unchanged. The ability to choose our basis vectors independently at each point in spacetime just means that there is no preferred basis of the internal space. Thus, gauge symmetry really just corresponds to local internal rotational symmetry within the internal space $\mathbb{V}$. Automatically the transformation (1.6.58) preserves our inner product conditions (1.6.51).

Now that we are thinking of the wavefunction at $x$ as an element of $\mathbb{V}(x)$, it is not meaningful to form linear combinations of $\Psi(x)$ at different values of $x$, as they live in different spaces. Thus, it is not meaningful to take naive partial derivatives of $\Psi(x)$. This is *exactly* like how in general relativity the components of a vector field at each different point in spacetime lie in a different tangent space, and we can solve this problem in exactly the same way. We need a notion of parallel transport, which will be provided by $\Upsilon(\Psi, x, x')$ so that for $x$ and $x'$ sufficiently close to each other, this map takes $\Psi \in \mathbb{V}(x)$ and produces a vector in $\mathbb{V}(x')$ and which continuously becomes the identity map as $x'$ and $x$ are taken to coincide. Just like in general relativity, in order to be compatible with the vector space structure of $\mathbb{V}(x)$ and to respect its inner product (which is what is meant by being parallel), we need:

$$\Upsilon(a\Psi_1 + b\Psi_2, x, x') = a\Upsilon(\Psi_1, x, x') + b\Upsilon(\Psi_2, x, x'), \qquad \Upsilon(\Psi_1, x, x')\cdot\Upsilon(\Psi_2, x, x') = \Psi_1\cdot\Psi_2 \tag{1.6.59}$$

for constants $a$ and $b$ and vectors $\Psi_1, \Psi_2 \in \mathbb{V}(x)$. Then, just like in Orig. (7.2.4), we define the covariant derivative of $\Psi(x)$ so that for any vector $v^\mu$:

$$v^\mu\mathscr{D}_\mu\Psi \stackrel{\text{def}}{=} \lim_{\varepsilon\to 0}\frac{\Upsilon(\Psi(x+\varepsilon v), x+\varepsilon v, x) - \Psi(x)}{\varepsilon}. \tag{1.6.60}$$

By following the steps to Orig. (7.2.8), the linearity of parallel transport written in components gives:

$$\mathscr{D}_\mu\Psi = \psi^a\mathscr{D}_\mu\vec{\xi}_a + (\partial_\mu\psi^a)\vec{\xi}_a \tag{1.6.61}$$

The quantity $v^\mu\mathscr{D}_\mu\vec{\xi}_a$ is the limit of a difference quotient of vectors in $\mathbb{V}(x)$ which all depend linearly on $v^\mu$ and so is also a vector in $\mathbb{V}(x)$ which depends linearly on $v^\mu$. That is, just like in general relativity, there exist connection coefficients ${\Gamma^b}_{a\mu}(x)$ so that:

$$\mathscr{D}_\mu\vec{\xi}_a(x) = {\Gamma^b}_{a\mu}(x)\vec{\xi}_b(x). \tag{1.6.62}$$

With this so-called **gauge connection**, the covariant derivative can be written as:

$$\mathscr{D}_\mu\Psi = (\mathscr{D}_\mu\psi^a)\vec{\xi}_a, \qquad \mathscr{D}_\mu\psi^a = \partial_\mu\psi^a + {\Gamma^a}_{b\mu}\psi^b. \tag{1.6.63}$$

Because the range of the $a$ and $b$ indices on ${\Gamma^a}_{b\mu}$ is $\{1,2\}$, the two linearly independent directions in $\mathbb{V}(x)$, at each $x$ ${\Gamma^a}_{b\mu}$ is a $2\times 2$ matrix of spacetime vectors. However, because the parallel transport map must preserve inner products, the components of that matrix are not independent of each other. In particular, the basis vectors satisfy $\vec{\xi}_a\cdot\vec{\xi}_b = \delta_{ab}$ by (1.6.51) by definition. Under parallel transport then:

$$\Upsilon(\vec{\xi}_a(x+v\varepsilon), x+v\varepsilon, x)\cdot\Upsilon(\vec{\xi}_b(x+v\varepsilon), x+v\varepsilon, x) = \delta_{ab}. \tag{1.6.64}$$

Differentiating both sides with respect to $\varepsilon$, evaluating at $\varepsilon = 0$, using that the result holds for all $v^\mu$, and applying the definition of the gauge connection:

$$\vec{\xi}_a\cdot\mathscr{D}_\mu\vec{\xi}_b + \vec{\xi}_b\cdot\mathscr{D}_\mu\vec{\xi}_a = 0 \implies \delta_{ac}{\Gamma^c}_{b\mu} + \delta_{bc}{\Gamma^c}_{a\mu} = 0. \tag{1.6.65}$$

Thus, the matrix $\delta_{ac}\Gamma^c{}_{b\mu}$ must be antisymmetric under swapping $a$ and $b$. The space of $2\times 2$ antisymmetric matrices is 1 dimensional, and so we can define the single vector field $A_\mu(x) = \frac{\hbar}{q}\Gamma^1{}_{2\mu}(x)$ from which it follows:

$$\Gamma^a{}_{b\mu} = -\frac{q}{\hbar}A_\mu \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^a{}_b. \tag{1.6.66}$$

The covariant derivatives of our basis vectors and field components become:

$$\mathscr{D}_\mu\vec{\xi}_1 = -\frac{q}{\hbar}A_\mu\vec{\xi}_2, \qquad \mathscr{D}_\mu\vec{\xi}_2 = \frac{q}{\hbar}A_\mu\vec{\xi}_1, \qquad \mathscr{D}_\mu\psi^1 = \partial_\mu\psi^1 + \frac{q}{\hbar}A_\mu\psi^2, \qquad \mathscr{D}_\mu\psi^2 = \partial_\mu\psi^2 - \frac{q}{\hbar}A_\mu\psi^1. \tag{1.6.67}$$

If we now collect our field components back into our complex representation in (1.6.53), it follows that:

$$\mathscr{D}_\mu\psi = \partial_\mu\psi - i\frac{q}{\hbar}A_\mu\psi \tag{1.6.68}$$

which is exactly the original definition of the gauge covariant derivative, confirming the geometric picture we are building is consistent with what we already know! Thus, the electromagnetic vector potential $A_\mu$ is actually geometrically the connection on the internal space on which $\Psi$ lives, the wavefunction $\psi$ is just a complex component representation of the abstract $\Psi$, and electromagnetic gauge transformations are just local changes of basis of the internal space. We can further confirm that this picture reproduces electromagnetism by looking at what a gauge transformation does to $A_\mu$. Taking the covariant derivative of the gauge transformation (1.6.58) and using (1.6.67), we have:

$$\mathscr{D}_\mu\vec{\xi}'_1 = -\frac{q}{\hbar}(A_\mu + \partial_\mu f)\vec{\xi}'_2, \qquad \mathscr{D}_\mu\vec{\xi}'_2 = \frac{q}{\hbar}(A_\mu + \partial_\mu f)\vec{\xi}'_1 \tag{1.6.69}$$

which is exactly of the same form as (1.6.67) but with the gauge transformed potential:

$$A'_\mu(x) = A_\mu(x) + \partial_\mu f(x) \tag{1.6.70}$$

as the electromagnetic potential should always have.

The combination of the collection of vector spaces $\mathbb{V}(x)$ among all $x$ and the parallel transport map $\Upsilon$ is called a **vector bundle** $\mathcal{V}$. The vector potential is then said to provide the connection on that vector bundle as it gives the geometric information that connects $\mathbb{V}(x)$ and $\mathbb{V}(x')$ for $x'$ very nearby $x$. Because we can now literally see the gauge transformation properties of the potentials as arising entirely from our ability to freely choose our basis vectors for $\mathbb{V}(x)$ at each $x$, we can further see that the outdated view that one sometimes encounters that the potentials are unphysical is as ridiculous as claiming that all vectors are unphysical because their components are basis dependent. The components of all vectors are not uniquely defined because I can choose my coordinate axes to be whatever I want! Similarly, the potentials are not uniquely defined because I can choose the axes of $\mathbb{V}(x)$ to be whatever I want. But, not uniquely defined is not the same as unphysical. While the $z$ component $v_z$ of a vector $\vec{v}$ is not uniquely defined, it does contain information that may be physical as it is necessary to compute $|\vec{v}|$ which *is* uniquely defined under changes of your coordinate axes. Similarly, while the numerical value of the potentials at a particular $x$ is not uniquely defined, the potentials as functions carry all of the physical information of electromagnetism!

You might worry that there is more to the geometry of the vector bundle $\mathcal{V}$ than what is in the connection. However, the connection determines the entire parallel transport map, just like Orig. (7.5.82) but easier because electromagnetism is Abelian. Consider a path $\gamma$ in spacetime with coordinates $x^\mu(\lambda)$ parameterized by a real parameter $\lambda$. Suppose we have a vector $\vec{v}(0)\in\mathbb{V}(x(0))$ which we want to parallel transport along the curve so that $v(\lambda)$ is the parallel transported vector in $\mathbb{V}(x(\lambda))$ for each $\lambda$. From $x(\lambda)$ we can transport to a nearby $x(\lambda+\varepsilon)$ using $\Upsilon$ to define a new vector $\vec{v}'_\varepsilon(\lambda)$:

$$\vec{v}'_\varepsilon(\lambda) = \Upsilon(\vec{v}(\lambda), x(\lambda), x(\lambda+\varepsilon)). \tag{1.6.71}$$

We can then parallel transport this back to $x(\lambda)$, which is the inverse process and so this is guaranteed to simply undo the original parallel transport:

$$\Upsilon(\vec{v}'_\varepsilon(\lambda), x(\lambda+\varepsilon), x(\lambda)) = \vec{v}(\lambda). \tag{1.6.72}$$

Since $\vec{v}(\lambda)$ represents the parallel transported vector for each $\lambda$, we have that for $\varepsilon$ sufficiently small, $\vec{v}'_\varepsilon(\lambda) = \vec{v}(\lambda+\varepsilon) + \mathcal{O}(\varepsilon^2)$. Then, in the $\varepsilon\to 0$ limit:

$$\lim_{\varepsilon\to 0}\frac{\Upsilon(\vec{v}(\lambda+\varepsilon, x(\lambda+\varepsilon), x(\lambda)) - \vec{v}(\lambda)}{\varepsilon} = 0 \implies \frac{dx^\mu}{d\lambda}\mathscr{D}_\mu\vec{v} = 0 \tag{1.6.73}$$

just like as in gravity. This differential equation can be solved to give the fully parallel transport map from the gauge covariant derivatives.

To solve (1.6.73), let $v(\lambda)$ be the complex representation of $\vec{v}(\lambda)$. Then, (1.6.73) becomes:

$$v(\lambda) = \frac{v^1(\lambda) + iv^2(\lambda)}{\sqrt{2}} \implies \frac{dv}{d\lambda} = i\frac{q}{\hbar}A_\mu(x(\lambda))\frac{dx^\mu}{d\lambda}v. \tag{1.6.74}$$

This is exactly solved by:

$$v_{\text{final}} = e^{i\theta[\gamma]}v_{\text{initial}} \tag{1.6.75}$$

where the path angle $\theta[\gamma]$ is defined by:

$$\theta[\gamma] = \frac{q}{\hbar}\int_\gamma A_\mu dx^\mu. \tag{1.6.76}$$

Thus in the complex representation $\Upsilon[\gamma] = e^{i\theta[\gamma]}$; electromagnetic parallel transport of the wavefunction just picks up a path dependent phase.

In general relativity, in Orig. (7.5.14), we derived the Riemann tensor and the notion of the local curvature of a connection as the measure of the failure of parallel transport around a closed loop to be trivial. Similarly, in electromagnetism parallel transporting a vector in $\mathcal{V}$ around a small loop will tell us about the electromagnetic curvature of $\mathcal{V}$. Parallel transporting a vector a small amount just changes it by the gauge covariant derivative along the transport path, and so going first in one direction, then in a second, backward along the first, and backward along the second to achieve a closed loop will result in the vector changing by the commutator of covariant derivatives in the two directions, just like in Orig. (1.11.25) and (7.5.13). (Because we are currently in flat spacetime we don't need the final correction to close the loop needed in Orig. (7.5.13)). So, let us compute the commutator of covariant derivatives to see what electromagnetic curvature really means:

$$[\mathscr{D}_\mu, \mathscr{D}_\nu]\psi = \left(\partial_\mu - i\frac{q}{\hbar}A_\mu\right)\left(\partial_\nu\psi - i\frac{q}{\hbar}A_\nu\psi\right) - \left(\partial_\nu - i\frac{q}{\hbar}A_\nu\right)\left(\partial_\mu\psi - i\frac{q}{\hbar}A_\mu\psi\right) \tag{1.6.77}$$

$$= -i\frac{q}{\hbar}(\partial_\mu A_\nu - \partial_\nu A_\mu)\psi = -i\frac{q}{\hbar}\mathcal{F}_{\mu\nu}\psi. \tag{1.6.78}$$

The electromagnetic field strength gives precisely the local curvature of $\mathcal{V}$! The magnetic field is the curvature of $\mathcal{V}$ in spatial directions while the electric field is the curvature in mixing space and time and directions.

We actually already organically found how important electromagnetic parallel transport is indirectly when studying classical electromagnetism in the Orig. Consider an area $\mathbb{A}$ bounded by a continuous closed loop $\gamma = \partial\mathbb{A}$ in space at fixed time. Then $\theta[\gamma]$ becomes, using Stokes' theorem:

$$\theta[\gamma] = \frac{q}{\hbar}\int_\gamma \vec{A}\cdot d\vec{x} = \frac{q}{\hbar}\int_{\mathbb{A}} \vec{B}\cdot d\vec{\mathbb{A}} \tag{1.6.79}$$

which is just the magnetic flux through the loop, up to an overall multiplier! That is, magnetic fluxes are the electromagnetic Wilson loops.

Both gravity and electromagnetism are fundamentally geometric theories, and the structures of the two are highly analogous. That analogy is made most clear when comparing to the gravitational spin connection. Vector fields $f^A\vec{\iota}_A$ in general relativity can be thought of at each point in spacetime as being valued in a copy of Minkowski space at that point. Each copy has Lorentz transformation symmetry so that the theory has local Lorentz gauge symmetry. These different copies of Minkowski space are connected by the spin connection $\omega^A{}_{B\mu}$. The curvature of that connection is the Riemann tensor $R^A{}_{B\mu\nu}$. Charged wavefunctions $\psi^a\vec{\xi}_a$ experiencing electromagnetism can be thought of at each point in spacetime as being valued in a copy of $\mathbb{C}$ at that point. Each copy has rotational symmetry so that the theory has local U(1) gauge symmetry. These different copies of $\mathbb{C}$ are connected by the electromagnetic connection $-i\frac{q}{\hbar}A_\mu$. The curvature of that connection is proportional to the electromagnetic field strength $-i\frac{q}{\hbar}\mathcal{F}_{\mu\nu}$. We could plainly generalize these structures to gauge theories of other Lie groups too, and doing so will actually discovery other forces of nature when we study the Standard Model of Particle Physics. There is only one additional structure that gravity has that gauge theories of generic Lie groups don't have, the vierbein. Uniquely, because the tangent space at any point in spacetime is itself Minkowski space, $V^A{}_\mu$ provides a way to relate the coordinate basis vectors $\vec{e}_\mu$ and the Minkowski basis vectors $\vec{\iota}_A$ of gravity's version of $\mathbb{V}(x)$. No such invertible map between the coordinate basis vectors and the basis vectors of a generic Lie algebra exists or even makes sense.

$\psi^*$ automatically has the opposite gauge transformation property to $\psi$, so that the inner product $\psi^*\psi$ is gauge invariant. That is, its equivalents of (1.6.43) and (1.6.68) are:

$$\psi'^* = e^{-\frac{iq}{\hbar}f}\psi^* \implies \mathscr{D}_\mu\psi'^* = e^{-\frac{iq}{\hbar}f}\mathscr{D}_\mu\psi^*, \qquad \mathscr{D}_\mu\psi^* = \partial_\mu\psi^* + i\frac{q}{\hbar}A_\mu\psi^*. \tag{1.6.80}$$

Importantly, note how the gauge covariant derivative operators are defined to act differently on $\psi^*$ than $\psi$, so that in either case it is compatible with the gauge transformation. Similarly, for objects which are gauge invariant we define the gauge covariant derivative to act on them just as the normal derivative, again so that it is compatible with its transformation property. With these definitions, the gauge covariant derivative automatically follows the product rule:

$$\mathscr{D}_\mu(AB) = \mathscr{D}_\mu AB + A\mathscr{D}_\mu B. \tag{1.6.81}$$

### 1.6.3 Charge Conservation

Now that the wavefunction in the Lagrangian density in theorem 1.10 is interacting with the electromagnetic potentials, the wavefunction carries an electromagnetic current. We can compute that current by seeing what terms the Schrödinger Lagrangian contributes to Maxwell's equations. By Orig. (3.1.18) the charge density and current density that appear in Maxwell's equations due to an interaction are given from the interaction terms in the action $S_{\text{int}}$ by (using (1.6.49)):

$$\rho = -\frac{\delta S_{\text{int}}}{\delta \phi} = -\frac{\partial \mathfrak{L}}{\partial \phi} = q|\psi|^2, \qquad \vec{J} = \frac{\delta S_{\text{int}}}{\delta \vec{A}^T} = \frac{\partial \mathfrak{L}}{\partial \vec{A}^T} = \frac{i\hbar q}{2m}\left(\psi \nabla \psi^* - \psi^* \nabla \psi\right) - \frac{q^2}{m}\vec{A}|\psi|^2. \tag{1.6.82}$$

In terms of the gauge covariant derivative:

$$\rho(t,\vec{x}) = q|\psi(t,\vec{x})|^2, \qquad \vec{J}(t,\vec{x}) = \frac{q\hbar}{m}\text{Im}\left(\psi^* \vec{\mathscr{D}}\psi\right). \tag{1.6.83}$$

These expressions quickly confirm that both $\rho$ and $\vec{J}$ are gauge invariant:

$$\rho' = q|\psi'|^2 = q\left|e^{i\frac{q}{\hbar}f}\psi\right|^2 = q|\psi|^2 = \rho \tag{1.6.84}$$

$$\vec{J}' = \frac{q\hbar}{m}\text{Im}\left(\psi'^* \vec{\mathscr{D}}\psi'\right) = \frac{q\hbar}{m}\text{Im}\left(e^{-i\frac{q}{\hbar}f}\psi^* e^{i\frac{q}{\hbar}f}\vec{\mathscr{D}}\psi\right) = \frac{q\hbar}{m}\text{Im}(\psi^*\vec{\mathscr{D}}\psi) = \vec{J}. \tag{1.6.85}$$

As well, $q$ is indeed the total charge of the wavefunction:

$$\text{Total Charge} = \int_{\mathbb{S}} \rho d\mathbb{S} = \int_{\mathbb{S}} q|\psi|^2 d\mathbb{S} = q\langle\psi|\psi\rangle = q \tag{1.6.86}$$

and the continuity equation is satisfied, as required by gauge invariance (using (1.6.44)):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{J} = \frac{\partial}{\partial t}\left(q\psi^*\psi\right) + \nabla \cdot \left(\frac{i\hbar q}{2m}\left(\psi(\vec{\mathscr{D}}\psi)^* - \psi^*\vec{\mathscr{D}}\psi\right)\right) \tag{1.6.87}$$

$$= \frac{q}{i\hbar}\psi^*\left(i\hbar\mathscr{D}_t\psi + \frac{\hbar^2}{2m}\mathscr{D}^2\psi\right) - \frac{q}{i\hbar}\left(-i\hbar\mathscr{D}_t\psi^* + \frac{\hbar^2}{2m}\mathscr{D}^2\psi^*\right)\psi = 0. \tag{1.6.88}$$

While Maxwell's equations are linear in the potentials at fixed electromagnetic current and the Schrödinger equation is linear in $\psi$ at fixed potentials, the coupled system of $\psi$ to Maxwell's equations is absolutely not linear because $\rho$, $\vec{J}$, and the covariant derivatives of $\psi$ contain quadratic and cubic terms in the full set of fields $(\psi, \psi^*, \phi, \vec{A})$. The interacting theory is nonlinear! The Lagrangian density in (1.6.49) contains cubic and even quartic terms in the fields ($\phi\psi^*\psi$ is cubic while $|\vec{A}|^2\psi^*\psi$ is quartic).

Using $\rho = q|\phi|^2$, we can find a quite nice formula for $\psi$'s electric dipole moment (recalling Orig. problem 3.68 for its definition):

$$\vec{\wp} = \int_{\mathbb{S}} \rho\vec{x}d\mathbb{S} = \int_{\mathbb{S}} q|\psi|^2\vec{x}d^3\vec{x} = q\langle\psi|\hat{\vec{x}}|\psi\rangle = q\langle\hat{\vec{x}}\rangle. \tag{1.6.89}$$

As well, we can find a nice formula for the magnetic dipole moment (defined in Orig. problem 3.77) with a little integration by parts:

$$\vec{\mu} = \frac{1}{2}\int_{\mathbb{S}} \vec{x}\times\vec{J}d\mathbb{S} = \frac{q\hbar}{2m}\epsilon^a{}_b{}^c\vec{\iota}_a\text{Im}\int_{\mathbb{S}} x^b\psi^*\mathscr{D}_c\psi d\mathbb{S} \tag{1.6.90}$$

$$= \frac{q\hbar}{4im}\epsilon^a{}_b{}^c\vec{\iota}_a\int_{\mathbb{S}} x^b\left(\psi^*\partial_c\psi - i\frac{q}{\hbar}\psi^*A_c\psi - \partial_c\psi^*\psi - i\frac{q}{\hbar}A_c\psi^*\psi\right)d\mathbb{S} \tag{1.6.91}$$

$$= \frac{q\hbar}{4im}\epsilon^a{}_b{}^c\vec{\iota}_a\int_{\mathbb{S}}\left(x^b\left(\psi^*\partial_c\psi - 2i\frac{q}{\hbar}\psi^*A_c\psi + \psi^*\partial_c\psi\right) + \delta^b_c\psi^*\psi\right)d\mathbb{S} \tag{1.6.92}$$

$$= \frac{q\hbar}{2im}\epsilon^a{}_b{}^c\vec{\iota}_a\int_{\mathbb{S}}\psi^*x^b\left(\partial_c\psi - i\frac{q}{\hbar}A_c\psi\right)d\mathbb{S} = \frac{q\hbar}{2im}\langle\psi|\hat{\vec{x}}\times\hat{\vec{\mathscr{D}}}|\psi\rangle = \frac{q}{2m}\langle\hat{\vec{x}}\times(-i\hbar\hat{\vec{\mathscr{D}}})\rangle \tag{1.6.93}$$

where $\hat{\vec{\mathscr{D}}}$ is defined so that $-i\hbar\hat{\vec{\mathscr{D}}} = \hat{\vec{p}} - q\hat{\vec{A}}$.

The operator $\hat{\vec{x}} \times (-i\hbar\hat{\vec{\mathscr{D}}})$ has a definite physical interpretation in terms of the angular momentum of $\psi$. Because $\psi(t, \vec{x})$ is a scalar field, under a rotation by a rotation matrix $\hat{R}$, $\psi'(t', \vec{x}') = \psi(t, \vec{x})$, thus:

$$\psi'(t, \vec{x}) = \psi(t, \hat{R}^{-1}\vec{x}). \tag{1.6.94}$$

Differentiating both sides with respect to the rotation angle $\theta^a$ near $\theta = 0$:

$$\left.\frac{d\psi}{d\theta^a}\right|_{\theta=0} = -\epsilon_{ab}{}^c \frac{\partial\psi}{\partial x^c} \implies \left.\frac{d|\psi\rangle}{d\theta^a}\right|_{\theta=0} = -\frac{i}{\hbar}\epsilon_{abc}\hat{x}^b\hat{p}^c|\psi\rangle. \tag{1.6.95}$$

The generator of rotations (and thus the angular momentum) of just the wavefunction is given by Noether's theorem (or its special case, theorem 1.4):

$$\mathrm{L}_a = \langle\pi|\left.\frac{d|\psi\rangle}{d\theta^a}\right|_{\theta=0} = \langle\psi|\epsilon_{abc}\hat{x}^b\hat{p}^c|\psi\rangle \implies \vec{\mathrm{L}} = \langle\psi|\hat{\vec{L}}|\psi\rangle = \langle\psi|(\hat{\vec{x}} \times \hat{\vec{p}})|\psi\rangle \implies \hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}}. \tag{1.6.96}$$

This $\vec{\mathrm{L}}$ is only the free angular momentum of the wavefunction, neglecting its interactions with other fields, which makes sense because in the Noether's theorem expression we used above we included no other fields. This $\vec{\mathrm{L}}$ is also specifically orbital angular momentum because it all comes from changing the evaluation point $\vec{x}$ and not from intrinsic changes in the direction of $\psi$ (because $\psi$ is a scalar under rotations, its direction does not change at all, i.e. it has no spin angular momentum).

To derive the total angular momentum of the interacting theory we have to use the action for the fully interacting theory (and thus include the Maxwell terms in our Lagrangian density in theorem 1.10). Then, Noether's theorem gives the total angular momentum:

$$\mathrm{J}_a = \langle\pi|\left.\frac{d|\psi\rangle}{d\theta^a}\right|_{\theta=0} + \int_{\mathbb{S}} \vec{P}_A \cdot \left.\frac{d\vec{A}}{d\theta^a}\right|_{\theta=0} d\mathbb{S} \tag{1.6.97}$$

where $\vec{P}_A = -\vec{E}$ is the conjugate momentum field to the vector potential. There is no surface term because the Lagrangian density is invariant under rotations. Recycling the results of Orig. problem 4.31, under a rotation the vector potential transforms as:

$$\vec{A}'(t, \vec{x}) = \hat{R}\vec{A}(t, \hat{R}^{-1}\vec{x}) \implies \left.\frac{d\vec{A}}{d\theta^a}\right|_{\theta=0} = \epsilon_a{}^{bc}A_b\vec{\iota_c} - \epsilon_{ab}{}^c x^b \frac{\partial\vec{A}}{\partial x^c}. \tag{1.6.98}$$

Then, following standard steps of calculation:

$$\mathrm{J}_a = \langle(\hat{\vec{x}} \times \hat{\vec{p}})_a\rangle + \int_{\mathbb{S}} (-\vec{E}) \cdot \left(\epsilon_a{}^{bc}A_b\vec{\iota_c} - \epsilon_{ab}{}^c x^b \frac{\partial\vec{A}}{\partial x^c}\right) d\mathbb{S} \tag{1.6.99}$$

$$= \langle(\hat{\vec{x}} \times \hat{\vec{p}})_a\rangle + \int_{\mathbb{S}} \left((\vec{E} \times \vec{A})_a + \epsilon_{ab}{}^c x^b \vec{E} \cdot \frac{\partial\vec{A}}{\partial x^c}\right) d\mathbb{S} + \int_{\mathbb{S}} \nabla \cdot \vec{E}\epsilon_{ab}{}^c x^b A_c d\mathbb{S} - \int_{\mathbb{S}} q|\psi|^2(\vec{x} \times \vec{A})_a d\mathbb{S} \tag{1.6.100}$$

$$= \langle(\hat{\vec{x}} \times (\hat{\vec{p}} - q\hat{\vec{A}}))_a\rangle + \int_{\mathbb{S}} \left((\vec{E} \times \vec{A})_a + \epsilon_{ab}{}^c x^b \vec{E} \cdot \frac{\partial\vec{A}}{\partial x^c} + \epsilon_{ab}{}^c x^b A_c \frac{\partial E^d}{\partial x^d}\right) d\mathbb{S} \tag{1.6.101}$$

$$= \langle(\hat{\vec{x}} \times (-i\hbar\hat{\vec{\mathscr{D}}}))_a\rangle + \int_{\mathbb{S}} \epsilon_{ab}{}^c x^b E^d \left(\frac{\partial A_d}{\partial x^c} - \frac{\partial A_c}{\partial x^d}\right) d\mathbb{S}. \tag{1.6.102}$$

So, we see that the total angular momentum is given as:

$$\vec{\mathrm{J}} = \langle(\hat{\vec{x}} \times (-i\hbar\hat{\vec{\mathscr{D}}})\rangle + \int_{\mathbb{S}} \vec{x} \times (\vec{E} \times \vec{B}) d\mathbb{S}. \tag{1.6.103}$$

The second integral is exactly the free electromagnetic angular momentum from Orig. problem 4.31 and the first term is the free orbital angular momentum from (1.6.96) except with $\hat{\vec{p}}$ replaced by the gauge covariant combination $\hat{\vec{p}} - q\hat{\vec{A}}$. Thus, the total orbital angular momentum of the wavefunction, including its contribution due to interacting with electromagnetism as:

$$\vec{\mathrm{L}} = \langle\hat{\vec{x}} \times (-i\hbar\hat{\vec{\mathscr{D}}})\rangle \tag{1.6.104}$$

with the previous formula (1.6.96) holding in the case of no vector potential. With (1.6.93), the magnetic

dipole moment of the wavefunction is given by:

$$\vec{\mu} = \frac{q\vec{\mathrm{L}}}{2m}. \tag{1.6.105}$$

## Exercises

**Problem 1.25: Free Quantum Wave.** Consider a free Galilean Schrödinger wave (meaning $U = 0$ and $\vec{K} = 0$) in $d$ spatial dimensions.

(a) Show that the free Schrödinger equation becomes:

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi.$$

(b) Fourier transform the free Schrödinger equation and solve it in Fourier space. Then, Fourier transform it back to position space to show that it has general solution given by:

$$\psi(t,\vec{x}) = \int_{\mathbb{S}} \Big(\frac{m}{2\pi i\hbar t}\Big)^{\frac{d}{2}} e^{\frac{im}{2\hbar t}|\vec{x}-\vec{x}\,'|^2}\psi(0,\vec{x}\,')d^d\vec{x}\,'.$$

(c) Let $\vec{x}_{\mathrm{cl}} = \langle\hat{\vec{x}}\rangle$, $\vec{p}_{\mathrm{cl}} = \langle\hat{\vec{p}}\rangle$. Show that they exactly satisfy:

$$\frac{d\vec{x}_{\mathrm{cl}}}{dt} = \frac{\vec{p}_{\mathrm{cl}}}{m}, \qquad \frac{d\vec{p}_{\mathrm{cl}}}{dt} = 0 \quad \Longrightarrow \quad \vec{x}_{\mathrm{cl}}(t) = \vec{x}_{\mathrm{cl}}(0) + \frac{\vec{p}_{\mathrm{cl}}(0)}{m}t, \qquad \vec{p}_{\mathrm{cl}}(t) = \vec{p}_{\mathrm{cl}}(0).$$

Thus, the center of mass of a free quantum wave packet moves exactly classically.

(d) Define:

$$X^{ab} = \mathrm{Cor}(\hat{x}^a,\hat{x}^b), \qquad P_{ab} = \mathrm{Cor}(\hat{p}_a,\hat{p}_b), \qquad C^a{}_b = \mathrm{Cor}(\hat{x}^a,\hat{p}_b).$$

Show that these exactly satisfy the equations of motion:

$$\frac{dX^{ab}}{dt} = \frac{C^{ab}+C^{ba}}{m}, \qquad \frac{dC^{ab}}{dt} = \frac{P^{ab}}{m}, \qquad \frac{dP^{ab}}{dt} = 0.$$

Finally then show that these are solved by:

$$X^{ab}(t) = X^{ab}(0) + \frac{C^{ab}(0)+C^{ba}(0)}{m}t + \frac{P^{ab}(0)}{m^2}t^2, \quad C^{ab}(t) = C^{ab}(0) + \frac{P^{ab}(0)}{m}t, \quad P_{ab}(t) = P_{ab}(0).$$

(e) Consider a single spatial dimension ($d = 1$) for simplicity and an initial wavefunction which is a coherent state of position space width $\ell$. Show that the position space spread at later times is:

$$\sigma_x(t) = \ell\sqrt{1+\left(\frac{\hbar t}{2m\ell^2}\right)^2}.$$

Thus show that after a long time the wave spread grows linearly with $t$.

**Problem 1.26: ⋆ Double-Slit Experiment Revisited.** Consider the setup of the a slit experiment with an otherwise free Galilean quantum wave passing through two slits on its way to a screen.

(a) First consider a single slit at the origin. Away from the origin $\psi(t,\vec{x})$ will satisfy the free Schrödinger equation. Thus, the problem has spherical symmetry about this origin and so we model $\psi(t,\vec{x})$ as being a function of $t, r$ only. Let $\phi(t,r) = r\psi$ and show that it satisfies:

$$\frac{\partial\phi}{\partial t} = \frac{i\hbar}{2m}\frac{\partial^2\phi}{\partial r^2}.$$

Now assume the incoming wave to the source has angular frequency $\omega$ and that the momentum of the outgoing solution should be directed outward from the origin (the wave is leaving the slit). Show that for some complex number $\xi$:

$$\psi(t,\vec{x}) = \frac{\xi}{r}e^{i\sqrt{\frac{2m\omega}{\hbar}}r - i\omega t}.$$

This is the simplest model of the outgoing wave from the slit, called a spherical wave.

(b) Now consider two slits, one placed at $\vec{x}_1 = (0,\frac{s}{2},0)$ and the other at $\vec{x}_2 = (0,-\frac{s}{2},0)$, with the screen at $z = d$. Show that the combined solution is:

$$\psi(t,\vec{x}) = \frac{\xi_1 e^{ik|\vec{x}-\vec{x}_1|-i\omega t}}{|\vec{x}-\vec{x}_1|} + \frac{\xi_2 e^{ik|\vec{x}-\vec{x}_2|-i\omega t}}{|\vec{x}-\vec{x}_2|}$$

where $k = \sqrt{\frac{2m\omega}{\hbar}}$. Show further that on the screen:

$$\psi(t,0,y,d) = e^{ik\left(d+\frac{s^2}{8d}+\frac{y^2}{2d}\right)-i\omega t} d\left(\xi_1 e^{-\frac{iksy}{2d}} + \xi_2 e^{\frac{iksy}{2d}}\right) + \mathcal{O}\left(\frac{1}{d^2}\right).$$

(c) The screen is able to measure impinging charges and so is sensitive to $|\psi|^2$ (as $\rho = q|\psi|^2$). Let $\mathcal{I} = |\psi|^2$ define the intensity of the incident wave and show that on the screen:

$$\mathcal{I} = \frac{1}{d^2}\left(|\xi_1|^2 + |\xi_2|^2 + 2\text{Re}\left(\xi_1^* \xi_2 e^{\frac{iksy}{d}}\right)\right) + \text{H.O.T.}$$

where the higher order terms fall off like $\frac{1}{d^3}$ or faster.

(d) Assume that the slits are equally spaced from the incoming wave source so that the components of the wave which reach each slit have equal magnitude and are in phase. Explain that $\xi_1 = \xi_2$ and let $A = |\xi_1|$. Show that the intensity pattern on the screen becomes:

$$\mathcal{I} = \frac{4A^2}{d^2}\cos^2\left(\frac{ksy}{2d}\right) + \text{H.O.T.}$$

This is precisely the same intensity pattern we found in Orig. section 4.6.4!

(e) Now assume that there is a detector at one of the slits (choose the second one for this part). Such a detector interacts with the incoming wave as it passes through, adding a local perturbation to the Hamiltonian operator. Argue using what we have found about adiabatic invariance in quantum mechanics that if this local perturbation is brief then it will cause the wave passing through the detector slit to maintain its magnitude but pick up a phase which depends very sensitively on the details of the detector. Thus, we can model the presence of the detector by setting $\xi_2$ and $\xi_1$ to be related by a randomly oscillating phase:

$$\xi_2 = e^{i\varphi(t)}\xi_1.$$

Show that now the time averaged resulting intensity pattern is:

$$\mathbb{T}(\mathcal{I}) = \frac{1}{d^2}\left(|\xi_1|^2 + |\xi_2|^2\right) + \text{H.O.T.}$$

Explain how this is the intensity pattern predicted from an incoming beam of classical particles. Thus, classical behavior can also often be recovered when there is **decoherence** between the phases of different components in a quantum system. (Decoherence here is modeling the detector interaction as introducing a rapidly changing phase, essentially random due to the unmodeled detector degrees of freedom.)

**Problem 1.27: Wave in an impenetrable box.** Consider a Galilean Schrödinger wave in one spatial dimension which is in an impenetrable box of side length $\ell$, meaning $\vec{K} = 0$ and $U = 0$ inside the box and $U \to \infty$ outside the box.

(a) Show that all finite energy eigenstates of the Hamiltonian operator must satisfy:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} = E\psi, \qquad \psi(0) = \psi(\ell) = 0$$

where $E$ is the energy eigenvalue corresponding to $\psi$. (Outside of the interval $0 \le x \le \ell$ the box enforces $\psi(t,x) = 0$, so all formulas in this question are assumed only to hold on the inside of the box.)

(b) Solve the eigenvector problem to show that the energy eigenstates are labeled by $n \in \{1,2,3,...\}$ and that once normalized are given by:

$$\psi_n(x) = \sqrt{\frac{2}{\ell}}\sin\left(\frac{n\pi x}{\ell}\right), \qquad E_n = \frac{n^2\pi^2\hbar^2}{2m\ell^2}.$$

(c) Show that the solution to the Schrödinger equation for arbitrary initial conditions is given by:

$$\psi(t,x) = \sum_{n=1}^{\infty} A_n e^{-\frac{in^2\pi^2\hbar t}{2m\ell^2}}\sqrt{\frac{2}{\ell}}\sin\left(\frac{n\pi x}{\ell}\right) \quad \text{where:} \quad A_n = \int_0^\ell \sqrt{\frac{2}{\ell}}\sin\left(\frac{n\pi x}{\ell}\right)\psi(0,x)dx.$$

(d) Notice that the eigenvalues of the Hamiltonian operator for this system are discrete. It is a very common error to say "the energy of the particle in a box is discretized" in quantum mechanics. This error is made by those who do not understand the difference between the energy (the conserved quantity due to time translation symmetry, given by $\text{E} = \mathcal{H} = \langle\psi|\hat{H}|\psi\rangle$) and an eigenvalue of $\hat{H}$. Show that the interpretation of the eigenvalues of $\hat{H}$ is that $\frac{E_n}{\hbar}$ gives the angular frequency of eigenstate $n$ and that

the actual energy E of the wave is in general not an eigenvalue of $\hat{H}$ and can take any continuous value greater than or equal to $E_1$.

**Problem 1.28: Wave in a penetrable box.** Consider a Galilean Schrödinger wave in one spatial dimension which is in an penetrable box of side length $\ell$ ($\vec{K} = 0$ and $U = 0$ inside the box and $U = \varepsilon$ outside).

(a) Show that the eigenstates of the Hamiltonian operator must satisfy:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} = E\psi \qquad (0 < x < \ell), \qquad\qquad -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} = (E-\varepsilon)\psi \qquad (x < 0 \text{ or } x > \ell)$$

where $E$ is the energy eigenvalue corresponding to $\psi$.

(b) Consider eigenstates of the Hamiltonian operator which decay to 0 far away from the box. Show that for some constants $\kappa > 0$, $A$, $B$, $C$, $D$:

$$\psi(x) = \begin{cases} Ae^{\kappa x}, & x < 0 \\ Be^{ix\sqrt{\frac{2m\varepsilon}{\hbar^2}-\kappa^2}} + Ce^{-ix\sqrt{\frac{2m\varepsilon}{\hbar^2}-\kappa^2}}, & 0 < x < \ell \\ De^{-\kappa x}, & x > \ell. \end{cases}$$

(c) Show using the boundary conditions at the box walls that:

$$B = \left(1 - \frac{i\kappa}{k}\right)\frac{A}{2}, \qquad C = \left(1 + \frac{i\kappa}{k}\right)\frac{A}{2}, \qquad D = e^{\kappa\ell}\left(\cos k\ell + \frac{k}{\kappa}\sin k\ell\right)A$$

where:

$$k = \sqrt{\frac{2m\varepsilon}{\hbar^2} - \kappa^2}.$$

Don't waste your time normalizing the solution. Further show that due to the boundary conditions the possible values of $k$ are the discrete values satisfying:

$$\tan k\ell = \frac{k\sqrt{\frac{2m\varepsilon}{\hbar^2} - k^2}}{k^2 - \frac{m\varepsilon}{\hbar^2}}.$$

Some people are very surprised when they learn quantum mechanics to see that the wave in the box is not fully contained in the box ($\psi$ is not zero outside). There is no reason to be surprised; that is how waves work and the exact same thing happens in electromagnetism. They are confused because they are imagining a point-particle, but no such point-particle exists here to be found. Tunneling isn't magic.

(d) Consider eigenstates of the Hamiltonian operator which do not decay to 0 far away from the box. Show that they take the form:

$$\psi(x) = \begin{cases} Ae^{ikx} + rAe^{-ikx}, & x < 0 \\ \frac{\sqrt{k^2+\delta^2}+k}{2\sqrt{k^2+\delta^2}}tAe^{i(k-\sqrt{k^2+\delta^2})\ell}e^{i\sqrt{k^2+\delta^2}x} + \frac{\sqrt{k^2+\delta^2}-k}{2\sqrt{k^2+\delta^2}}tAe^{i(k+\sqrt{k^2+\delta^2})\ell}e^{-i\sqrt{k^2+\delta^2}x}, & 0 < x < \ell \\ tAe^{ikx}, & x > \ell \end{cases}$$

where $k$ is a real parameter, $A$ is a coefficient, and:

$$\delta = \sqrt{\frac{2m\varepsilon}{\hbar^2}}, \qquad r = \frac{i\delta^2\tan(\ell\sqrt{k^2+\delta^2})}{2k\sqrt{k^2+\delta^2} - i(2k^2+\delta^2)\tan(\ell\sqrt{k^2+\delta^2})},$$

$$t = \frac{k\sqrt{k^2+\delta^2}e^{-ik\ell}}{k\sqrt{k^2+\delta^2}\cos(\ell\sqrt{k^2+\delta^2}) - i\left(k^2+\frac{\delta^2}{2}\right)\sin(\ell\sqrt{k^2+\delta^2})}.$$

(e) Recall the definitions of reflection and transmission coefficients at an interface in electromagnetism. Show that $r$ and $t$ given above correspond to the reflection and transmission coefficients of this system when a wavepacket is sent toward from the box from far away to the left (for $k > 0$). Show that in this case these coefficients satisfy:

$$|r|^2 + |t|^2 = 1.$$

Later we will prove that this must hold very generally, not simply for this system.

**Problem 1.29: Bloch's Theorem.** Spatially periodic Hamiltonian operators occur in many applications (for example a crystalline material will produce a periodic potential). Let $V(x)$ be analytic and decay to 0

sufficiently quickly for large $|x|$. Consider, for a constant $\ell$, in 1 spatial dimension:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \sum_{n\in\mathbb{Z}} V(\hat{x} - n\ell\mathbb{1})$$

(a) Define the lattice translation operator:

$$\hat{T} = e^{-\frac{i}{\hbar}\ell\hat{p}}.$$

Show that $\hat{T}$ is unitary and that this system has symmetry under the transformation $|\psi'\rangle = \hat{T}|\psi\rangle$.

(b) Show that a general eigenstate of $\hat{T}$ takes the form:

$$|\psi\rangle = \sum_{n\in\mathbb{Z}} c_n |p_n\rangle$$

for some coefficients $c_n$ and where $|p_n\rangle$ are momentum eigenstates with momentum eigenvalue:

$$p_n = \hbar k + \frac{2\pi\hbar n}{\ell}$$

and $0 \leq k < \frac{2\pi}{\ell}$. Show that $\hat{T}$ and $\hat{H}$ are simultaneously diagonalizable and use this to explain that there exists a choice of the $c_n$ so that $|\psi\rangle$ is also an eigenstate of $\hat{H}$.

(c) Prove Bloch's theorem: each eigenfunction of a spatially periodic Hamiltonian operator can be written, for $0 \leq k < \frac{2\pi}{\ell}$ and a $u(x)$ with the same spatial period as the Hamiltonian operator, as:

$$\psi(x) = e^{-ikx}u(x).$$

**Problem 1.30: Amplitude and Phase Variables.** Define the real valued fields $n(t,\vec{x})$ and $S(t,\vec{x})$ by:

$$\psi(t,\vec{x}) = \sqrt{n(t,\vec{x})}e^{\frac{i}{\hbar}S(t,\vec{x})}.$$

(a) Show that the Lagrangian density in theorem 1.10 can be written in these coordinates as:

$$\mathfrak{L} = -n\left(\dot{S} + q\phi + \frac{|\nabla S - q\vec{A}|^2}{2m}\right) - \frac{\hbar^2|\nabla n|^2}{8mn}.$$

Show that this Lagrangian density produces the equations of motion:

$$\frac{\partial n}{\partial t} + \nabla\cdot\left(n\frac{\nabla S - q\vec{A}}{m}\right) = 0, \qquad \frac{\partial S}{\partial t} + q\phi + \frac{|\nabla S - q\vec{A}|^2}{2m} - \frac{\hbar^2\nabla^2\sqrt{n}}{2m\sqrt{n}} = 0.$$

Notice that the first is a continuity equation with velocity field $\vec{v} = \frac{\nabla S - q\vec{A}}{m}$ and the second is almost the Hamilton-Jacobi equation for a classical charge particle's canonical action $S$.

(b) Assume that $\frac{\hbar\nabla n}{n}$ is everywhere small compared to $\nabla S - q\vec{A}$. Explain that this means that we assume $|\psi|^2$ is approximately steady on the length scale of the gauge invariant fluctuations of the phase of $\psi$. Show that under these circumstances the above equations of motion have a smooth $\hbar \to 0$ limit and reduce exactly to the results of Orig. problem 6.13. Because $\hbar \to 0$ formally represents the classical limit, it is common the consider a series solution for $n$ and $S$ in powers of $\hbar$ with the lowest order solutions being these classical results. This is known as the Wentzel–Kramers–Brillouin (WKB) approximation.

**Problem 1.31: ⋆ Aharanov-Bohm Effect.** Consider a double-slit experiment with the same setup as in problem 1.26 except that now between the two slits is a closed impenetrable solenoid of cross sectional area $\mathbb{A}$ with a uniform magnetic field of magnitude $B$ pointing in the $x$ direction inside it.

(a) Show that it is possible to choose the gauge of $\vec{A}$ so that the part of the wave that passes through the top slit is always in a region of $\vec{A} = 0$. Show that it is possible to choose instead the gauge so that the corresponding result is true for the bottom slit. Show that it is not possible to pick the gauge so that this is true for both of them.

(b) Show, using (1.6.79), that a gauge invariant phase difference is accumulated between the parts of the wave from each slit, given by:

$$\theta = -\frac{qB\mathbb{A}}{\hbar}.$$

Show under the assumptions of problem 1.26(d) that this produces the new interference pattern:

$$\mathcal{I} = \frac{4A^2}{d^2}\cos^2\left(\frac{ksy}{2d} - \frac{qB\mathbb{A}}{2\hbar}\right) + \mathcal{O}\left(\frac{1}{d^3}\right).$$

The spatial shift in the interference pattern detected on the screen is known as the Aharanov-Bohm effect. We see that even though the wavefunction is only nonzero in regions of space with $\vec{E} = 0$ and $\vec{B} = 0$ it is still affected by electromagnetism. This is to be expected because the Schrödinger equation depends directly on $\vec{A}$ and $\phi$, not only on $\vec{E}$ or $\vec{B}$. The well-established observation of this effect serves as experimental proof that the potentials are physical, as here they are interacted with directly[9].

**Problem 1.32: ⋆ Generator of Gauge Transformations.** Show using Orig. (4.4.14) that the generator of gauge transformations for the Schrödinger equation is:

$$G = -\int f(t,\vec{x})(q|\psi|^2 + \nabla\cdot\vec{P}_A)d^3\vec{x}.$$

Further, show that the resulting gauge transformation property produced by this generator is:

$$\left.\frac{d\phi}{d\varepsilon}\right|_{\varepsilon=0} = -\frac{\partial f}{\partial t}, \qquad \left.\frac{d\vec{A}}{d\varepsilon}\right|_{\varepsilon=0} = \{\vec{A}, G\} = \nabla f, \qquad \left.\frac{d\vec{P}_A}{d\varepsilon}\right|_{\varepsilon=0} = \{\vec{P}_A, G\} = 0, \qquad \left.\frac{d\psi}{d\varepsilon}\right|_{\varepsilon=0} = \{\psi, G\} = \frac{iq}{\hbar}f\psi.$$

Show that when integrated these produce the correct gauge transformation property for each of these quantities.

**Problem 1.33: ⋆ Quantum Harmonic Oscillator (QHOI) - Creation and Annihilation Operators.** Consider a Galilean Schrödinger wave in 1 spatial dimension subject to the harmonic oscillator potential so that, for positive constants $m$ and $\omega$:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2.$$

(a) We found in (1.2.3) a useful coordinate transformation for a harmonic oscillator. Taking that transformation as inspiration, define the "annihilation operator":

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + \frac{i}{\sqrt{2m\omega\hbar}}\hat{p}.$$

Show that $\hat{a}$ is not self-adjoint and is not a normal operator. In particular, show that:

$$[\hat{a}, \hat{a}^\dagger] = 1.$$

Therefore, by the spectral theorem $\hat{a}$ cannot be diagonalized by an orthonormal basis of eigenvectors.

(b) Show that $\hat{H}$ can be expressed in terms of $\hat{a}$ by:

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\mathbb{1}\right).$$

Define the number operator $\hat{N}$ by:

$$\hat{N} = \hat{a}^\dagger\hat{a}.$$

Show that $\hat{N}$ is self-adjoint.

(c) Let $|\psi\rangle$ be an eigenstate of $\hat{N}$ with eigenvalue $\lambda$. Show that $|\psi\rangle$ is an eigenstate of $\hat{H}$ with eigenvalue $\hbar\omega\left(\lambda + \frac{1}{2}\right)$. Further, show that $\hat{a}|\psi\rangle$ is an eigenstate of $\hat{N}$ with eigenvalue $\lambda - 1$ and that $\hat{a}^\dagger|\psi\rangle$ is an eigenstate of $\hat{N}$ with eigenvalue $\lambda + 1$. Because $\hat{a}$ lowers the eigenvalue of $\hat{N}$ it is called the annihilation operator (or lowering operator) and because $\hat{a}^\dagger$ raises the eigenvalue of $\hat{N}$ it is called the creation operator (of raising operator). As a pair, $\hat{a}^\dagger$ and $\hat{a}$ are sometimes called ladder operators.

(d) Let $|\psi\rangle$ be a normalized eigenstate of $\hat{N}$ with eigenvalue $\lambda$. By considering $\|\hat{a}|\psi\rangle\|$, show that $\lambda \geq 0$.

(e) Use the results of the previous two parts to show that there exists a normalized eigenstate $|\phi\rangle$ of $\hat{N}$ with eigenvalue $\lambda \geq 0$ such that $\hat{a}|\phi\rangle = 0$. Show further that this $|\phi\rangle$ has eigenvalue $\lambda = 0$.

(f) Show that the $|\phi\rangle$ described in the previous part in position space satisfies the differential equation:

$$\frac{\partial\phi}{\partial x} + \frac{m\omega x}{\hbar}\phi = 0.$$

[9] Historically there was a lot of confusion over this result and it was suggested that instead this was proof that quantum mechanics is non-local. The thinking here starts from a priori assuming that $\psi, \vec{A}, \phi$ are unphysical but $|\psi|^2, \vec{E}, \vec{B}$ are physical because they are gauge invariant. There is no local equation of motion equivalent to the Schrödinger equation that can couple $|\psi|^2, \vec{E}, \vec{B}$ and so since they experimentally are coupled, this must mean quantum mechanics is not local. This thinking is ridiculous because the Schrödinger equation is entirely local as a field theory. This silliness stemmed from a prejudice against the potentials which has since been rendered incompatible with a basic understanding of field theory. In the modern day there are many experiments in which the potentials are directly essential. Grow up.

Solve this equation for $\phi$ and normalize it. Show that:

$$\phi(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} e^{-\frac{m\omega x^2}{2\hbar}}.$$

Call this state $|0\rangle = |\phi\rangle$ because its $\hat{N}$ eigenvalue is 0.

(g) Show that all of the eigenstates of $\hat{N}$ are given by:

$$|n\rangle = C_n \left(\hat{a}^\dagger\right)^n |0\rangle$$

for some coefficients $C_n$ and that $\hat{N}$ has only nonrepeated eigenvalues given by $\{0, 1, 2, 3, 4, ...\}$. Show that once normalized:

$$|n\rangle = \frac{\left(\hat{a}^\dagger\right)^n}{\sqrt{n!}}|0\rangle, \qquad \hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \qquad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$$

and that these $|n\rangle$ are an orthonormal and complete basis for the Hilbert space. The value $n$ is called the excitation number or occupation number of the state $|n\rangle$.

(h) Show that in position space:

$$\psi_n(x) = \langle x|n\rangle = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) e^{-\frac{m\omega x^2}{2\hbar}}$$

where $H_n(z)$ is the $n^{\text{th}}$ Hermite polynomial defined by:

$$H_n(z) = (-1)^n e^{z^2} \left(\frac{\partial}{\partial z}\right)^n e^{-z^2}.$$

Show that $H_n(z)$ is a polynomial in $z$ of degree $n$. Plot $\psi_n(x)$ as a function of $x$ for several different $n$. Notice that $\psi_n(x)$ always decays for large $|x|$ and has $n+1$ bumps.

(i) Show that the $|n\rangle$ are the eigenstates of $\hat{H}$ and have associated eigenvalues:

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right).$$

Use the familiar fact that the Hamiltonian may be shifted by a constant with no physical consequences and the interpretation of $\frac{E_n}{\hbar}$ as the angular frequencies of the system to show that the quantum harmonic oscillator (QHO) can be thought of as having angular frequencies coming in integer multiples of the classical frequency $\omega$. Use problem 1.3 to show that the QHO's time evolution is given by:

$$|\psi(t)\rangle = \sum_{n=0}^{\infty} e^{-i(n+\frac{1}{2})\omega t}|n\rangle\langle n|\psi(0)\rangle.$$

(j) Let $\ell = \sqrt{\frac{\hbar}{2m\omega}}$ and recall problems 1.21 and 1.24. Show that coherent states may be written:

$$|\mathfrak{q}, \mathfrak{p}\rangle = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}|0\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle \quad \text{where:} \quad \alpha = \sqrt{\frac{m\omega}{2\hbar}}\mathfrak{q} + \frac{i}{\sqrt{2m\omega\hbar}}\mathfrak{p}.$$

As well, show that coherent states are eigenstates of $\hat{a}$ with eigenvalue $\alpha$:

$$\hat{a}|\mathfrak{q}, \mathfrak{p}\rangle = \alpha|\mathfrak{q}, \mathfrak{p}\rangle.$$

**Problem 1.34: ⋆ QHOII - Moment Evolution Equations.** Consider again the quantum harmonic oscillator. Define the first and second order moments of the wavefunction:

$$\bar{x} = \langle\hat{x}\rangle, \quad \bar{p} = \langle\hat{p}\rangle, \quad \sigma_x^2 = \text{Spr}(\hat{x}), \quad \sigma_{xp} = \text{Cor}(\hat{x}, \hat{p}), \quad \sigma_p^2 = \text{Spr}(\hat{p}).$$

(a) Show these exactly satisfy the equations of motion:

$$\frac{d\bar{x}}{dt} = \frac{\bar{p}}{m}, \quad \frac{d\bar{p}}{dt} = -m\omega^2\bar{x}, \quad \frac{d\sigma_x^2}{dt} = \frac{2}{m}\sigma_{xp}, \quad \frac{d\sigma_{xp}}{dt} = \frac{\sigma_p^2}{m} - m\omega^2\sigma_x^2, \quad \frac{d\sigma_p^2}{dt} = -2m\omega^2\sigma_{xp}.$$

(b) Solve those equations of motion to show that:

$$\bar{x}(t) = \bar{x}(0)\cos\omega t + \frac{\bar{p}(0)}{m\omega}\sin\omega t, \qquad \bar{p}(t) = \bar{p}(0)\cos\omega t - m\omega\bar{x}(0)\sin\omega t,$$

$$\sigma_x^2(t) = \sigma_x^2(0)\cos^2(\omega t) + \frac{2\sigma_{xp}(0)}{m\omega}\sin(\omega t)\cos(\omega t) + \frac{\sigma_p^2(0)}{m^2\omega^2}\sin^2(\omega t),$$

$$\sigma_{xp}(t) = \sigma_{xp}(0)\cos(2\omega t) + \frac{\sigma_p^2(0) - m^2\omega^2\sigma_x^2(0)}{2m\omega}\sin(2\omega t),$$
$$\sigma_p^2(t) = \sigma_p^2(0)\cos^2(\omega t) - 2m\omega\sigma_{xp}(0)\sin(\omega t)\cos(\omega t) + m^2\omega^2\sigma_x^2(0)\sin^2(\omega t).$$

Interpret what these tell you about how the quantum harmonic oscillator wave sloshes around.

(c) Show that the QHO active picture operators $(\hat{x}_A, \hat{p}_A)$ exactly satisfy the classical equations of motion:

$$\frac{d\hat{x}_A}{dt} = \frac{\hat{p}_A}{m}, \qquad \frac{d\hat{p}_A}{dt} = -m\omega^2\hat{x}_A$$

and that these are solved by (taking $t_0 = 0$):

$$\hat{x}_A(t) = \hat{x}\cos\omega t + \frac{\hat{p}}{m\omega}\sin\omega t, \qquad \hat{p}_A = \hat{p}\cos\omega t - m\omega\hat{x}\sin\omega t.$$

(d) Define the normalized central moments $N_{k,n}$ by:

$$N_{k,n} = (m\omega)^{\frac{k-n}{2}}\langle \mathrm{Sym}((\hat{x} - \bar{x}\mathbb{1}); k, (\hat{p} - \bar{p}\mathbb{1}); n)\rangle.$$

Show that their time dependence is given by:

$$N_{k,n} = \sum_{l=0}^{k}\sum_{m=0}^{n}(-1)^m\binom{k}{l}\binom{n}{m}\cos^{k+n-l-m}(\omega t)\sin^{l+m}(\omega t)N_{k-l+m,n+l-m}(0).$$

Interpret what this tells you about how the quantum harmonic oscillator wave sloshes around.

**Problem 1.35: ⋆ QHOIII - Evolution of Coherent States and Propagator.** Consider again the quantum harmonic oscillator and use the results of the previous QHO problems.

(a) Show that if the initial state vector is a coherent state $|\mathfrak{q}_0, \mathfrak{p}_0\rangle$ with $\ell = \sqrt{\frac{\hbar}{2m\omega}}$ then the second order moments of the wavefunction satisfy for all $t$:

$$\sigma_x^2(t) = \frac{\hbar}{2m\omega}, \qquad \sigma_{xp}(t) = 0, \qquad \sigma_p^2(t) = \frac{\hbar m\omega}{2}.$$

Use this and the fact that a state of minimal uncertainty is necessarily a coherent state to show that for all $t$ the wavefunction is, for some real functions $\varphi(t), \mathfrak{q}(t), \mathfrak{p}(t)$, of the form:

$$|\psi(t)\rangle = e^{i\varphi(t)}|\mathfrak{q}(t), \mathfrak{p}(t)\rangle$$

(b) Show that $\varphi(t), \mathfrak{q}(t), \mathfrak{p}(t)$ satisfy:

$$\dot{\varphi} = -\frac{\omega}{2}, \qquad \dot{\mathfrak{q}} = \frac{\mathfrak{p}}{m}, \qquad \dot{\mathfrak{p}} = -m\omega^2\mathfrak{q}.$$

Solve these to show that:

$$\varphi(t) = -\frac{\omega t}{2}, \qquad \mathfrak{q}(t) = \mathfrak{q}_0\cos(\omega t) + \frac{\mathfrak{p}_0}{m\omega}\sin(\omega t), \qquad \mathfrak{p}(t) = \mathfrak{p}_0\cos(\omega t) - m\omega\mathfrak{q}_0\sin(\omega t).$$

(c) Now consider a wavefunction with arbitrary initial conditions and use that the coherent states are an over-complete set to show that:

$$\psi(t, x) = \int_{-\infty}^{\infty} U(t, x, x')\psi(0, x')dx'$$

where $U(t, x, x')$, called the propagator, is given in terms of the time evolution operator by:

$$U(t, x, x') = \langle x|\hat{U}(t)|x'\rangle = \sqrt{\frac{m\omega}{2\pi i\hbar\sin(\omega t)}}e^{-\frac{m\omega((x^2+x'^2)\cos(\omega t) - 2xx')}{2i\hbar\sin(\omega t)}}.$$

Thus, this propagator fully solves the quantum harmonic oscillator.

(d) Consider a classical harmonic oscillator with the same value of $m$ and $\omega$ as our quantum oscillator. Let $S(t, x, x')$ be the value of its action along the classical trajectory with initial condition $x'$ at time 0 and final condition $x$ at time $t$. Show that:

$$S(t, x, x') = \frac{m\omega((x^2 + x'^2)\cos(\omega t) - 2xx')}{2\sin(\omega t)}.$$

Use this to show that the propagator for the quantum harmonic oscillator may be written exactly as:

$$U(t, x, x') = \frac{1}{\sqrt{2\pi i\hbar}}\sqrt{-\frac{\partial^2 S}{\partial x\partial x'}}e^{\frac{i}{\hbar}S(t,x,x')}.$$

**Problem 1.36: ⋆ QHOIV - Driven Oscillations.** Consider yet again the quantum harmonic oscillator.

(a) Assume that a linear force $F(t)$ is applied to the oscillator. Show that this can be modeled by the Hamiltonian operator:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 - F(t)\hat{x}.$$

(b) Show that under such driven oscillations, a coherent state still remains coherent and so has solution:

$$|\psi(t)\rangle = e^{i\varphi(t)}|\mathfrak{q}(t), \mathfrak{p}(t)\rangle \quad \text{where:} \quad \dot{\varphi} = -\frac{\omega}{2} + \frac{\mathfrak{q}(t)F(t)}{2\hbar}, \quad \dot{\mathfrak{q}} = \frac{\mathfrak{p}}{m}, \quad \dot{\mathfrak{p}} = -m\omega^2\mathfrak{q} + F(t).$$

Solve these to show that:

$$\mathfrak{q}(t) = \mathfrak{q}(0)\cos(\omega t) + \frac{\mathfrak{p}(0)}{m\omega}\sin(\omega t) + \frac{1}{m\omega}\int_0^t F(t')\sin(\omega(t-t'))dt'$$

$$\mathfrak{p}(t) = \mathfrak{p}(0)\cos(\omega t) - m\omega\mathfrak{q}(0)\sin(\omega t) + \int_0^t F(t')\cos(\omega(t-t'))dt'$$

$$\varphi(t) = -\frac{\omega t}{2} + \frac{1}{2\hbar}\int_0^t \mathfrak{q}(t')F(t')dt'.$$

(c) Following the steps used for the free quantum harmonic oscillator, show that the propagator for the driven quantum harmonic oscillator is:

$$U(t,x,x') = \sqrt{\frac{m\omega}{2\pi i\hbar\sin(\omega t)}}e^{-\frac{m\omega(((x-Q)^2+x'^2)\cos(\omega t)-2(x-Q)x')}{2i\hbar\sin(\omega t)} + \frac{im\omega}{\hbar}x\dot{Q} - \frac{im\omega}{2\hbar}Q\dot{Q} + \frac{i}{2\hbar}\int_0^t Q(t')F(t')dt'}$$

$$\text{where:} \qquad Q(t) = \int_0^t \frac{F(t')}{m\omega}\sin(\omega(t-t'))dt'.$$

(d) Consider a classical harmonic oscillator with the same value of $m$ and $\omega$ as our quantum oscillator experiencing the same linear force $F(t)$ so that its Hamiltonian is given by:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 - F(t)x.$$

Let $S(t,x,x')$ be the value of its action along the classical trajectory with initial condition $x'$ at time 0 and final condition $x$ at time $t$. Show that:

$$S(t,x,x') = \frac{m\omega(((x-Q)^2+x'^2)\cos(\omega t) - 2(x-Q)x')}{2\sin(\omega t)} + mx\dot{Q} - \frac{1}{2}mQ\dot{Q} + \frac{1}{2}\int_0^t Q(t')F(t')dt'.$$

Use this to show that the propagator for the driven QHO is exactly:

$$U(t,x,x') = \frac{1}{\sqrt{2\pi i\hbar}}\sqrt{-\frac{\partial^2 S}{\partial x\partial x'}}e^{\frac{i}{\hbar}S(t,x,x')}.$$

**Problem 1.37: ⋆ QHOV - Multivariable Oscillator.** Consider a quantum system of $d$ coupled quantum harmonic oscillators with undriven Hamiltonian operator $\hat{H}_0$ experiencing a linear force $F_a(t)$ so that the driven system has Hamiltonian operator:

$$\hat{H} = \hat{H}_0 - F_a(t)\hat{x}^a.$$

Consider as well the classical system with the Hamiltonian $\mathcal{H}$ given by replacing $\hat{x}$ and $\hat{p}$ in $\hat{H}$ with classical $x$ and $p$. Let $S(t,x,x')$ be the value of that classical system's action along the classical trajectory with initial condition $x'$ at time 0 and final condition $x$ at time $t$. Define the Van Vleck matrix $V$ as having components:

$$V_{ab}(t,x,x') = -\frac{1}{2\pi i\hbar}\frac{\partial^2 S}{\partial x^a\partial x'^b}.$$

The determinant of this matrix is called the Van Vleck determinant. Show, using the theorem of normal modes, that the propagator for the general quantum harmonic oscillator system may be written exactly as:

$$U(t,x,x') = \langle x|\hat{U}|\hat{x}'\rangle = e^{\frac{i}{\hbar}S(t,x,x')}\sqrt{\det V}.$$

**Problem 1.38: ⋆ QHOVI - Fully general quadratic Hamiltonian.** Consider a quantum system of $d$ coupled coordinate operators $\hat{x}^a$ and $d$ momenta operators $\hat{p}_a$ with a Hamiltonian operator that is any time dependent quadratic polynomial:

$$\hat{H} = E(t) - F_a(t)\hat{x}^a + v^a(t)\hat{p}_a + \frac{1}{2}K_{ab}(t)\hat{x}^a\hat{x}^b + \frac{1}{2}\Gamma^a{}_b(t)(\hat{p}_a\hat{x}^b + \hat{x}^b\hat{p}_a) + \frac{1}{2}W^{ab}(t)\hat{p}_a\hat{p}_b.$$

Consider as well the classical system with the Hamiltonian $\mathcal{H}$ given by replacing $\hat{x}$ and $\hat{p}$ in $\hat{H}$ with classical $x$ and $p$. Let $S(t_f, x_f, t_0, x_0)$ be the value of that classical system's action along the classical trajectory with initial condition $x_0$ at time $t_0$ and final condition $x_f$ at time $t_f$. Define the Van Vleck matrix $V$ as:

$$V_{ab}(t_f, x_f, t_0, x_0) = -\frac{1}{2\pi i\hbar}\frac{\partial^2 S}{\partial x_f^a \partial x_0^b}.$$

(a) Show that for some functions $A_{ab}(t_f,t_0), B_{ab}(t_f,t_0), C_{ab}(t_f,t_0), I_a(t_f,t_0), J_a(t_f,t_0), L(t_f,t_0)$ the solved action $S(t_f, x_f, t_0, x_0)$ can be written:

$$S(t_f, x_f, t_0, x_0) = \frac{1}{2}A_{ab}x_f^a x_f^b + B_{ab}x_f^a x_0^b + \frac{1}{2}C_{ab}x_0^a x_0^b + I_a x_f^a + J_a x_0^a + L.$$

Do not try to calculate these functions explicitly; simply show that they exist.

(b) For now, take the Hamiltonian operator and corresponding $\mathcal{H}$ to be piecewise constant, meaning that there is a list of times $t_n$ with $t_0 < t_1 < t_2 < ... < t_f$ so that between any two adjacent times in this list, its coefficients are constant. Show (using all of our previous harmonic oscillator results) that:

$$U(t_{n+1}, x_{n+1}, t_n, x_n) = e^{\frac{i}{\hbar}S(t_{n+1}, x_{n+1}, t_n, x_n)}\sqrt{\det V(t_{n+1}, x_{n+1}, t_n, x_n)}.$$

(c) Continuing with piecewise constancy, show that:

$$S(t_{n+2}, x_{n+2}, t_n, x_n) = S(t_{n+2}, x_{n+2}, t_{n+1}, \bar{x}_{n+1}) + S(t_{n+1}, \bar{x}_{n+1}, t_n, x_n)$$

where $\bar{x}_{n+1}$ satisfies:

$$\frac{\partial}{\partial x_{n+1}^a}(S(t_{n+2}, x_{n+2}, t_{n+1}, x_{n+1}) + S(t_{n+1}, x_{n+1}, t_n, x_n))\Big|_{x_{n+1}=\bar{x}_{n+1}} = 0.$$

Further, show that:

$$\det\frac{\partial^2 S(t_{n+2}, x_{n+2}, t_n, x_n)}{\partial x_{n+2}\partial x_n} = (-1)^d \frac{\det\frac{\partial^2 S(t_{n+2}, x_{n+2}, t_{n+1}, x_{n+1})}{\partial x_{n+2}\partial x_{n+1}} \det\frac{\partial^2 S(t_{n+1}, x_{n+1}, t_n, x_n)}{\partial x_{n+1}\partial x_n}}{\det\frac{\partial^2 (S(t_{n+2}, x_{n+2}, t_{n+1}, x_{n+1}) + S(t_{n+1}, x_{n+1}, t_n, x_n))}{\partial x_{n+1}^2}}.$$

(d) Continuing with piecewise constancy, using the composition property of the time evolution operator from problem 1.7, the results of the previous parts, and the general Gaussian integral, show that:

$$U(t_{n+2}, x_{n+2}, t_n, x_n) = e^{\frac{i}{\hbar}S(t_{n+2}, x_{n+2}, t_n, x_n)}\sqrt{\det V(t_{n+2}, x_{n+2}, t_n, x_n)}.$$

By repeatedly applying the above arguments, show that:

$$U(t_f, x_f, t_0, x_0) = e^{\frac{i}{\hbar}S(t_f, x_f, t_0, x_0)}\sqrt{\det V(t_f, x_f, t_0, x_0)}.$$

(e) Drop piecewise constancy and now assume that all the functions of time in the Hamiltonian are continuous. Explain that because a continuous function may be arbitrarily well approximated by a piecewise constant function, that even for our fully general quadratic Hamiltonian operator, we have:

$$U(t_f, x_f, t_0, x_0) = e^{\frac{i}{\hbar}S(t_f, x_f, t_0, x_0)}\sqrt{\det V(t_f, x_f, t_0, x_0)}.$$

**Problem 1.39: Landau Levels.** Consider a Schrödinger wave with electric charge $q$ in a uniform magnetic field $\vec{B}$. For convenience, define your coordinates so that $\vec{B}$ points in the $z$ direction with magnitude $B$. Show that one may choose the gauge of the vector potential so that the Hamiltonian operator is given by:

$$\vec{A} = Bx\vec{\imath}_y \implies \hat{H} = \frac{\hat{p}_x^2}{2m} + \frac{(\hat{p}_y - qBx)^2}{2m} + \frac{\hat{p}_z^2}{2m}.$$

Show that the eigenvectors (up to a normalization factor) and eigenvalues of the Hamiltonian operator are labeled by a nonnegative integer $n$ and real numbers $k_y, k_z$ by:

$$\psi_{nk_yk_z}(x,y,z) = Ne^{ik_y y + ik_z z}\Psi_n\left(x - \frac{\hbar k_y}{m\omega}\right), \qquad E_{nk_yk_z} = \hbar\omega\left(n + \frac{1}{2}\right) + \frac{\hbar^2 k_z^2}{2m}$$

where $\omega = \frac{qB}{m}$ is called the cyclotron frequency and $\Psi_n(x)$ is the $n^{\text{th}}$ quantum harmonic oscillator eigenfunction for an oscillator of mass $m$ and frequency $\omega$.

**Problem 1.40: Time Reversal Pseudo-Operator and Time Reversal Symmetry.** Let $\hat{\mathcal{K}}$ denote the

simple time reversal pseudo-operator, defined so that (in some specific orthonormal basis $|n\rangle$):

$$\hat{\mathcal{K}}|\psi\rangle = |\psi^*\rangle \implies \hat{\mathcal{K}}\left(\sum_n \psi_n|n\rangle\right) = \sum_n \psi_n^*|n\rangle.$$

The object $\hat{\mathcal{K}}$ is called a pseudo-operator because it is not complex linear.

(a) Show that $\hat{\mathcal{K}}$ is complex "anti-linear" or conjugate linear, meaning that for complex numbers $a$ and $b$ and vectors in Hilbert space $|\psi_1\rangle$ and $|\psi_2\rangle$:

$$\hat{\mathcal{K}}(a|\psi_1\rangle + b|\psi_2\rangle) = a^*\hat{\mathcal{K}}|\psi_1\rangle + b^*\hat{\mathcal{K}}|\psi_2\rangle.$$

(b) Show, using Wigner's theorem, that if $\hat{\mathcal{T}}$ is a time reversal pseudo-operator, meaning the transformation $|\psi'\rangle = \hat{\mathcal{T}}|\psi\rangle$ is a time reversing transformation then there exists a unitary operator $\hat{V}$ so that:

$$\hat{\mathcal{T}} = \hat{V}\hat{\mathcal{K}}.$$

Show that $\hat{\mathcal{T}}$ is also conjugate linear. Show that $\hat{\mathcal{T}}$ can be understood as a linear operator over the real numbers when $|\psi\rangle$ is broken up into real and imaginary components.

(c) Show that:

$$\hat{\mathcal{T}}i\hat{\mathcal{T}}^{-1} = -i.$$

(d) For an operator $\hat{\mathcal{O}}$, define the time reversed operator $\hat{\mathcal{O}}_{\mathcal{T}}$ as the operator which when acting on a time reversed state gives the time reversed behavior of $\hat{\mathcal{O}}$ acting on a non-time reversed state. Show that:

$$\hat{\mathcal{O}}_{\mathcal{T}}(\hat{\mathcal{T}}|\psi\rangle) = \hat{\mathcal{T}}(\hat{\mathcal{O}}\psi\rangle), \qquad \hat{\mathcal{O}}_{\mathcal{T}} = \hat{\mathcal{T}}\hat{\mathcal{O}}\hat{\mathcal{T}}^{-1}.$$

(e) Let $|n\rangle$ be an orthonormal basis of states which are invariant under time reversal: $\hat{\mathcal{T}}|n\rangle = |n\rangle$. Show that for any operator $\hat{\mathcal{O}}$, the matrix elements of the time reversed operator are related to the matrix elements of the original operator by:

$$\langle n|\hat{\mathcal{O}}_{\mathcal{T}}|n'\rangle = (\langle n|\hat{\mathcal{O}}|n'\rangle)^*.$$

(f) Explain why, based on the time reversal properties of spatial position $\vec{x}$, it is natural to define:

$$\hat{\mathcal{T}}|\vec{x}\rangle = |\vec{x}\rangle.$$

Show that this implies:

$$\hat{\mathcal{T}}\hat{\vec{x}}\hat{\mathcal{T}}^{-1} = \hat{\vec{x}}, \qquad \hat{\mathcal{T}}\hat{\vec{p}}\hat{\mathcal{T}}^{-1} = -\hat{\vec{p}}.$$

(g) Show that if the quantum system has time reversal symmetry then its Hamiltonian operator satisfies:

$$\hat{\mathcal{T}}\hat{H}\hat{\mathcal{T}}^{-1} = \hat{H}.$$

Suppose that $\hat{H}$ can be written solely as a function of the $\hat{\vec{x}}$ and $\hat{\vec{p}}$ operators and has time reversal symmetry. Show that this requires $\hat{H}$ to be an even function of $\hat{\vec{p}}$.

(h) Consider a scalar Schrödinger field. Show that for such a field we can take the time reversal transformation to simply be given by complex conjugation, by confirming that it satisfies the conditions in part (f):

$$|\psi'(t')\rangle = \hat{\mathcal{T}}|\psi(t)\rangle \implies \psi'(t,\vec{x}) = \psi^*(-t,\vec{x}).$$

**Problem 1.41: Schrödinger Transport Equations.** Consider the Schrödinger Lagrangian density (1.6.49) together with that for Maxwell's equations (ignoring for now their differing boost symmetries).

(a) Show that the total energy and total linear momentum of the interacting theory are given by our familiar expressions from electromagnetism plus the contributions:

$$\mathrm{E}_{\mathrm{qm}} = \int \mathcal{E}d\mathbb{S}, \qquad \vec{\mathrm{P}}_{\mathrm{qm}} = \int \vec{\wp}d\mathbb{S} \quad \text{where:} \quad \mathcal{E} = \frac{\hbar^2}{2m}|\vec{\mathscr{D}}\psi|^2, \qquad \vec{\wp} = \hbar\mathrm{Im}(\psi^*\vec{\mathscr{D}}\psi).$$

Thus $\mathcal{E}, \vec{\wp}$ are the real gauge invariant energy and linear momentum densities carried by $\psi$.

(b) Because the Galilean Schrödinger equation is Galilean symmetric, it must satisfy the nonrelativistic transport equations in Orig. theorem 6.1. Show that it does, in particular:

$$\frac{\partial\rho_m}{\partial t} + \nabla\cdot\vec{\wp} = 0, \qquad \frac{\partial\wp^a}{\partial t} - \nabla_b\sigma^{ab} = \rho_q E^a + (\vec{J}_q\times\vec{B})^a, \qquad \frac{\partial\mathcal{E}}{\partial t} + \nabla\cdot\vec{\mathcal{I}} = \vec{J}_q\cdot\vec{E}$$

where $\rho_m = m|\psi|^2$ is $\psi$'s mass density, $\sigma^{ab} = \sigma^{ba}$ is its total stress tensor, $\rho_q$ its charge density, $\vec{J}_q$ its

charge current density, $\vec{\mathcal{I}}$ its energy current density, and $\sigma^{ab}$ and $\vec{\mathcal{I}}$ are given by:

$$\sigma_{ab} = -\frac{\hbar^2}{2m}\text{Re}(\mathscr{D}_a\psi^*\mathscr{D}_b\psi - \psi^*\mathscr{D}_{(a}\mathscr{D}_{b)}\psi), \qquad \mathcal{I}_a = \frac{\hbar^3}{2m^2}\text{Im}(\mathscr{D}^b\psi^*\mathscr{D}_b\mathscr{D}_a\psi).$$

## 1.7 Angular Momentum

The most important and common internal degree of freedom is spin. In Orig. (2.8.84) we found that the total angular momentum of a rigid body decomposes into a sum of orbital angular momentum (that due to center of mass motion about a chosen origin) and spin angular momentum (that due to the rotation of the body about its own center of mass). This orbital and spin decomposition appeared later in electromagnetism in Orig. problem 4.42.(i). That same decomposition exists in quantum mechanics as well. There is orbital angular momentum due to the wavefunction flowing about around a location in space. The wavefunction can also have at each location in space an orientation in its internal space, called its polarization, which rotates under spatial rotations. The generator of such transformations on the internal space is the spin angular momentum[10]. In this section we will classify the representations of angular momentum in quantum mechanics.

Under a rotation of the wavefunction, the different components will get rotated in some way and thus become some linear combination of each other. As well, due to rotating the space by rotation matrix $\hat{R}$, the underlying point of evaluation will be rotated as well so that $\vec{x}' = \hat{R}\vec{x}$. That is, the rotated wavefunction $\psi'^a(t,\vec{x}')$ will then be related to the unrotated wavefunction by:

$$\psi'^a(t,\vec{x}) = V^a{}_b(\theta)\psi^b(t,\hat{R}^{-1}\vec{x}). \tag{1.7.1}$$

for some coefficients $V^a{}_b(\theta)$ which depend on the rotation applied. Then, if we say that the wavefunction has $D$ components and orthonormal basis vectors for the internal space $|n\rangle$ where $n$ goes from 1 to D, we can write the wavefunction as:

$$|\psi\rangle = \sum_{n=1}^{D}\int_{\mathbb{S}}\psi^n(t,\vec{x})|\vec{x}\rangle\otimes|n\rangle d^3\vec{x} = \sum_{n=1}^{D}\int_{\mathbb{S}}\psi^n(t,\vec{x})|\vec{x},n\rangle d^3\vec{x} \tag{1.7.2}$$

where we use the tensor product notation:

$$|\vec{x},n\rangle = |\vec{x}\rangle\otimes|n\rangle. \tag{1.7.3}$$

Differentiating (1.7.1) at the identity with respect to the rotation angles (with $\mathscr{D}_a$ meaning the $\theta^a$ derivative at the identity):

$$\mathscr{D}_b\psi^a(t,\vec{x}) = \left.\frac{d}{d\theta^b}\left(V^a{}_c\psi^c(t,\hat{R}^{-1}\vec{x})\right)\right|_{\theta=0} = -\epsilon_{bc}{}^d x^c\frac{\partial\psi^a}{\partial x^d} + \left(\mathscr{D}_b V^a{}_c\right)\psi^c(t,\vec{x}). \tag{1.7.4}$$

Then, we have:

$$\mathscr{D}_a|\psi\rangle = \sum_{n=1}^{D}\int_{\mathbb{S}}\mathscr{D}_a\psi^n(t,\vec{x})|\vec{x},n\rangle d^3\vec{x} = \sum_{n=1}^{D}\int_{\mathbb{S}}\left(-\epsilon_{ab}{}^c x^b\frac{\partial\psi^n}{\partial x^c} + \left(\mathscr{D}_a V^n{}_b\right)\psi^b(t,\vec{x})\right)|\vec{x},n\rangle d^3\vec{x} \tag{1.7.5}$$

$$= \sum_{n=1}^{D}\int_{\mathbb{S}}|\vec{x},n\rangle\left(-\epsilon_{ab}{}^c x^b\frac{\partial}{\partial x^c}\langle\vec{x},n| + \sum_{m=1}^{D}\left(\mathscr{D}_a V^n{}_m\right)\langle\vec{x},m|\right)d^3\vec{x}|\psi\rangle \tag{1.7.6}$$

$$= -\left(\sum_{n=1}^{D}|n\rangle\langle n|\right)\otimes\epsilon_{ab}{}^c\hat{x}^b\int_{\mathbb{S}}|\vec{x}\rangle\frac{\partial}{\partial x^c}\langle\vec{x}|d^3\vec{x}|\psi\rangle + \left(\sum_{\substack{n=1\\m=1}}^{D,D}\mathscr{D}_a V^n{}_m|n\rangle\langle m|\right)\otimes\int_{\mathbb{S}}|\vec{x}\rangle\langle\vec{x}|d^3\vec{x}|\psi\rangle \tag{1.7.7}$$

[10]It is a very common error among quantum mechanics books to say that spin has no classical analog. This is entirely wrong but is a second-order error (it is a wrong result derived from a more basic error). The first-order error which is the parent of this mistake is the belief that the quantum mechanics of a single particle is about a single *point-particle*. Hopefully by now you understand why that belief if fundamentally misguided and know that quantum mechanics of a single particle is about the dynamics of an extended wave-packet. It is true that point particles do not have any spin (spin is finitely proportional to the moment of inertia which itself scales like the squared radius of the body, so 0 in the point case) but neither the classical mechanics of spinning bodies nor quantum mechanics of a particle have anything to do with point particles. Obviously classical bodies of finite size do have spin.

$$= -\mathbb{1}_{\text{int}} \otimes \epsilon_{ab}{}^{c}\hat{x}^{b}\frac{i}{\hbar}\hat{p}_{c}|\psi\rangle + \left(\sum_{\substack{n=1\\m=1}}^{D,D} \mathscr{D}_{a}V^{n}{}_{m}|n\rangle\langle m|\right) \otimes \mathbb{1}_{\mathbb{S}}|\psi\rangle \tag{1.7.8}$$

$$= -\frac{i}{\hbar}(\hat{\vec{x}} \times \hat{\vec{p}})_{a}|\psi\rangle + \left(\sum_{\substack{n=1\\m=1}}^{D,D} \mathscr{D}_{a}V^{n}{}_{m}|n\rangle\langle m|\right)|\psi\rangle. \tag{1.7.9}$$

In the last line we have applied a common abuse of notation. If $\hat{A}$ is an operator in Hilbert space $\mathbb{H}_1$ and $\mathbb{1}_2$ is the identity on Hilbert space $\mathbb{H}_2$, when acting the operator $\hat{A} \otimes \mathbb{1}_2$ on a vector $|\psi\rangle$ in the Hilbert space $\mathbb{H}_1 \otimes \mathbb{H}_2$, it is common to write $\hat{A}|\psi\rangle$ even though more rigorously $\hat{A} \otimes \mathbb{1}_2|\psi\rangle$ is intended.

From the change in $|\psi\rangle$ under a rotation and (1.4.5), we have that the generator of rotations (the total angular momentum $\mathrm{J}_a$) of the wavefunction and the corresponding total angular momentum operator $\hat{J}_a$ are given by:

$$\mathrm{J}_{a} = i\hbar\langle\psi|\mathscr{D}_{a}|\psi\rangle = \langle\psi|\hat{J}_{a}|\psi\rangle, \qquad \hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}} \tag{1.7.10}$$

where the orbital angular momentum operator $\hat{\vec{L}}$, which acts only on spatial degrees of freedom, is the same as we found in (1.6.96) and the spin angular momentum operator, which acts only on the wavefunction's internal degrees of freedom, is given by:

$$\hat{\vec{S}} = -\frac{i}{\hbar}\sum_{\substack{n=1\\m=1}}^{D,D} \vec{\mathscr{D}}V^{n}{}_{m}|n\rangle\langle m|. \tag{1.7.11}$$

This $\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}}$ decomposition is just as a special case of theorem 1.5. The calculation as done here is only of the free angular momentum of the wavefunction, since we didn't include possible contributions from any fields it might interact with. As always for the wavefunction, the generating quantities are given by the operator summaries of their corresponding operators:

$$\vec{\mathrm{J}} = \langle\hat{\vec{J}}\rangle, \qquad \vec{\mathrm{L}} = \langle\hat{\vec{L}}\rangle, \qquad \vec{\mathrm{S}} = \langle\hat{\vec{S}}\rangle, \qquad \vec{\mathrm{J}} = \vec{\mathrm{L}} + \vec{\mathrm{S}}. \tag{1.7.12}$$

We found in Orig. (2.4.53) that angular momentum must satisfy the so(3) algebra and consequently due to Theorem 1.5 we must have:

$$\{\mathrm{J}_a, \mathrm{J}_b\} = \epsilon_{ab}{}^{c}\mathrm{J}_c \implies [\hat{J}_a, \hat{J}_b] = i\hbar\epsilon_{ab}{}^{c}\hat{J}_c. \tag{1.7.13}$$

We will call these commutators the su(2) algebra (wait! What? Why isn't it called the so(3) algebra?? What gives? Don't worry. Just be patient for a few pages). By direct computation, the orbital angular momentum operator $\hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}}$ we found already satisfies:

$$[\hat{L}_a, \hat{L}_b] = i\hbar\epsilon_{ab}{}^{c}\hat{L}_c \tag{1.7.14}$$

as it should. Because the orbital angular momentum operator acts only on the spatial degrees of freedom of $|\psi\rangle$ and the spin angular momentum operator acts only on the internal degrees of freedom of $|\psi\rangle$, they commute automatically:

$$[\hat{L}_a, \hat{S}_b] = 0. \tag{1.7.15}$$

Then, because $\hat{J}_a = \hat{L}_a + \hat{S}_a$, in order for $\hat{J}_a$ to satisfy the su(2) algebra, $\hat{S}_a$ must satisfy it as well:

$$[\hat{S}_a, \hat{S}_b] = i\hbar\epsilon_{ab}{}^{c}\hat{S}_c. \tag{1.7.16}$$

This must be true for the spin operators of any quantum system! As well, from Theorem 1.5, $\hat{J}_a$, $\hat{L}_a$, and $\hat{S}_a$ must all be self-adjoint.

### 1.7.1 Representations of SU(2)

Let us now classify all of the ways in which the su(2) algebra can be satisfied and the SU(2) group can be represented. To do so, we will use $J$ but because all of the results follow from the algebra, everything applies for $L$ and $S$ as well until noted otherwise. Representations of a Lie algebra can be classified as reducible or as irreducible. A representation is irreducible if there is no subset of that representation which is itself a nontrivial representation of the algebra. A representation is reducible if there is such a subset.

For example, $3 \times 3$ rotation matrices form an irreducible representation of the group SO(3). There is no subset of those matrices which is also a representation of SO(3) (except the identity matrix which gives a trivial representation of SO(3)). The $3 \times 3$ rotation matrices represent how a single vector is rotated. An ordered pair of $3 \times 3$ rotation matrices represents how two separate vectors are rotated. That ordered pair of $3 \times 3$ rotation matrices forms a reducible representation of the group SO(3) (the ordered pair $(\hat{R}_1, \hat{R}_2)$ can be reduced into $(\hat{R}_1, \mathbb{1})$, $(\mathbb{1}, \hat{R}_2)$ which are both irreducible representations). We only need to understand all irreducible representations of a group or algebra and can build the reducible representations by just combinations of the irreducible ones.

Now turning our attention to angular momentum, for convenience define $\hat{J}^2, \hat{J}_\pm$ as:

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2, \qquad \hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y. \tag{1.7.17}$$

Automatically $\hat{J}^2$ is self-adjoint and $\hat{J}_- = \hat{J}_+^\dagger$. As well, by (1.7.13):

$$[\hat{J}^2, \hat{J}_a] = [\hat{J}_b\hat{J}^b, \hat{J}_a] = \hat{J}_b[\hat{J}^b, \hat{J}_a] + [\hat{J}_b, \hat{J}_a]\hat{J}^b = \hat{J}^b i\hbar\epsilon_{ba}{}^c\hat{J}_c + i\hbar\epsilon_{ba}{}^c\hat{J}_c\hat{J}^b \implies [\hat{J}^2, \hat{J}_a] = 0 \tag{1.7.18}$$

Hence $\hat{J}^2$ commutes with the angular momentum operator along any particular axis. By convention we tend to designate the $z$ axis when making explicit calculations, just like with spherical coordinates. From (1.7.18) we conclude in particular that:

$$[\hat{J}^2, \hat{J}_z] = 0. \tag{1.7.19}$$

Because $\hat{J}^2$ and $\hat{J}_z$ are self-adjoint, they are diagonalizable by the spectral theorem and because they commute they are simultaneously diagonalizable by some orthonormal eigenbasis $|j, m\rangle$. Here $j$ labels the eigenvalue of $\hat{J}^2$ and $m$ labels the eigenvalue of $\hat{J}_z$. When I say "labels" I don't mean that $j$ and $m$ are necessarily equal to the eigenvalues of $\hat{J}^2$ and $\hat{J}_z$, simply that there will be an invertible map between $j$ and the eigenvalue of $\hat{J}^2$ and between $m$ and the eigenvalue of $\hat{J}_z$. We will choose that map for convenience as we continue. In particular, we choose $\hbar m$ to be the eigenvalue of $\hat{J}_z$ but we use $\lambda(j)$ for the eigenvalue of $\hat{J}^2$, which will be specified later. Because the $\hat{J}_a$ are all self-adjoint, they all have real eigenvalues and hence $\hat{J}_a^2$ are all nonnegative operators. Therefore,

$$\langle \hat{J}^2 \rangle = \langle \hat{J}_x^2 \rangle + \langle \hat{J}_y^2 \rangle + \langle \hat{J}_z^2 \rangle \geq \langle \hat{J}_z^2 \rangle \tag{1.7.20}$$

for all $|\psi\rangle$. In particular, for our eigenstates:

$$\langle j, m|\hat{J}_z^2|j, m\rangle = \hbar^2 m^2, \qquad \langle j, m|\hat{J}^2|j, m\rangle = \lambda(j) \implies \hbar^2 m^2 \leq \lambda(j). \tag{1.7.21}$$

Hence the eigenvalues of $\hat{J}_z$ are bounded at least this strongly by the eigenvalues of $\hat{J}^2$. From the definitions of $\hat{J}_\pm$, a short calculation gives:

$$[\hat{J}^2, \hat{J}_\pm] = 0, \qquad [\hat{J}_z, \hat{J}_\pm] = \pm\hbar\hat{J}_\pm. \tag{1.7.22}$$

Now consider the action of $\hat{J}_\pm$ on the $|j, m\rangle$ eigenstates:

$$\hat{J}_z|j, m\rangle = \hbar m|j, m\rangle \implies \hat{J}_\pm\hat{J}_z|j, m\rangle = \hbar m\hat{J}_\pm|j, m\rangle \tag{1.7.23}$$

$$\implies -[\hat{J}_z, \hat{J}_\pm]|j, m\rangle + \hat{J}_z\hat{J}_\pm|j, m\rangle = \hbar m\hat{J}_\pm|j, m\rangle \implies \hat{J}_z\left(\hat{J}_\pm|j, m\rangle\right) = \hbar(m \pm 1)\hat{J}_\pm|j, m\rangle. \tag{1.7.24}$$

Therefore, $\hat{J}_\pm|j, m\rangle$ is an eigenstate of $\hat{J}_z$ with eigenvalue $\hbar(m \pm 1)$. Completing the same with $\hat{J}^2$:

$$\hat{J}^2|j, m\rangle = \lambda(j)|j, m\rangle \implies \hat{J}_\pm\hat{J}^2|j, m\rangle = \lambda(j)\hat{J}_\pm|j, m\rangle \implies \hat{J}^2\hat{J}_\pm|j, m\rangle = \lambda(j)\hat{J}_\pm|j, m\rangle. \tag{1.7.25}$$

Hence $\hat{J}_\pm|j, m\rangle$ is an eigenstate of $\hat{J}^2$ with eigenvalue $\lambda(j)$. With these two facts together, we conclude that up to normalization $N_\pm(j, m)$:

$$\hat{J}_\pm|j, m\rangle = N_\pm(j, m)|j, m \pm 1\rangle. \tag{1.7.26}$$

We can see that $\hat{J}_+$ acts as a raising operator, increasing $m$ at fixed $j$, and that $\hat{J}_-$ acts as a lowering operator, decreasing $m$ at fixed $j$. It appears that we can keep raising or lowering the eigenvalues of $\hat{J}_z$ indefinitely using $\hat{J}_\pm$. However, because $\hbar^2 m^2 \leq \lambda$ by (1.7.21) this raising and lowering must terminate if pushing further would cause this bound to be violated. Since we are to only be able to raise $m$ a finite amount, there must be some $m_+$ so that attempting to raise it any further causes the right hand side of (1.7.26) to go to 0 (else we could solve for the higher $m$ state appearing on that side). That is, there must exist an $m_+$ so that $\hat{J}_+|j, m_+\rangle = 0$. Similarly if we are only able to lower $m$ a finite amount, there must be some $m_-$ so that $\hat{J}_-|j, m_-\rangle = 0$. Starting with $|j, m_-\rangle$ and iteratively applying the raising operator must eventually give 0 which is precisely the condition which defines $m_+$. However, iteratively applying the raising operator will only

increment by $m_-$ by 1 each time. Hence, by incrementing $m_-$ by 1 sufficiently one reaches $m_+$ and therefore $m_+ - m_-$ is a positive integer. (You may argue that there may be two different values of $m_+$. However, having two such branches would then give two families of eigenvectors $|j,m\rangle$ which do not mix under the algebra and so would give us a reducible representation.) Further, because we go through one state for each $m$, we have $m_+ - m_- + 1$ eigenvectors for $\hat{J}_z$ at fixed $j$. So the dimension $D(j)$ of the subspace of the vector space with a particular value of $j$ is $D(j) = m_+ - m_- + 1$.

Now rewrite $\hat{J}^2$ in terms of the $\hat{J}_\pm$ operators instead of $\hat{J}_x$ and $\hat{J}_y$ and use the commutation relation between $\hat{J}_+$ and $\hat{J}_-$. Short calculations give:

$$[\hat{J}_+, \hat{J}_-] = 2\hbar\hat{J}_z, \qquad \hat{J}_x = \frac{\hat{J}_+ + \hat{J}_-}{2}, \qquad \hat{J}_y = \frac{\hat{J}_+ - \hat{J}_-}{2i} \tag{1.7.27}$$

$$\hat{J}^2 = \hat{J}_+\hat{J}_- - \hbar\hat{J}_z + \hat{J}_z^2 = \hat{J}_-\hat{J}_+ + \hbar\hat{J}_z + \hat{J}_z^2. \tag{1.7.28}$$

Now let us act these relationships on our eigenstates $|j,m_\pm\rangle$.

$$\hat{J}^2|j,m_+\rangle = \hat{J}_-\hat{J}_+|j,m_+\rangle + \hbar\hat{J}_z|j,m_+\rangle + \hat{J}_z^2|j,m_+\rangle \tag{1.7.29}$$

$$\implies \lambda(j)|j,m_+\rangle = 0 + \hbar^2 m_+|j,m_+\rangle + \hbar^2 m_+^2|j,m_+\rangle = \hbar^2 m_+(m_+ + 1)|j,m_+\rangle \tag{1.7.30}$$

$$\implies \lambda(j) = \hbar^2 m_+(m_+ + 1) \tag{1.7.31}$$

$$\hat{J}^2|j,m_-\rangle = \hat{J}_+\hat{J}_-|j,m_-\rangle - \hbar\hat{J}_z|j,m_-\rangle + \hat{J}_z^2|j,m_-\rangle \tag{1.7.32}$$

$$\implies \lambda(j)|j,m_-\rangle = 0 - \hbar^2 m_-|j,m_-\rangle + \hbar^2 m_-^2|j,m_-\rangle = \hbar^2 m_-(m_- - 1)|j,m_-\rangle \tag{1.7.33}$$

$$\implies \lambda(j) = \hbar^2 m_-(m_- - 1). \tag{1.7.34}$$

Hence we have two expression for $\lambda$. Equating them:

$$\hbar^2(m_+^2 + m_+) = \hbar^2(m_-^2 - m_-) \implies (m_+ + m_-)(m_+ - m_- + 1) = 0 \tag{1.7.35}$$

$$\implies D(j)(m_+ + m_-) = 0 \implies m_- = -m_+ \tag{1.7.36}$$

because $D(j)$ is never 0 (even if the internal polarization space of the wavefunction is trivial so that there is only 1 single state in it, $D = 1$). We now define $j = m_+ = -m_-$. Then:

$$D(j) = m_+ - m_- + 1 = 2j + 1 \implies j = \frac{D-1}{2}. \tag{1.7.37}$$

Hence the value $j$ comes in discrete half-integer values (because $D$ comes in natural numbers). This is purely necessitated by the algebra. For $D$ as $1, 2, 3$ we have $j$ as $0, 1/2, 1$ and so on. These different $j$ give different representations of the angular momentum and some common ones are named. In the case that we are discussing spin angular momentum in particular, we write $s$ instead of $j$ and systems with $s = 0$ are called spin-0 or scalars, systems with $s = 1/2$ are called spin-1/2 or spinors (although spinor is sometimes used for any half-integer spin), systems with $s = 1$ are called spin-1 or vectors. Because $D = 2s + 1$ gives the dimension of the angular momentum space at fixed $s$, different fields with different spins have different spin spaces. From (1.7.31), $\lambda = \hbar^2 j(j+1)$. Therefore, we have that the angular momentum eigenvectors are fundamentally defined to satisfy:

$$\hat{J}^2|j,m\rangle = \hbar^2 j(j+1)|j,m\rangle, \qquad \hat{J}_z|j,m\rangle = \hbar m|j,m\rangle \qquad -j \le m \le j. \tag{1.7.38}$$

We can now compute the normalization factors $N_\pm$ to get exact expressions for the raising and lowering operators.

$$\hat{J}_\pm|j,m\rangle = N_\pm(j,m)|j,m\pm1\rangle \implies \|\hat{J}_+|j,m\rangle\|^2 = \|N_+(j,m)|j,m+1\rangle\|^2 \tag{1.7.39}$$

$$\implies \langle j,m|\hat{J}_+^\dagger\hat{J}_+|j,m\rangle = \langle j,m|\hat{J}_-N_+(j,m)|j,m+1\rangle = N_+(j,m)N_-(j,m+1) = N_+^*(j,m)N_+(j,m) \tag{1.7.40}$$

$$\implies N_-(j,m+1) = N_+^*(j,m). \tag{1.7.41}$$

Now using the commutator between $\hat{J}_+$ and $\hat{J}_-$:

$$\hat{J}_-\hat{J}_+|j,m\rangle = \hat{J}_-N_+(j,m)|j,m+1\rangle = N_+(j,m)N_-(j,m+1)|j,m\rangle = |N_+(j,m)|^2|j,m\rangle \tag{1.7.42}$$

$$\hat{J}_+\hat{J}_-|j,m\rangle = \hat{J}_+N_-(j,m)|j,m-1\rangle = N_-(j,m)N_+(j,m-1)|j,m\rangle = |N_+(j,m-1)|^2|j,m\rangle \tag{1.7.43}$$

$$\implies 2\hbar^2 m|j,m\rangle = 2\hbar\hat{J}_z|j,m\rangle = [\hat{J}_+,\hat{J}_-]|j,m\rangle = \left(|N_+(j,m-1)|^2 - |N_+(j,m)|^2\right)|j,m\rangle \tag{1.7.44}$$

$$\implies 2\hbar^2 m = |N_+(j,m-1)|^2 - |N_+(j,m)|^2. \tag{1.7.45}$$

We know that $N_+(j,j) = 0$ because the $|j,j\rangle$ state when raised must go to 0 by definition. Now take (1.7.45)

and sum from $m = l$ to $m = j$:

$$2\hbar^2 \sum_{m=l}^{j} m = \sum_{m=l}^{j} |N_+(j, m-1)|^2 - \sum_{m=l}^{j} |N_+(j, m)|^2 \tag{1.7.46}$$

$$\implies 2\hbar^2 \left( \frac{j(j+1)}{2} - \frac{(l-1)l}{2} \right) = |N_+(j, l-1)|^2 \tag{1.7.47}$$

$$\implies N_+(j, m) = \hbar e^{i\varphi(j,m)} \sqrt{j(j+1) - m(m+1)} \tag{1.7.48}$$

for some phases $\varphi(j, m)$ that are undetermined because the phase of eigenvectors is set only by convention even after normalization. From a given single state, $|j, j\rangle$, if we define all the other eigenvectors $|j, m\rangle$ for $m < j$ by lowering without a phase, then we absorb these phases into the definitions of the eigenstates, leaving us with:

$$\hat{J}_\pm |j, m\rangle = \hbar \sqrt{j(j+1) - m(m \pm 1)} |j, m \pm 1\rangle. \tag{1.7.49}$$

Now that we know exactly how the $\hat{J}_z$ and raising and lower operators act, we can find explicit forms for all of the angular momentum operators.

Here we will split our discussion into considering spin and orbital angular momenta separately. For spin angular momentum, if the wavefunction has $D$ components which all mutually become linear combinations of each other under rotations we have that it is described by spin $s = \frac{D-1}{2}$. Then, because $\hat{S}_z$ is self-adjoint and therefore normal, its orthonormal eigenvectors form a basis for the spin space. This allows us to use the completeness relation:

$$\sum_{m=-s}^{s} |s, m\rangle\langle s, m| = \mathbb{1}. \tag{1.7.50}$$

Beginning with the $\hat{S}^2$ operator:

$$\hat{S}^2 |s, m\rangle = \hbar^2 s(s+1) |s, m\rangle \implies \hat{S}^2 \sum_{m=-s}^{s} |s, m\rangle\langle s, m| = \hbar^2 s(s+1) \sum_{m=-s}^{s} |s, m\rangle\langle s, m| \tag{1.7.51}$$

$$\implies \hat{S}^2 = \hbar^2 s(s+1) \mathbb{1}. \tag{1.7.52}$$

The next easiest is $\hat{S}_z$:

$$\hat{S}_z |s, m\rangle = \hbar m |s, m\rangle \implies \hat{S}_z \sum_{m=-s}^{s} |s, m\rangle\langle s, m| = \sum_{m=-s}^{s} \hbar m |s, m\rangle\langle s, m| \tag{1.7.53}$$

$$\implies \hat{S}_z = \sum_{m=-s}^{s} \hbar m |s, m\rangle\langle s, m|. \tag{1.7.54}$$

The raising and lower operators are slightly nastier because they are not diagonal:

$$\hat{S}_+ |s, m\rangle = \hbar \sqrt{s(s+1) - m(m+1)} |s, m+1\rangle \tag{1.7.55}$$

$$\implies \hat{S}_+ \sum_{m=-s}^{s} |s, m\rangle\langle s, m| = \sum_{m=-s}^{s} \hbar \sqrt{s(s+1) - m(m+1)} |s, m+1\rangle\langle s, m| \tag{1.7.56}$$

$$\implies \hat{S}_+ = \sum_{m=-s}^{s-1} \hbar \sqrt{s(s+1) - m(m+1)} |s, m+1\rangle\langle s, m| \tag{1.7.57}$$

$$\implies \hat{S}_- = S_+^\dagger = \sum_{m=-s+1}^{s} \hbar \sqrt{s(s+1) - m(m-1)} |s, m-1\rangle\langle s, m|. \tag{1.7.58}$$

Finally we can use these to complete the $\hat{S}_x$ and $\hat{S}_y$ operators by (1.7.27). Hence we can compute the spin matrices in general and in principle have all of the setup for use with any spin $s$.

**Theorem 1.11: Spin Representation Theorem**

The spin angular momentum of a wavefunction in quantum mechanics is given by $\mathrm{S}_a = \langle \hat{S}_a \rangle$ where $\hat{S}_a$

are the internal self-adjoint spin operators which satisfy the su(2) commutation relations:

$$[\hat{S}_a, \hat{S}_b] = i\hbar\epsilon_{ab}{}^c\hat{S}_c. \tag{1.7.59}$$

The spin operators have irreducible representations labeled by $s$, the total spin, which takes values $0, 1/2, 1, 3/2, 2, ...$ determined by the eigenvalue of the $\hat{S}^2$ operator as:

$$\hat{S}^2|\psi\rangle = \left(\hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2\right)|\psi\rangle = \hbar^2 s(s+1)|\psi\rangle. \tag{1.7.60}$$

The representation of spin $s$ has dimension $D = 2s + 1$ and is spanned by the orthonormal states $|s, m\rangle$, $-s \leq m \leq s$, defined by:

$$\hat{S}^2|s, m\rangle = \hbar^2 s(s+1)|s, m\rangle, \qquad \hat{S}_z|s, m\rangle = \hbar m|s, m\rangle. \tag{1.7.61}$$

In terms of these states, the spin operators have the explicit forms:

$$\hat{S}_x = \sum_{m=-s}^{s-1} \hbar\sqrt{s(s+1) - m(m+1)}\frac{|s, m+1\rangle\langle s, m| + |s, m\rangle\langle s, m+1|}{2} \tag{1.7.62}$$

$$\hat{S}_y = \sum_{m=-s}^{s-1} \hbar\sqrt{s(s+1) - m(m+1)}\frac{|s, m+1\rangle\langle s, m| - |s, m\rangle\langle s, m+1|}{2i} \tag{1.7.63}$$

$$\hat{S}_z = \sum_{m=-s}^{s} \hbar m|s, m\rangle\langle s, m|. \tag{1.7.64}$$

An initially curious (and later highly consequential) fact comes from looking at what happens when we rotate about a particular axis by $2\pi$. Take it to be the $z$ axis for convenience. By (1.7.1), we can see that the $V^a{}_b$ appearing there are simply the components of the operator $\hat{U}_{\text{int}}$ appearing in (1.4.34). Thus, we have that:

$$\hat{U}_{\text{int}}(\theta) = e^{-\frac{i}{\hbar}\theta\hat{S}_z} \tag{1.7.65}$$

for a $z$ axis rotation by $\theta$ from Theorem 1.4. We can compute this matrix exponential explicitly by using the fact that $\hat{S}_z$ is diagonal:

$$\hat{U}_{\text{int}}(\theta)|s, m\rangle = e^{-\frac{i}{\hbar}\theta\hat{S}_z}|s, m\rangle = e^{-im\theta}|s, m\rangle \implies \hat{U}_{\text{int}}(\theta) = \sum_{m=-s}^{s} e^{-im\theta}|s, m\rangle\langle s, m|. \tag{1.7.66}$$

Now let us look at what this becomes for a $2\pi$ rotation. If $s$ is an integer then so is $m$ and $e^{-2\pi im}$ for all integer $m$ is just 1. If $s$ is an integer plus a half then so is $m$ and $e^{-2\pi im}$ for all half-integers $m$ is $-1$. These can both be captured by $(-1)^{2s}$, and so:

$$\hat{U}_{\text{int}}(2\pi) = (-1)^{2s}\mathbb{1} \implies V^a{}_b = (-1)^{2s}\delta^a_b. \tag{1.7.67}$$

Then, under a rotation by $2\pi$, (1.7.1) becomes:

$$\psi'^a(t, \vec{x}) = (-1)^{2s}\psi^a(t, \vec{x}) \implies |\psi'\rangle = (-1)^{2s}|\psi\rangle. \tag{1.7.68}$$

Here $\vec{x}' = \vec{x}$ because under a $2\pi$ rotation any point in space returns to itself. Under such a rotation the wavefunction may or may not pick up a minus sign, depending on its spin! This fact surprises everyone the first time they see it, because shouldn't a $2\pi$ rotation always return the world back to the way it was? Those who aren't careful think of the possibility of a minus sign on $|\psi\rangle$ after a $2\pi$ rotation as one of the great mysteries of quantum mechanics, proof that it can never be understood. It is true that you cannot understand quantum mechanics if you ignore how quantum mechanics works. But, in fact that sign makes total sense. Recall that the wavefunction carries one single overall meaningless phase and a minus sign is precisely a phase of $\pi$. In particular, $|\psi'\rangle$ and $|\psi\rangle$ here produce all of the same operator summaries:

$$\langle\psi'|\hat{G}|\psi'\rangle = (-\langle\psi|)\hat{G}(-|\psi\rangle) = \langle\psi|\hat{G}|\psi\rangle \tag{1.7.69}$$

and so because these are the quantities which we relate to experiment, all of the physical properties of the system are returned exactly to what they were after a $2\pi$ rotation, exactly as you would expect. This sign does have consequences though for the global structure of the group of rotations in quantum mechanics; it is why (1.7.13) is actually the su(2) algebra and not the so(3) algebra. Due to this minus sign, it actually takes a rotation by $4\pi$ to return the wavefunction back to itself in full generality. The group SO(3) (of orthogonal $3 \times 3$ matrices with determinant one) only has representations which return to themselves after a $2\pi$ rotation

(and thus only have integer $s$) whereas the group SU(2) (of unitary $2\times 2$ matrices with determinant 1) actually has the exact same Lie algebra as SO(3) but has group representations that allow for half-integer $s$. These half-integer representations are allowed for spin because $|\psi\rangle$ and $-|\psi\rangle$ correspond to the same physical state.

Let us unpack all of that a bit further by looking at the simplest representations of our algebra. We start with $D=1$, and we see from Theorem 1.11 that:

$$s=0 \implies \hat{S}_x=\hat{S}_y=\hat{S}_z=0. \tag{1.7.70}$$

These operators do in fact satisfy the commutation relations trivially and so this representation is called the trivial representation. The smallest representation which satisfies the commutation relations without making all of the operators 0 (called the fundamental representation) is for $D=2$ and consequently $s=\frac{1}{2}$. Then, we see from Theorem 1.11 that:

$$s=\frac{1}{2} \implies \hat{S}_x=\frac{\hbar}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad S_y=\frac{\hbar}{2}\begin{pmatrix}0&-i\\i&0\end{pmatrix},\quad S_z=\frac{\hbar}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix}. \tag{1.7.71}$$

These are all $2\times 2$ self-adjoint matrices with 0 trace and so the rotations they create on the internal space are $2\times 2$ unitary matrices:

$$\hat{U}_{\text{int}}=e^{-\frac{i}{\hbar}(\theta_x\hat{S}_x+\theta_y\hat{S}_y+\theta_z\hat{S}_z)} \implies \hat{U}^\dagger=e^{\frac{i}{\hbar}(\theta_x\hat{S}_x+\theta_y\hat{S}_y+\theta_z\hat{S}_z)}=\hat{U}^{-1}. \tag{1.7.72}$$

As well, because these $\hat{S}_a$ are traceless, the determinant of $\hat{U}_{\text{int}}$ is automatically 1. This can be shown by using that the determinant of a diagonalizable matrix is given by the product of its eigenvalues and the trace is given by the sum of its eigenvalues:

$$\det(\hat{U}_{\text{int}})=\det\left(e^{-\frac{i}{\hbar}\theta^a\hat{S}_a}\right)=\prod_n\lambda_n=\prod_n e^{\ln\lambda_n}=e^{\sum_n\ln\lambda_n}=e^{\text{Tr}(\ln\hat{U})}=e^{\text{Tr}\left(-\frac{i}{\hbar}\theta^a\hat{S}_a\right)}=e^0=1. \tag{1.7.73}$$

Thus, the Lie group made by the internal rotations is the Lie group of $2\times 2$ Unitary matrices of determinant 1, i.e. SU(2). All of the other angular momentum representations are just different representations of SU(2). All the way back in (2.4.53) when we found the Poisson brackets for angular momentum we did so considering $3\times 3$ rotation matrices, which form the SO(3) group. Because the generators of SO(3) and SU(2) have the same Poisson brackets, they have the same Lie algebras, i.e. $\mathfrak{so}(3)=\mathfrak{su}(2)$. However, the Lie groups are not quite the same. For every SO(3) matrix there are two SU(2) matrices which correspond to it (a particular matrix and its negative). So, the Lie group SU(2) is the "double-cover" of SO(3) as there is a map from SU(2) into SO(3) which preserves the commutation relations but it is 2-to-1, not 1-to-1. Since we tend to think in 3 dimensions rather than $2\times 2$ matrices, we also call the group of rotations as given here by SU(2) as the group Spin(3).

A very important advantage of SU(2) over SO(3) is that SU(2) is simply connected, while SO(3) is not. A $3\times 3$ rotation matrix is determined by a unit vector $\vec{n}$ as the axis of rotation and a magnitude $\theta$ as the angle of rotation. We only ever need to define rotations with magnitude $\leq\pi$ because a rotation by an angle of greater than $\pi$ is equivalent to a rotation in the opposite direction by an angle less than $\pi$. With this, we can visualize SO(3) as a ball in 3 spatial dimensions of radius $\pi$. Every point in the ball specifies $\theta$ by its distance from the origin and $\vec{n}$ as the unit vector from the origin to that point. However, antipodal points on the surface of the ball (points which are directly opposite each other, like the north pole and the south pole) actually correspond to the same rotation. A rotation in the $z$ direction by $\pi$ is the same as a rotation in the $-z$ direction by $\pi$. One can consider a path running directly from the north pole of the ball to the south pole. This path is technically a closed loop because the north and south poles are the same rotation. However, there is no way to continuously deform this closed loop within the ball so that it shrinks down to a single point. Because of this, SO(3) is not simply connected. Because SU(2) is the double-cover of SO(3), we can visualize it as two balls of radius $\pi$ where the north pole of one ball is identified with the south pole of the other (and similarly so for each antipodal pair). In this two ball picture, every closed loop can be continuously deformed to a point, and so SU(2) is simply connected.

### 1.7.2 Spherical Harmonics

Let us now study the angular momentum commutation relations with specific attention to orbital angular momentum. To do so it is nicest to use spherical coordinates $(r,\theta,\varphi)$. Then, the orbital angular momentum operators become:

$$\langle\vec{x}|\hat{L}_x|\psi\rangle=\langle\vec{x}|(\hat{y}\hat{p}_z-\hat{z}\hat{p}_y)|\psi\rangle=-i\hbar\left(y\frac{\partial\psi}{\partial z}-z\frac{\partial\psi}{\partial y}\right)=i\hbar\left(\sin\varphi\frac{\partial\psi}{\partial\theta}+\frac{\cos\varphi}{\tan\theta}\frac{\partial\psi}{\partial\varphi}\right) \tag{1.7.74}$$

$$\langle\vec{x}|\hat{L}_y|\psi\rangle = \langle\vec{x}|(\hat{z}\hat{p}_x - \hat{x}\hat{p}_z)|\psi\rangle = -i\hbar\left(z\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial z}\right) = i\hbar\left(-\cos\varphi\frac{\partial\psi}{\partial\theta} + \frac{\sin\varphi}{\tan\theta}\frac{\partial\psi}{\partial\varphi}\right) \tag{1.7.75}$$

$$\langle\vec{x}|\hat{L}_z|\psi\rangle = \langle\vec{x}|(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x)|\psi\rangle = -i\hbar\left(x\frac{\partial\psi}{\partial y} - y\frac{\partial\psi}{\partial x}\right) = -i\hbar\frac{\partial\psi}{\partial\varphi}. \tag{1.7.76}$$

As well, $\hat{L}^2$ becomes after a bit of effort:

$$\langle\vec{x}|\hat{L}^2|\psi\rangle = -\hbar^2\left(\frac{\partial^2\psi}{\partial\theta^2} + \frac{1}{\tan\theta}\frac{\partial\psi}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2}\right) \tag{1.7.77}$$

Let us define $\hat{r} = \sqrt{\hat{x}^2 + \hat{y}^2 + \hat{z}^2}$, which we can immediately see commutes with all 3 $\hat{L}_a$ because they contain no $r$ derivatives when put into coordinates. Since they commute, we can find a simultaneous eigenbasis. In fact, at fixed $r$, $\psi$ is only a function of $\theta$ and $\varphi$ and so we can define the Hilbert space $\mathbb{H}_\Omega$ of (square-integrable) functions of $(\theta,\varphi)$, i.e. functions on the unit sphere $\Omega$. Let $Y_\alpha(\theta,\varphi)$ be an orthonormal basis of such functions. At fixed $(t,r)$, $\psi(t,\vec{x})$ can necessarily be written as a linear combination of them and for different values of $r$ and $t$ the coefficients of linear combination may be different, so we have:

$$\psi(t,\vec{x}) = \sum_\alpha R_\alpha(t,r)Y_\alpha(\theta,\varphi). \tag{1.7.78}$$

A point on $\Omega$ is specified by a unit vector $\vec{n}(\theta,\varphi) = \vec{\iota}_r = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$ from the origin to that point. We let $|\vec{n}(\theta,\varphi)\rangle$ be the surface area Dirac delta on $\Omega$, so that for any function $\psi(\theta,\varphi)$ on $\Omega$:

$$\psi(\theta,\varphi) = \langle\vec{n}(\theta,\varphi)|\psi\rangle = \int_\Omega \delta(\cos\theta - \cos\theta')\delta(\varphi-\varphi')\psi(\theta',\varphi')d\Omega' \tag{1.7.79}$$

where $d\Omega' = \sin\theta d\theta' d\varphi'$. The $\delta(\cos\theta - \cos\theta')$ as opposed to $\delta(\theta-\theta')$ is required for the surface area Dirac delta to cancel the $\sin\theta$ factor in the surface area measure. Immediately, these deltas satisfy the orthonormality and completeness relations on $\mathbb{H}_\Omega$:

$$\langle\vec{n}(\theta,\varphi)|\vec{n}(\theta',\varphi')\rangle = \delta(\cos\theta-\cos\theta')\delta(\varphi-\varphi'), \qquad \int_\Omega |\vec{n}(\theta,\varphi)\rangle\langle\vec{n}(\theta,\varphi)|d\Omega = \mathbb{1} \tag{1.7.80}$$

Our angular momentum operators can be considered as operators on only $\mathbb{H}_\Omega$. We let $|l,m\rangle$ be the eigenbasis of angular momentum in that space:

$$\hat{L}^2|l,m\rangle = \hbar^2 l(l+1)|l,m\rangle, \qquad \hat{L}_z|l,m\rangle = \hbar m|l,m\rangle \tag{1.7.81}$$

and we let $Y_{lm}(\theta,\varphi)$, called the spherical harmonics, be the corresponding wavefunctions:

$$Y_{lm}(\theta,\varphi) = \langle\vec{n}(\theta,\varphi)|l,m\rangle, \quad \frac{\partial^2 Y_{lm}}{\partial\theta^2} + \frac{1}{\tan\theta}\frac{\partial Y_{lm}}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2 Y_{lm}}{\partial\varphi^2} + l(l+1)Y_{lm} = 0, \quad \frac{\partial Y_{lm}}{\partial\varphi} = imY_{lm}. \tag{1.7.82}$$

The last of these is easily solved to give the $\varphi$ dependence of the spherical harmonics as:

$$Y_{lm}(\theta,\varphi) = e^{im\varphi}P_{lm}(\theta) \tag{1.7.83}$$

for some $P_{lm}(\theta)$, called the associated Legendre functions. Equation (1.7.82) demands that the $Y_{lm}$ be differentiable, which first demands they be continuous. Because the points at $\varphi = 0$ versus $\varphi = 2\pi$ for fixed $\theta$ are identically the same, this requires $Y_{lm}(\theta,2\pi) = Y_{lm}(\theta,0)$. By (1.7.83), this demands $e^{2\pi im} = 1$, which holds if and only if $m$ is an integer. $m$ is an integer if and only if $l$ is, therefore **while half-integer angular momentum is perfectly allowed for spin, for orbital angular momentum only integer eigenvalues are allowed**. The orthogonality and completeness of the eigenbasis $|l,m\rangle$ then in components becomes:

$$\int Y^*_{lm}(\theta,\varphi)Y_{l'm'}(\theta,\varphi)d\Omega = \delta_{ll'}\delta_{mm'}, \quad \sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}(\theta,\varphi)Y^*_{lm}(\theta',\varphi') = \delta(\cos\theta-\cos\theta')\delta(\varphi-\varphi'). \tag{1.7.84}$$

And, (1.7.78) can be written more specifically, for some radial wavefunctions $R_{lm}(t,r)$, as:

$$\psi(t,\vec{x}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} R_{lm}(t,r)Y_{lm}(\theta,\varphi) \tag{1.7.85}$$

Having solved for the $\varphi$ dependence of $Y_{lm}$, (1.7.82) becomes:

$$\frac{\partial^2 P_{lm}}{\partial\theta^2} + \frac{1}{\tan\theta}\frac{\partial P_{lm}}{\partial\theta} + \left(l(l+1) - \frac{m^2}{\sin^2\theta}\right)P_{lm} = 0. \tag{1.7.86}$$

We can solve this differential equation for the $P_{lm}(\theta)$ by using our $\hat{L}_\pm$ operators. By definition:

$$0 = \langle \vec{n}(\theta,\varphi)|\hat{L}_+|l,l\rangle = i\hbar e^{i\varphi}\left(\frac{1}{\tan\theta}\frac{\partial Y_{ll}}{\partial\varphi} - i\frac{\partial Y_{ll}}{\partial\theta}\right) \implies \frac{\partial P_{ll}}{\partial\theta} = \frac{l}{\tan\theta}P_{ll} \implies P_{ll}(\theta) = N\sin^l\theta \quad (1.7.87)$$

for some normalization factor $N$, which we can set by using (1.7.84) with $l = l' = m = m'$. Doing so gives:

$$1 = |N|^2\int_0^{2\pi}d\varphi\int_0^{\pi}\sin^{2l+1}\theta d\theta \implies |N| = \frac{1}{2^l l!}\sqrt{\frac{(2l+1)!}{4\pi}} \quad (1.7.88)$$

were we have used an integral identity which is proven in problem 1.45

$$\int_0^{\pi}\sin^{2l+1}\theta d\theta = 2^{2l+1}\frac{l!^2}{(2l+1)!}. \quad (1.7.89)$$

We choose $N$ to be real and to have sign so that it is positive for even $l$ and negative for odd $l$, giving:

$$Y_{ll}(\theta,\varphi) = \frac{(-1)^l}{2^l l!}\sqrt{\frac{(2l+1)!}{4\pi}}e^{il\varphi}\sin^l\theta. \quad (1.7.90)$$

We motivate this alternating sign convention in 1.44. We can lower $m$ using (1.7.49):

$$Y_{l(m-1)} = \frac{1}{\hbar\sqrt{l(l+1)-m(m-1)}}i\hbar e^{-i\varphi}\left(\frac{1}{\tan\theta}\frac{\partial Y_{lm}}{\partial\varphi} + i\frac{\partial Y_{lm}}{\partial\theta}\right) \quad (1.7.91)$$

$$\implies P_{l(m-1)}(\theta) = -\frac{1}{\sqrt{l(l+1)-m(m-1)}}\left(\frac{\partial P_{lm}}{\partial\theta} + \frac{m}{\tan\theta}P_{lm}\right). \quad (1.7.92)$$

Starting from $P_{ll}(\theta)$, we can use equation this to recursively compute all of the $P_{lm}$ for $m < l$. In particular, using an integrating factor we can write the right hand side as:

$$P_{l(m-1)}(\theta) = -\frac{1}{\sqrt{l(l+1)-m(m-1)}}\frac{1}{\sin^m\theta}\frac{\partial}{\partial\theta}\left(\sin^m\theta P_{lm}\right). \quad (1.7.93)$$

Then, defining $x = \cos\theta$ and $Q_{lm}(x) = \sin^m(\theta)P_{lm}(\theta)$ this becomes:

$$Q_{l(m-1)}(x) = \frac{1}{\sqrt{l(l+1)-m(m-1)}}\frac{\partial Q_{lm}}{\partial x} \quad (1.7.94)$$

$$\implies Q_{lm}(x) = \prod_{n=m}^{l-1}\left(\frac{1}{\sqrt{l(l+1)-n(n-1)}}\right)\left(\frac{\partial}{\partial x}\right)^{l-m}Q_{ll}(x). \quad (1.7.95)$$

This last formula is gotten by applying (1.7.94) $l-m$ times. Renaming the product variable to $k = l-n$, we have:

$$\prod_{n=m}^{l-1}\left(\frac{1}{\sqrt{l(l+1)-n(n-1)}}\right) = \prod_{k=1}^{l-m}\frac{1}{\sqrt{k(2l+1-k)}} = \frac{1}{\sqrt{(l-m)!}}\sqrt{\frac{(l+m)!}{(2l)!}}. \quad (1.7.96)$$

Then, recalling $x = \cos\theta$ and simplifying (1.7.95) and (1.7.83):

$$\sin^m(\theta)P_{lm}(\theta) = Q_{lm}(x) = \sqrt{\frac{(l+m)!}{(l-m)!(2l)!}}\left(\frac{\partial}{\partial x}\right)^{l-m}\left(\frac{(-1)^l}{2^l l!}\sqrt{\frac{(2l+1)!}{4\pi}}(1-x^2)^l\right) \quad (1.7.97)$$

$$\implies \boxed{Y_{lm}(\theta,\varphi) = \frac{1}{2^l l!}\sqrt{\frac{2l+1}{4\pi}\frac{(l+m)!}{(l-m)!}}\frac{e^{im\varphi}}{\sin^m\theta}\left(\frac{\partial}{\partial\cos\theta}\right)^{l-m}(\cos^2\theta-1)^l.} \quad (1.7.98)$$

## Exercises

**Problem 1.42: Spin Precession.** Consider a quantum system with a spin $s$ wavefunction and, for a constant $\omega$ and constant unit vector $\vec{n}$, the Hamiltonian operator:

$$\hat{H} = \omega\vec{n}\cdot\hat{\vec{S}}.$$

(a) Show using Ehrenfest's theorem that the total spin angular momentum of the wavefunction satisfies:

$$\frac{d\vec{S}}{dt} = \omega\vec{n}\times\vec{S}.$$

(b) Solve this equation of motion to show that the spin evolves in time as:

$$\vec{S}(t) = \vec{S}(0)\cos\omega t + \vec{n}\vec{n}\cdot\vec{S}(0)(1-\cos\omega t) + \vec{n}\times\vec{S}(0)\sin\omega t.$$

Interpret this result geometrically.

**Problem 1.43: Spin-1/2 Rotations.** Recall the results of 1.4. By the logic leading to (1.7.65), for a spin $s$ wavefunction, a rotation by angle $\theta$ in the unit direction $\vec{n}$ is represented on the internal spin space by:

$$\hat{U} = e^{-\frac{i}{\hbar}\theta\vec{n}\cdot\hat{\vec{S}}}.$$

In this question we will specialize our interest to wavefunctions of spin $s = 1/2$.

(a) Let $\hat{S}_n = \vec{n}\cdot\hat{\vec{S}}$. Show that the eigenstates of $\hat{S}_n$ are given by:

$$|+\rangle = \begin{pmatrix} e^{-\frac{i\varphi}{2}}\cos\frac{\theta}{2} \\ e^{\frac{i\varphi}{2}}\sin\frac{\theta}{2}\end{pmatrix}, \qquad |-\rangle = \begin{pmatrix} -e^{-\frac{i\varphi}{2}}\sin\frac{\theta}{2} \\ e^{\frac{i\varphi}{2}}\cos\frac{\theta}{2}\end{pmatrix}, \qquad \hat{S}_n|\pm\rangle = \pm\frac{\hbar}{2}|\pm\rangle$$

where $\vec{n} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$ as usual in spherical coordinates.

(b) Show that the general rotation operator $\hat{U}$ on the internal space can be written explicitly as:

$$\hat{U} = \cos\left(\frac{\theta}{2}\right)\mathbb{1} - i\sin\left(\frac{\theta}{2}\right)\vec{n}\cdot\hat{\vec{\sigma}}$$

where $\hat{\vec{\sigma}}$ is the vector of Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0\end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0\end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1\end{pmatrix}.$$

**Problem 1.44: Spin-1 and electromagnetic polarization vectors.** A three dimensional vector rotates irreducibly and carries 3 independent components, thus it must correspond in spin representation to spin-1, as it is the irreducible representation of dimension 3. Suppose we have a spin-1 (vectorial) wavefunction.

(a) Explicitly compute the spin matrices for spin-1 using Theorem 1.11 to show that:

$$\begin{aligned}
\hat{S}_x =& \frac{\hbar}{\sqrt{2}}(|1,-1\rangle\langle 1,0| + |1,0\rangle\langle 1,-1| + |1,1\rangle\langle 1,0| + |1,0\rangle\langle 1,1|), \\
\hat{S}_y =& \frac{i\hbar}{\sqrt{2}}(|1,-1\rangle\langle 1,0| - |1,0\rangle\langle 1,-1| + |1,0\rangle\langle 1,1| - |1,1\rangle\langle 1,0|), \\
\hat{S}_z =& \hbar|1,1\rangle\langle 1,1| - \hbar|1,-1\rangle\langle 1,-1|.
\end{aligned}$$

(b) Show that each of $\hat{S}_x$, $\hat{S}_y$, $\hat{S}_z$ have an eigenvector with eigenvalue of 0. Explain that consequently any rotation on the internal space leaves a particular direction vector unchanged. Define $|x\rangle, |y\rangle, |z\rangle$ as the unit vectors in the directions left unchanged by $x, y, z$ rotations respectively. Explain why this means $|x\rangle$, $|y\rangle$, $|z\rangle$ must correspond geometrically to the $x, y, z$ coordinate directions. Show that we can choose the phases of $|x\rangle, |y\rangle, |z\rangle$ so that:

$$|x\rangle = \frac{1}{\sqrt{2}}(|1,-1\rangle + |1,-1\rangle), \qquad |y\rangle = \frac{i}{\sqrt{2}}(|1,1\rangle + |1,-1\rangle), \qquad |z\rangle = |1,0\rangle$$

and that these can be inverted to give:

$$|1,1\rangle = -\frac{|x\rangle + i|y\rangle}{\sqrt{2}}, \qquad |1,0\rangle = |z\rangle, \qquad |1,-1\rangle = \frac{|x\rangle - i|y\rangle}{\sqrt{2}}.$$

Let $\vec{\xi}_m = \xi^a_m \vec{\iota}_a$ be the electromagnetic helicity polarization vectors as in Orig. problems 4.38 and 4.42. Show:

$$\xi^a_m = \langle x^a|1,m\rangle.$$

Thus explain how the $|1,m\rangle$ states are precisely these polarization vectors in Dirac notation.

(c) Show that in terms of $|x\rangle, |y\rangle, |z\rangle$, the spin-1 matrices can be written as:

$$\hat{S}_x = i\hbar(|z\rangle\langle y| - |y\rangle\langle z|) \qquad \hat{S}_y = i\hbar(|x\rangle\langle z| - |z\rangle\langle x|) \qquad \hat{S}_z = i\hbar(|y\rangle\langle x| - |x\rangle\langle y|)$$

and show that using the $\hat{J}$ matrices originally defined in Orig. section 2.4 these are given simply by:

$$\hat{S}_a = i\hbar\hat{J}_a.$$

Show that the internal rotation matrix $\hat{U}$ for spin-1 can be identified precisely with our previous

definition of a $3 \times 3$ rotation matrix $\hat{R}$:

$$\hat{U} = e^{-\frac{i}{\hbar}\theta^a \hat{S}_a} = e^{\theta^a \hat{J}_a} = \hat{R}.$$

(d) Show that the $l = 1$ spherical harmonics can be written as:

$$Y_{1m}(\theta, \varphi) = \sqrt{\frac{3}{4\pi}} \vec{\iota}_r \cdot \vec{\xi}_m.$$

The $(-1)^l$ was included as a convention in (1.7.90) so that $Y_{00}$ is a positive and $Y_{1m}$ is positively proportional to $\vec{\iota}_r \cdot \vec{\xi}_m$.

**Problem 1.45: An Integral Identity.** For this question it is useful to recall the definitions of the Gamma function $\Gamma(x)$ and the Beta function $\mathrm{B}(x, y)$. For nonnegative integer $l$, consider the integral:

$$I(l) = \int_0^\pi \sin^{2l+1}\theta d\theta.$$

(a) Let $x = \cos^2\theta$ to show that the integral becomes:

$$I(l) = \int_0^1 x^{-\frac{1}{2}}(1-x)^l dx = \mathrm{B}\left(\frac{1}{2}, l+1\right) = \frac{\Gamma\left(\frac{1}{2}\right)\Gamma(l+1)}{\Gamma\left(l+\frac{3}{2}\right)}.$$

(b) Use that the Gamma function satisfies $\Gamma(x+1) = x\Gamma(x)$ and the definition of the factorial to show:

$$\Gamma\left(l+\frac{3}{2}\right) = \frac{\Gamma\left(\frac{1}{2}\right)}{2^{l+1}} \prod_{k=0}^{l}(2k+1), \quad \prod_{k=0}^{l}(2k+1) = \frac{(2l+1)!}{2^l l!} \quad \Longrightarrow \quad I(l) = 2^{2l+1}\frac{l!^2}{(2l+1)!}.$$

**Problem 1.46: The Schwinger Model.** Consider a quantum system with Hilbert space spanned by the orthonormal states labeled $|n_a, n_b\rangle$ where $n_a$ and $n_b$ each run from $0, 1, 2, 3, 4, ...$ which represent the eigenstates of two independent quantum harmonic oscillators with annihilation operators $\hat{a}$ and $\hat{b}$. That is:

$$[\hat{a}, \hat{a}^\dagger] = [\hat{b}, \hat{b}^\dagger] = 1, \quad [\hat{a}, \hat{b}] = [\hat{a}^\dagger, \hat{b}] = 0, \quad \hat{a}|n_a, n_b\rangle = \sqrt{n_a}|n_a - 1, n_b\rangle, \quad \hat{b}|n_a, n_b\rangle = \sqrt{n_b}|n_a, n_b - 1\rangle.$$

Define the operators:

$$\hat{J}_x = \frac{\hbar}{2}(\hat{a}^\dagger\hat{b} + \hat{b}^\dagger\hat{a}), \qquad \hat{J}_y = \frac{\hbar}{2i}(\hat{a}^\dagger\hat{b} - \hat{b}^\dagger\hat{a}), \qquad \hat{J}_z = \frac{\hbar}{2}(\hat{a}^\dagger\hat{a} - \hat{b}^\dagger\hat{b}), \qquad \hat{J} = \frac{\hbar}{2}(\hat{a}^\dagger\hat{a} + \hat{b}^\dagger\hat{b}).$$

(a) Show that $\hat{J}_x, \hat{J}_y, \hat{J}_z$ satisfy the su(2) algebra and that $\hat{J}$ satisfies:

$$[\hat{J}_a, \hat{J}_b] = i\hbar\epsilon_{ab}{}^c\hat{J}_c, \qquad [\hat{J}_a, \hat{J}] = 0, \qquad \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 = \hat{J}(\hat{J} + \hbar\mathbb{1}).$$

(b) Show using knowledge about the quantum harmonic oscillator that the eigenvalues of $\hat{J}$ come in half-integers (0, 1/2, 1, 3/2, ...) and that the eigenspace of $\hat{J}$ with eigenvalue $\hbar j$ forms an irreducible representation of the space of angular momentum $j$. Show explicitly that the $|j, m\rangle_{\mathrm{AM}}$ angular momentum states can be related to the $|n_a, n_b\rangle_{\mathrm{HO}}$ harmonic oscillator states by:

$$|j, m\rangle_{\mathrm{AM}} = |j + m, j - m\rangle_{\mathrm{HO}}.$$

(c) Show that this two harmonic oscillator model of angular momentum, called the Schwinger model, contains precisely one representation of every possible angular momentum.

**Problem 1.47: Central Potential Problem.** Consider a mass $M$ nonrelativistic scalar Schrödinger wave in 3 spatial dimensions subject to a spherically symmetric potential $U(r)$ so that it has Hamiltonian operator:

$$\hat{H} = \frac{\hat{\vec{p}}^2}{2M} + U(\hat{r}).$$

(a) Let $\hat{p}_r$ be the natural self-adjoint operator which is conjugate to the radial coordinate $\hat{r}$. Show that:

$$[\hat{r}, \hat{p}_r] = i\hbar\mathbb{1}, \qquad [\hat{p}_r, \hat{L}_a] = 0, \qquad \langle\vec{x}|\hat{p}_r|\psi\rangle = -i\hbar\left(\frac{\partial\psi}{\partial r} + \frac{\psi}{r}\right), \qquad \hat{\vec{p}}^2 = \hat{p}_r^2 + \frac{\hat{L}^2}{\hat{r}^2}.$$

(b) Show using the decomposition in (1.7.85) that the eigenfunctions of this Hamiltonian operator with eigenvalue $E$ can be written in the form:

$$\psi_{Elm}(\vec{x}) = R_{El}(r)Y_{lm}(\theta, \varphi)$$

where $Y_{lm}$ are the spherical harmonics and the radial functions $R_{El}$ satisfy:

$$\frac{\partial^2 R_{El}}{\partial r^2} + \frac{2}{r}\frac{\partial R_{El}}{\partial r} - \frac{l(l+1)}{r^2}R_{El} - \frac{2M}{\hbar^2}(U(r) - E)R_{El} = 0.$$

Thus, the spherically symmetric eigenvalue problem reduces to this radial ordinary differential equation.

**Problem 1.48: Eccentricity Vector in Classical Mechanics.** Let us study a purely classical mechanics system to develop insight for a corresponding quantum system that will be of interest. Consider a nonrelativistic point particle with mass $M$ evolving according to the Hamiltonian:

$$\mathcal{H} = \frac{|\vec{p}|^2}{2M} - \frac{K}{r}$$

for a positive constant $K$. Define the eccentricity vector $\vec{\varepsilon}$ and the Laplace-Runge-Lenz vector $\vec{D}$ by:

$$\vec{\varepsilon} = \frac{1}{MK}\vec{p} \times \vec{\mathrm{L}} - \frac{\vec{x}}{r}, \qquad \vec{D} = K\sqrt{\frac{M}{2|\mathcal{H}|}}\vec{\varepsilon}.$$

(a) Show that $\vec{\varepsilon}$ is conserved in time, that it is colinear with the vector from the point $r = 0$ to the point on the orbit at which $r$ is smallest, and that $|\vec{\varepsilon}|$ is the orbit's eccentricity.

(b) Show that:

$$\{D_a, \mathcal{H}\} = 0, \qquad \{D_a, D_b\} = -\mathrm{sgn}(\mathcal{H})\epsilon_{ab}{}^c \mathrm{L}_c, \qquad \{D_a, \mathrm{L}_b\} = \epsilon_{ab}{}^c D_c, \qquad \{\mathrm{L}_a, \mathrm{L}_b\} = \epsilon_{ab}{}^c \mathrm{L}_c.$$

(c) Explain that for unbound orbits, $D_a$ and $\mathrm{L}_a$ form precisely the Lie algebra so(3,1), the same algebra as that of the connected Lorentz group. Interpret this geometrically within the phase space of the system.

(d) Explain that for bound orbits, $D_a$ and $\mathrm{L}_a$ form precisely the Lie algebra so(4), the same algebra as that of rotations in 4 dimensions. Interpret this geometrically within the phase space of the system.

**Problem 1.49: ⋆ Solving the Hydrogen Atom.** A hydrogen atom consists of a proton of positive charge $q_e$ and mass $m_p$ and electron of negative charge $-q_e$ and mass $m_e$. (The proton charge magnitude is by all experiments equal to the electron charge magnitude. This enormous "coincidence" will get more attention in later chapters.) The electron-proton separation is very large compared to the size of either and the proton mass is about 2000 the electron mass. So, in the neighborhood of the electron the electric field produced by the proton is approximately that of a point charge $q_e$ (using the exterior multipole expansion) and we can approximate the proton as not moving.

(a) Explain based on the above why we can approximate the electron's Hamiltonian as:

$$\hat{H} = \frac{\hat{\vec{p}}^2}{2m_e} - \frac{q_e^2}{4\pi\epsilon_0 \hat{r}}.$$

For the rest of this question let $M = m_e$ and $K = \frac{q_e^2}{4\pi\epsilon_0}$. (Like in classical mechanics, the reduced mass $\frac{m_e m_p}{m_e + m_p}$ would give a better approximation than $m_e$.) Our goal is to find all of the eigenstates $\hat{H}$ for bound states of the system (states in which the electron stays finitely close to the proton).

(b) Using the radial equation from problem 1.47, show that bound orbits must have energy eigenvalue $E < 0$ as for $E > 0$ the solutions become asymptotically plane waves (and so are not spatially bounded; the electron extends arbitrarily far from the proton). The magnitude of the energy of a bound state is called the binding energy $\Phi$ of that state. The unbound "eigenstates" are actually generalized eigenstates in the rigged Hilbert space, in the sense that they are not square-integrable. They have take any positive energy as an eigenvalue. For the rest of this problem, we consider only bound states.

(c) Define the eccentricity vector operator $\hat{\vec{\varepsilon}}$ and $\hat{\vec{D}}$ exactly analogously to the classical quantities:

$$\hat{\vec{\varepsilon}} = \frac{1}{2MK}(\hat{\vec{p}} \times \hat{\vec{L}} - \hat{\vec{L}} \times \hat{\vec{p}}) - \frac{\hat{\vec{x}}}{\hat{r}}, \qquad \hat{\vec{D}} = K\sqrt{\frac{M}{2|\hat{H}|}}\hat{\vec{\varepsilon}}.$$

Show that these operators are self-adjoint and that:

$$\hat{\vec{\varepsilon}}^2 = \mathbb{1} + \frac{2\hat{H}}{MK^2}(\hat{\vec{L}}^2 + \hbar^2 \mathbb{1}), \qquad [\hat{\vec{D}}, \hat{H}] = 0, \qquad \hat{\vec{L}} \cdot \hat{\vec{D}} = \hat{\vec{D}} \cdot \hat{\vec{L}} = 0.$$

(d) Let:

$$\hat{\vec{B}}_\pm = \frac{\hat{\vec{L}} \pm \hat{\vec{D}}}{2}, \qquad \hat{C}_+ = \hat{\vec{B}}_+^2 + \hat{\vec{B}}_-^2, \qquad \hat{C}_- = \hat{\vec{B}}_+^2 - \hat{\vec{B}}_-^2.$$

Show that:

$$[\hat{\vec{B}}_\pm, \hat{H}] = 0, \qquad [\hat{B}_{+a}, \hat{B}_{+b}] = i\hbar\epsilon_{ab}{}^c \hat{B}_{+c}, \qquad [\hat{B}_{+a}, \hat{B}_{-b}] = 0, \qquad [\hat{B}_{-a}, \hat{B}_{-b}] = i\hbar\epsilon_{ab}{}^c \hat{B}_{-c}$$

and that on the genuine Hilbert space, which consists only of bound states, there exist $|\psi\rangle$ which are simultaneous eigenstates of $\hat{H}$, $\hat{C}_+$, $\hat{L}^2$, $\hat{L}_z$ and that these satisfy, for $b$ an arbitrary half-integer:

$$\hat{C}_- = 0, \qquad [\hat{C}_+, \hat{H}] = [\hat{C}_+, \hat{L}_a] = 0, \qquad \mathbb{1} + \frac{2\hat{H}}{MK^2}(\hbar^2 \mathbb{1} + 2\hat{C}_+) = 0, \qquad \hat{C}_+|\psi\rangle = 2\hbar^2 b(b+1)|\psi\rangle$$

(e) Consider simultaneous eigenstates of $\hat{H}, \hat{C}_+, \hat{L}^2, \hat{L}_z$. Let $n = 2b + 1$ be an arbitrary positive integer with $b$ as previously. Show that the energy eigenvalues of the hydrogen atom are given by:

$$\boxed{E_n = -\frac{MK^2}{2\hbar^2 n^2} = -\frac{m_e q_e^4}{32\pi^2 \epsilon_0^2 \hbar^2 n^2}}$$

and that the total orbital angular momentum eigenvalue $l$ must satisfy $l < n$.

(f) Show that the radial part of the wavefunction $R_{nl}$ for eigenstates of $\hat{H}$ satisfies:

$$\frac{\partial^2 R_{nl}}{\partial r^2} + \frac{2}{r}\frac{\partial R_{nl}}{\partial r} - \frac{l(l+1)}{r^2}R_{nl} + \frac{2}{a_0 r}R_{nl} - \frac{1}{n^2 a_0^2}R_{nl} = 0 \qquad \text{where:} \qquad a_0 = \frac{\hbar^2}{MK} = \frac{4\pi\epsilon_0\hbar^2}{q_e^2 m_e}$$

known as the Bohr radius is the characteristic length scale of the hydrogen atom. Experimentally, its value is $a_0 = 5.29177210903(80) \times 10^{-11}\text{m}$ [37]. While we have established the values of $q_e$, $\epsilon_0$, and $m_e$ by previous experiments or choices of units, in order to calculate $a_0$ we need the value of $\hbar$, which we will determine in the next problem.

(g) The radial equation is still quite nasty. Change variables to $x = \frac{2}{na_0}r$ and define $f_{nl}(x)$ so that:

$$R_{nl}(r) = \left(\frac{2r}{na_0}\right)^l e^{-\frac{r}{na_0}} f_{nl}\left(\frac{2r}{na_0}\right).$$

Show that $f_{nl}(x)$ satisfies:

$$x f''_{nl}(x) + (2l + 2 - x) f'_{nl}(x) + (n - l - 1) f_{nl}(x) = 0.$$

(h) The differential equation satisfied by $f_{nl}$ can be put in a more standard mathematical form by defining $\alpha = 2l + 1$ and $k = n - l - 1$. Then, let $f(x)$ be a generic function satisfying the differential:

$$x f''(x) + (\alpha + 1 - x) f'(x) + k f(x) = 0.$$

This can be solved by the Frobenius method. Near $x = 0$, $x^\beta f(x)$ is guaranteed to have a well defined Taylor series for sufficiently large $\beta$. Explain why, by returning to the original wavefunction, that for a physically reasonable bound (i.e. square-integrable) solution we must have $\beta \leq l + 1$. Using this, show by inserting the Taylor series for $x^\beta f(x)$ into the differential equation that $\beta = 0$ and that:

$$f(x) = \sum_{j=0}^{\infty} c_j x^j, \qquad c_{j+1} = -\frac{k - j}{(j+1)(j+\alpha+1)} c_j.$$

(i) Show that the $c_j$ are solved by:

$$c_j = \frac{(-1)^j}{j!}\binom{k+\alpha}{k-j} c_0$$

and that $\frac{f(x)}{c_0}$ is a polynomial (known as the generalized Laguerre polynomial of $L_k^{(\alpha)}(x)$) given by:

$$\frac{f(x)}{c_0} = L_k^{(\alpha)}(x) = \sum_{j=0}^{k} \frac{(-1)^j}{j!}\binom{k+\alpha}{k-j} x^j = \frac{x^{-\alpha}e^x}{k!}\left(\frac{\partial}{\partial x}\right)^k (e^{-x}x^{\alpha+k}) = \frac{x^{-\alpha}}{k!}\left(\frac{\partial}{\partial x} - 1\right)^k x^{\alpha+k}.$$

(j) Choose $c_0$ to normalize the $\psi$ to finally show that the radial solution to the Hydrogen atom is given by:

$$R_{nl}(r) = \left(\frac{2}{na_0}\right)^{\frac{3}{2}} \sqrt{\frac{(n-l-1)!}{2n(n+l)!}} \left(\frac{2r}{na_0}\right)^l L_{n-l-1}^{(2l+1)}\left(\frac{2r}{na_0}\right) e^{-\frac{r}{na_0}}.$$

(k) Show that each hydrogen energy level (determined by $n$) has degeneracy $n^2$ and that that, for some initial conditions $C_{nlm}$, any of its bound state solutions can be written as:

$$\psi(t,\vec{x}) = \sum_{n=1}^{\infty}\sum_{l=0}^{n-1}\sum_{m=-l}^{l} C_{nlm}e^{-\frac{i}{\hbar}E_n t}R_{nl}(r)Y_{lm}(\theta,\varphi)$$

**Problem 1.50: ⋆ Atomic Spectral Lines and Measuring $\hbar$.** This problem is based on chapter 9 of [30]. Consider an electron in an atom, modeled so that the electron has wavefunction $\psi(t,\vec{x})$ with Hamiltonian operator $\hat{H}_0$. This Hamiltonian operator will take the form:

$$\hat{H}_0 = \frac{\hat{\vec{p}}^2}{2m} + \hat{U}(\hat{\vec{x}})$$

where $U(\vec{x})$ is some potential which encodes the atomic structure. The particular details of this potential will be generally irrelevant for this question. We assume that we know the normalized eigenstates $|n\rangle$ of $\hat{H}_0$ and that they have (possibly degenerate) eigenvalues $E_n$. Now consider this electron to be perturbed by an incoming electromagnetic plane wave. Choose the gauge of the plane wave so that it takes the form:

$$\phi = 0, \qquad \vec{A} = \epsilon \mathrm{Re}\left(\vec{\xi}e^{i(\vec{k}\cdot\vec{x}-\omega t)}\right)$$

where $\epsilon$ is a constant encoding the strength of the plane wave, $\vec{\xi}$ is a complex unit vector encoding its polarization with $\vec{k}\cdot\vec{\xi} = 0$, and $\omega = c|\vec{k}|$ due to Maxwell's equations in free space.

(a) Show that the perturbed electron now satisfies the Schrödinger equation with a new Hamiltonian operator $\hat{H}$ which can be put in the form:

$$\hat{H} = \hat{H}_0 + \epsilon\hat{H}_1 + \mathcal{O}(\epsilon^2) \quad \text{where:} \quad \hat{H}_1 = -\frac{q}{2m}e^{i\vec{k}\cdot\hat{\vec{x}}-i\omega t}\vec{\xi}\cdot\hat{\vec{p}} - \frac{q}{2m}e^{-i\vec{k}\cdot\hat{\vec{x}}+i\omega t}\vec{\xi}^*\cdot\hat{\vec{p}}.$$

(b) Assume that the electromagnetic plane-wave has wavelength much longer than the atomic size. Show that this allows us to approximate $\hat{H}_1$ as:

$$\hat{H}_1 = -\frac{q}{m}\mathrm{Re}\left(e^{-i\omega t}\vec{\xi}\right)\cdot\hat{\vec{p}}.$$

Show that for the hydrogen atom this approximation is valid for $\omega$ such that:

$$\omega \ll \frac{q_e^2 m_e c}{2\epsilon_0\hbar^2}.$$

(c) While we have not yet discussed precisely how to apply thermodynamics and statistical mechanics to quantum mechanics, you should be unsurprised to learn that in thermodynamic equilibrium the distribution of states of a statistical sample of atoms will be proportional to $e^{-\frac{\mathrm{E}}{T}}$ where E is the energy of that state and $T$ is the temperature. If there is a significant gap between the ground state (the lowest energy state) energy eigenvalue of the atom and the energy eigenvalue of the next lowest state (called the first excited state) compared to the temperature of the system, the atoms will overwhelmingly be in the ground state. For now, we will simply assume this to be true (physically it is very accurate for ordinary atoms at or below ordinary temperatures[11]). Let $|g\rangle$ be the ground state of the atom and let it have energy eigenvalue $E_g$. Assume that the atom is initially in the ground state before being perturbed by the electromagnetic field. Show that the solution to the Schrödinger equation has the form:

$$|\psi(t)\rangle = e^{-i\frac{E_g}{\hbar}t}|g\rangle + \epsilon|\delta\psi(t)\rangle + \mathcal{O}(\epsilon^2)$$

where $|\delta\psi(t)\rangle$ evolves according to:

$$i\hbar\frac{d|\delta\psi\rangle}{dt} = \hat{H}_0|\delta\psi\rangle - \frac{q}{m}e^{-i\frac{E_g}{\hbar}t}\mathrm{Re}\left(e^{-i\omega t}\vec{\xi}\right)\cdot\hat{\vec{p}}|g\rangle.$$

Solve this differential equation to show that $|\delta\psi(t)\rangle$ is given explicitly by:

$$|\delta\psi(t)\rangle = \frac{q}{2m}e^{-i\frac{E_g}{\hbar}t}\sum_n\left(\frac{e^{-i\omega t}\vec{\xi}}{E_n - E_g - \hbar\omega} + \frac{e^{i\omega t}\vec{\xi}^*}{E_n - E_g + \hbar\omega}\right)\cdot\langle n|\hat{\vec{p}}|g\rangle|n\rangle.$$

[11] Those who do know how to treat this with statistical mechanics may object that the partition function of the hydrogen atom is divergent. This divergence is due to the very slow fall off of the Coulomb interaction and can be removed as a problem by cutting off the involved integrals at a finite volume, computing physical quantities from the cutoff results, and then taking the cutoff to infinity at the end. See [5]

(d) Using (1.6.83), compute the integral of the current density $\vec{J}$ over all space. Show that:

$$\begin{aligned}\int_{\mathbb{S}} \vec{J}(t,\vec{x})d^3\vec{x} &= \frac{q}{m}\langle\psi|(\hat{\vec{p}} - q\hat{\vec{A}})|\psi\rangle \\ &= \frac{q}{m}\langle g|\hat{\vec{p}}|g\rangle + \frac{q^2\epsilon}{m^2}\text{Re}\left(\sum_n \left(\frac{e^{-i\omega t}\vec{\xi}}{E_n - E_g - \hbar\omega} + \frac{e^{i\omega t}\vec{\xi}^*}{E_n - E_g + \hbar\omega}\right)\cdot\langle n|\hat{\vec{p}}|g\rangle\langle g|\hat{\vec{p}}|n\rangle\right) \\ &\quad - \frac{q^2\epsilon}{m}\text{Re}\left(\vec{\xi}e^{-i\omega t}\right) + \mathcal{O}(\epsilon^2).\end{aligned}$$

Using the results of Orig. problems 4.53 and 4.54, show that (in the fully nonrelativistic limit), the electromagnetic radiation vector $\vec{\alpha}(t,\theta,\varphi)$ becomes:

$$\vec{\alpha} = \frac{q^2\omega\epsilon}{m^2}\text{Im}\left(\sum_n \left(\frac{e^{-i\omega t}\vec{\xi}}{E_n - E_g - \hbar\omega} - \frac{e^{i\omega t}\vec{\xi}^*}{E_n - E_g + \hbar\omega}\right)\cdot\langle n|\hat{\vec{p}}|g\rangle\langle g|\hat{\vec{p}}|n\rangle\right) - \frac{q^2\omega\epsilon}{m}\text{Im}\left(\vec{\xi}e^{-i\omega t}\right).$$

(e) Show using the result of Orig. problem 4.57 that the total electromagnetic scattering cross section of light off of the atom is given by:

$$\sigma = \frac{\mu_0^2 q^4}{12\pi m^2}\text{Tr}(\perp \hat{K}^2(\omega))$$

where $\perp$ is the $3\times 3$ matrix that projects out components of a vector in the direction of the incoming wavenumber $\vec{k}$, i.e. $\perp^{ab} = \delta^a b - \iota_k^a \iota_k^b$ where $\vec{\iota}_k = \frac{\vec{k}}{|\vec{k}|}$, and $\hat{K}$ is the matrix with components:

$$K_{ab}(\omega) = \delta_{ab} - \frac{1}{m}\sum_n \left(\frac{\langle g|\hat{p}_a|n\rangle\langle n|\hat{p}_b|g\rangle}{E_n - E_g - \hbar\omega} + \frac{\langle g|\hat{p}_b|n\rangle\langle n|\hat{p}_a|g\rangle}{E_n - E_g + \hbar\omega}\right).$$

Notice that the cross section becomes very large whenever $E_g + \hbar\omega$ approaches an eigenvalue of $\hat{H}_0$ (due to the denominator $E_n - E_g - \hbar\omega$ in the definition of $\hat{K}$). Notice further that because $E_g$ is the lowest eigenvalue of $\hat{H}_0$ and $\omega > 0$ by definition, the denominator $E_n - E_g + \hbar\omega$ never gets near 0 and so never contributes the same peak behavior that the $E_n - E_g - \hbar\omega$ denominator does. Because of this, when the incoming beam of light on the atom is tuned to have appropriate frequency, a stimulated emission of light from the atom will occur, which is confirmed by experiment. That significantly increased response of the atom thus occurs at $\hbar\omega = E_n - E_g$ for some $n$ and gives a direct probe of the atomic energy eigenvalues and a way to measure the value of $\hbar$. These peaks in the emitted light of stimulated atoms are called the atomic spectral lines. They provide a direct access point to the (relative to the ground state) energy eigenvalues of an atomic system. Note that this amplified response at discrete separated values has nothing to do with any magical fundamental quantum discreteness. All results are derived from the fully continuous equations of motion.

(f) Suppose we target hydrogen atoms with our light. Let $\omega_n$ be the angular frequency at which peak emission occurs for transitions from the hydrogen ground state with energy eigenvalue $E_g = E_1$ to a state of hydrogen energy eigenvalue $E_n$. Show that the value of $\hbar$ is given by:

$$\hbar = \left(\frac{m_e q_e^4}{32\pi^2\epsilon_0^2}\frac{1 - \frac{1}{n^2}}{\omega_n}\right)^{\frac{1}{3}}$$

and that by scanning over a range of such $n$ allows the accurate measurement of the value of $\hbar$. In SI units the value of $\hbar$ is exactly [15]:

$$\hbar = \frac{6.62607015}{2\pi}\times 10^{-34}\frac{\text{kg m}^2}{\text{s}}$$

(The value of $\hbar$ is exact because as of 2019 SI units defines the unit of the kilogram so that $\hbar$ takes this value. Historical SI units chose a particular arbitrary object as having a mass of a kilogram, in terms of which the value of $\hbar$ could be measured.) Explain how alternatively using natural units in which $\hbar = \epsilon_0 = 1$, this same experiment provides a measurement of the electron mass in natural units.

(g) Use the given value of $\hbar$ to check the value of the Bohr radius given in problem 1.49.

(h) Show using the formula for the hydrogen spectrum that the hydrogen ground state energy is [37]:

$$E_1 = -13.605693122994(26)\text{ eV}.$$

Compute the energy of the first excited state of hydrogen and show that at any reasonable temperature our assumption that the hydrogen is in the ground state is valid.

**Problem 1.51: ⋆ Photoelectric Effect.** This problem is based on [29]. When we solved the hydrogen atom previously, we only considered states in which the electron was bound to the proton. Consequently, the electron wavefunction decayed quickly at large distances. It is also possible to consider unbound states in which the electron becomes arbitrarily far away from the proton.

(a) Consider the setup of the the previous problem with the electron in the atom beginning in its ground state $|g\rangle$ with energy $E_g = -\Phi$ subject to an incoming electromagnetic plane wave but also subject to a uniform voltage $V$. Generalize the result of Fermi's Golden Rule from problem 1.19 by considering the time dependent perturbation found in the previous problem to show that to lowest order in $\varepsilon$ the transition rate for the electron from $|g\rangle$ to a different eigenstate $|n\rangle$ due to an unpolarized light beam is:

$$\Gamma_n = \frac{\pi q^2 \epsilon^2}{4\hbar m^2}\delta(E_n + \Phi - \hbar\omega)\langle n|\hat{\vec{p}}|g\rangle \perp \langle g|\hat{\vec{p}}|n\rangle.$$

(b) Let $\Gamma$ be the total transition rate for electrons to go from the ground state to an unbound state and suppose that in the unbound region the uniform potential has dropped off to zero. Let $T(E, V) = |t|^2$ where $t$ is the transmission coefficient for an unbound electron in the region of voltage $V$ to escape to the region of voltage 0 so that its final energy is $E$. Show that $\Gamma$ is given by:

$$\Gamma = \int_{\mathbb{R}^3} \frac{\pi q^2 \epsilon^2}{4\hbar m^2} T(E(\vec{p}), V)\delta(E(\vec{p}) + \Phi - qV - \hbar\omega)\langle\psi_{\vec{p}}|\hat{\vec{p}}|g\rangle \perp \langle g|\hat{\vec{p}}|\psi_{\vec{p}}\rangle d^3\vec{p}$$

where $E(\vec{p}) = \frac{|\vec{p}|^2}{2m}$ and $|\psi(\vec{p})\rangle$ is the unbound atomic eigenstate which asymptotically at large distances becomes a conventionally normalized eigenvector of the momentum operator $\hat{\vec{p}}$.

(c) Suppose that an unpolarized beam of light is shone on a material so that effectively $N$ electrons in a material are subject to the perturbation and suppose that the electrons which escape due to the light (called photoelectrons) travel toward a conducting wire so that they produce an electrical current $I$ which can be measured. Show that the electrical current is given by:

$$I = qN\Gamma.$$

Notice that the current is proportional to the intensity of the incoming light and that there is a critical frequency $\omega_{\min} = \frac{\Phi}{\hbar}$ below which no current is produced at $V = 0$. Now suppose we adjust the voltage $V$ until the electrical current drops to 0. The voltage at which this occurs is called the stopping voltage $V_{\text{stop}}$. Show that the stopping voltage is given by:

$$V_{\text{stop}} = \frac{\hbar\omega - \Phi}{q_e}.$$

Explain that given that we know the electron charge magnitude from previous experiments, by measuring the stopping voltage across many different frequencies of incoming light we can measure the values of $\hbar$ and the binding energy $\Phi$ of our material by plotting $V_{\text{stop}}$ versus $\omega$. Notice that the stopping voltage depends only on the frequency of the light and not at all on its intensity[12].

**Problem 1.52: ⋆ Fine-Structure Constant.** Let $\mathrm{E}_e = m_e c^2$ be the electron rest energy (using the relativistic energy zero-point).

(a) Using the previously established value of the electron mass, show that $\mathrm{E}_e$ is given by [37]:

$$\mathrm{E}_e = 510.99895000(15)\ \text{keV}$$

(keV = kiloelectronvolts = $10^3$ eV). This is the value of the electron mass in natural units.

(b) Show that the energy eigenvalues of the hydrogen atom can be expressed as:

$$E_n = -\frac{\alpha^2 \mathrm{E}_e}{2n^2}$$

where $\alpha$, known as the **fine-structure constant**, is a dimensionless physical constant defined by:

$$\alpha = \frac{q_e^2}{4\pi\epsilon_0\hbar c}.$$

(c) Explain from all three perspectives of historical SI units in which $q_e, \epsilon_0, \hbar, c$ are measured, modern SI in

[12] Readers who have heard introductory explanations of the photoelectric effect before are likely surprised to see an explanation of the effect without any reference to the concept of a photon. This is because the photon concept has nothing to do with the photoelectric effect; the common confusion to the contrary is a historical accident.

which $\epsilon_0$ is measured and $q_e, \hbar, c$ are defined by exact values, and natural units in which $q_e$ is measured and $\epsilon_0 = \hbar = c = 1$ that by combining the results of previously described experiments we can measure the value of $\alpha$. Show using the established values of $q_e, \epsilon_0, \hbar, c$ that $\alpha$ takes the value [37]:

$$\alpha = 0.0072973525693(11).$$

From this show that the magnitude of the electron charge in natural units, more generally called the **elementary charge** $\mathsf{e}$ is:

$$\mathsf{e} = \sqrt{4\pi\alpha} = 0.302822120872(22).$$

$\alpha$ or equivalently $\mathsf{e}$ is a dimensionless fundamental parameter of our universe; it cannot be gotten rid of by any choice of unit system. It is determined only by experiment. If you found a theory that could derive the value of $\alpha$ mathematically, you would have an instant Nobel prize. The elementary charge (also called the electromagnetic coupling constant) $\mathsf{e}$ is always unitless. It is related to the magnitude of the electron charge $q_e$ in SI units by:

$$q_e = \mathsf{e}\sqrt{\epsilon_0 \hbar c}.$$

Because the proton has exactly charge $q_e$, the neutron has exactly charge 0, and the electron has exactly charge $-q_e$ and all atomic matter is comprised of these constituents, the electric charge $q$ of any object consisting of these particles is of the form:

$$q = Qq_e = Q\mathsf{e}\sqrt{\epsilon_0 \hbar c}$$

where $Q$ is an **integer** known as the "charge number" of the object (we will see in later chapter that particles other than protons, neutrons, and electrons exist, but even for those particles $3Q$ is experimental always an integer). Essentially, $Q$ is the charge of the object "in units of $q_e$". Unfortunately, myself and other physicists will lazily call $q$, $\mathsf{e}$, or $Q$ simply "charge" from time to time, and one has to understand which is intended from context.

**Problem 1.53: Legendre Polynomials.** Often problems have an axis of symmetry, and then there is a natural reduction of the spherical harmonics, which are a basis of functions of $(\theta, \varphi)$, to the so-called Legendre polynomials $P_l(\cos\theta)$, which are a basis of functions of $\theta$.

(a) Consider an azimuthally symmetric function on $\Omega$, $\psi(\theta)$. Show, using (1.7.84), that:

$$\psi(\theta) = \sum_{l=0}^{\infty} \frac{2l+1}{2} \psi_l P_l(\cos\theta) \quad \text{where:} \quad \psi_l = \int_0^1 \psi(\theta) P_l(\cos\theta) \sin\theta d\theta, \qquad P_l(\cos\theta) = \sqrt{\frac{4\pi}{2l+1}} Y_{l0}(\theta, \varphi).$$

(b) Explain why the $P_l(x)$ are naturally viewed as functions on the interval $[-1, 1]$, that they are polynomials of degree $l$ which are even/odd for even/odd $l$, that $P_l(1) = 1$, that $P_0(x) = 1$, and that they have explicit form and orthogonal and completeness relations:

$$P_l(x) = \frac{1}{2^l l!} \left(\frac{d}{dx}\right)^l (x^2 - 1)^l, \quad \int_{-1}^{1} P_l(x) P_{l'}(x) dx = \frac{2}{2l+1} \delta_{ll'}, \quad \sum_{l=0}^{\infty} \frac{2l+1}{2} P_l(x) P_l(x') = \delta(x - x').$$

(c) Using their explicit formula, show that the Legendre polynomials satisfy the recurrence relation:

$$(l+1)P_{l+1}(x) + lP_{l-1}(x) - (2l+1)xP_l(x) = 0.$$

(d) With $f(t, x)$ as defined below, show from the above recurrence relation that $f$ satisfies the shown differential equation and has the solution:

$$f(t,x) = \sum_{l=0}^{\infty} P_l(x) t^l \implies (1 - 2xt + t^2)\frac{\partial f}{\partial t} + (t - x)f = 0 \implies f(t,x) = \frac{1}{\sqrt{1 - 2xt + t^2}}.$$

This $f(t, x)$ is called the generating function function for the Legendre polynomials, as its derivatives at $t = 0$ contain the full list of them.

**Problem 1.54: Spherical Harmonic Projection Theorem.** For spherical parameters $(\theta, \varphi)$ and $(\theta', \varphi')$, let $\vec{\imath} = \vec{n}(\theta, \varphi)$ and $\vec{\imath}' = \vec{n}(\theta', \varphi')$ be the corresponding points on the unit sphere. Let $\hat{\pi}_l$ be the operator which projects functions on $\mathbb{H}_\Omega$ into the subspace of functions for which $\hat{L}^2$ has eigenvalue $\hbar^2 l(l+1)$.

(a) Show that:

$$\hat{\pi}_l\hat{\pi}_{l'} = \delta_{ll'}\hat{\pi}_l, \qquad \sum_{l=0}^{\infty}\hat{\pi}_l = \mathbb{1}, \qquad \hat{\pi}_l = \sum_{m=-l}^{l}|l,m\rangle\langle l,m|, \qquad [\hat{\pi}_l, \hat{L}_a] = 0,$$

$$\pi_l(\theta,\varphi,\theta',\varphi') \overset{\text{def}}{=} \langle\vec{n}(\theta,\varphi)|\hat{\pi}_l|\vec{n}(\theta',\varphi')\rangle = \sum_{m=-l}^{l} Y_{lm}(\theta,\varphi)Y^*_{lm}(\theta',\varphi').$$

(b) Show, using either (1.7.98) or the continuity of the $Y_{lm}$ as functions, that $Y_{lm}(0,\varphi) = 0$ for all $m \neq 0$.

(c) Explain why $\hat{\pi}_l$ is rotationally invariant, i.e. for $\hat{U}$ the representation of any rotation, $\hat{U}\hat{\pi}_l\hat{U}^\dagger = \hat{\pi}_l$. Using this, explain why necessarily $\pi_l(\theta,\varphi,\theta',\varphi') = F_l(\vec{\imath}\cdot\vec{\imath}')$ for some function $F_l$. By considering $\vec{\imath}' = \vec{\imath}_z$, explain that $\pi_l(\theta,\varphi,0,\varphi') = F_l(\cos\theta)$ and show using the previous results and the definition of the Legendre polynomials that therefore:

$$P_l(\vec{\imath}\cdot\vec{\imath}') = \frac{4\pi}{2l+1}\sum_{m=-l}^{l} Y_{lm}(\theta,\varphi)Y^*_{lm}(\theta',\varphi').$$

**Problem 1.55: Spherical Harmonic Conjugation.**

(a) Explain why it is natural to define the Dirac delta basis $|\vec{n}\rangle$ to be real: $|\vec{n}\rangle^* = |\vec{n}\rangle$ and show that with this definition, $|l,0\rangle^* = |l,0\rangle$ because $Y^*_{l0} = Y_{l0}$.

(b) Show that the raising and lowering operators are given in components as below, and by conjugating that expression show that they satisfy:

$$\langle\vec{n}|\hat{L}_\pm|\psi\rangle = \hbar e^{\pm i\varphi}\left(\pm\frac{\partial\psi}{\partial\theta} + \frac{i}{\tan\theta}\frac{\partial\psi}{\partial\varphi}\right) \implies \hat{L}^*_\pm = -\hat{L}_\mp.$$

(c) By iteratively applying $\hat{L}_\pm$ to $|l,0\rangle$, show using the above that the spherical harmonics satisfy:

$$Y^*_{lm} = (-1)^m Y_{l(-m)}.$$

**Problem 1.56: Spherical Harmonics and Symmetric Traceless Tensors.** Let $\psi(\theta,\varphi)$ be a function in $\mathbb{H}_\Omega$. This function could be written in terms of the components of $\vec{n}$. By the Stone-Weierstrass theorem and a little analysis, because the unit sphere is compact, any such $\psi$ can be written as a power series[13] in the components of $\vec{n}$.

(a) Show, using that $\vec{n}\cdot\vec{n} = 1$, that we can take the coefficients of expansion of $\psi$ to be fully symmetric and traceless between any pair of indices:

$$\psi(\theta,\varphi) = \sum_{l=0}^{\infty}\psi_l^{a_1\ldots a_l}n_{a_1}(\theta,\varphi)\ldots n_{a_l}(\theta,\varphi), \qquad \psi_l^{a_1\ldots a_l} = \psi_l^{(a_1\ldots a_l)}, \qquad \psi_l^{a_1\ldots a_l}\delta_{a_{l-1}a_l} = 0.$$

(b) Show by explicit calculation that:

$$\langle\vec{n}|\hat{L}^2|\psi\rangle = \sum_{l=0}^{\infty}\hbar^2 l(l+1)\psi_l^{a_1\ldots a_l}n_{a_1}(\theta,\varphi)\ldots n_{a_l}(\theta,\varphi).$$

Explain based off of the above that the term for each $l$ in the expansion of $\psi$ is in the eigenspace of $\mathbb{H}_\Omega$ on which $\hat{L}^2$ has eigenvalue $\hbar^2 l(l+1)$. Explain why this means each such term can be written as a sum over spherical harmonics with that value of $l$. Show that consistent with this, there are $2l+1$ linearly independent symmetric traceless tensors of rank $l$.

(c) Let ST denote the symmetric traceless part of a tensor defined by first symmetrizing over all of its indices and then symmetrically adding terms that are products of Kronecker deltas and lower order tensors until the result is traceless in all pairs of indices. Based off of the above results and the consideration of having the correct $\varphi$ dependence, show that for some constants $N_{lm}$:

$$Y_{lm}(\theta,\varphi) = N_{lm}\mathrm{ST}(\xi^{a_1}_{\mathrm{sgn}(m)}\ldots\xi^{a_{|m|}}_{\mathrm{sgn}(m)}\xi_0^{a_{|m|+1}}\ldots\xi_0^{a_l})n_{a_1}\ldots n_{a_l}$$

with $\vec{\xi}_h$ the polarization vectors as defined in Orig. problems 4.38 and 4.42.

(d) For $m < 0$, consider the highest power of $\cos\theta$ that occurs in the above right hand side and in $Y_{lm}$ as

[13] Up to differing on sets of measure 0.

given by (1.7.98). By matching the coefficients of the two and using the result of 1.55(c), demonstrate:

$$Y_{lm}(\theta,\varphi)=\frac{(2l)!}{2^l l!}\sqrt{\frac{2l+1}{4\pi}}\sqrt{\frac{2^{|m|}}{(l+m)!(l-m)!}}\mathrm{ST}(\xi^{a_1}_{\mathrm{sgn}(m)}...\xi^{a_{|m|}}_{\mathrm{sgn}(m)}\xi_0^{a_{|m|+1}}...\xi_0^{a_l})n_{a_1}...n_{a_l}.$$

(e) By using the above and the original definition of the Legendre polynomials, show that:

$$P_l(\vec{\iota}\cdot\vec{\iota}')=\frac{(2l)!}{2^l l!^2}\mathrm{ST}(\iota^{a_1}...\iota^{a_l})\mathrm{ST}(\iota'_{a_1}...\iota'_{a_l}).$$

**Problem 1.57: Multipole Expansion in Spherical Harmonics.**

(a) Show, with $r=|\vec{x}|$, $r'=|\vec{x}'|$, $\vec{\iota}=\frac{\vec{x}}{r}$, $\vec{\iota}'=\frac{\vec{x}'}{r'}$ and using the results of problem 1.53, that the Coulomb Green's function for the Poisson equation can be written using the Legendre polynomials as:

$$G(\vec{x},\vec{x}')=\frac{1}{4\pi|\vec{x}-\vec{x}'|}=\sum_{l=0}^{\infty}\frac{r'^l}{4\pi r^{l+1}}P_l(\vec{\iota}\cdot\vec{\iota}').$$

(b) Show, using the results of problem 1.54, that the electrostatic multipole expansion can be written in spherical harmonic components as:

$$\phi(\vec{x})=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\frac{Q_{lm}}{4\pi\epsilon_0}\frac{Y_{lm}(\theta,\varphi)}{r^{l+1}}\qquad\text{where:}\qquad Q_{lm}=\frac{4\pi}{2l+1}\int_{\mathbb{S}}\rho(\vec{x})r^l Y^*_{lm}(\theta,\varphi)d\mathbb{S}.$$

(c) Show, using the results of problem 1.56, that the electrostatic multipole expansion can be written in terms of true symmetric traceless multipole moments as:

$$\phi(\vec{x})=\sum_{l=0}^{\infty}\frac{Q_l^{a_1...a_l}\iota_{a_1}...\iota_{a_l}}{4\pi\epsilon_0 r^{l+1}}\qquad\text{where:}\qquad Q_l^{a_1...a_l}=\frac{(2l)!}{2^l l!^2}\int_{\mathbb{S}}\rho(\vec{x})\mathrm{ST}(x^{a_1}...x^{a_l})d\mathbb{S}.$$

**Problem 1.58: Spherical Bessel Functions.**

(a) Using the original definition of the Legendre polynomials in terms of $Y_{l0}$ and (1.7.82), show that the Legendre polynomials satisfy the differential equation:

$$(x^2-1)P_l''(x)+2xP_l'(x)-l(l+1)P_l(x)=0.$$

(b) Define the spherical Bessel function of the first kind $j_l(s)$ by, up to normalization as below, the inverse Fourier transform of $P_l(x)$ over its interval of definition. Show that $j_l(s)$ satisfies:

$$j_l(s)=\frac{1}{2i^l}\int_{-1}^{1}e^{isx}P_l(x)dx\implies s^2j_l''(s)+2sj_l'(s)+(s^2-l(l+1))j_l(s)=0.$$

(c) Consider the above differential equation for $l=0$. Show that two linearly independent solutions are:

$$j_0(s)=\frac{\sin s}{s},\qquad y_0(s)=-\frac{\cos s}{s}$$

where $j_0(s)$ is as defined before and $y_0(s)$ is called a spherical Bessel function of the second kind.

(d) Show, using their explicit form in problem 1.53(b), that the Legendre polynomials satisfy the relation below use this fact to show that the $j_l(s)$ satisfy the recurrence relation:

$$P'_{l+1}(x)=xP_l'(x)+(l+1)P_l(x)\implies j_{l+1}(s)=-j_l'(s)+\frac{l}{s}j_l(s).$$

Define $y_l(s)$ to satisfy the same recurrence relation replacing $j_l$. Show that $j_l(s)$ and $y_l(s)$ provide a basis of solutions to the differential equation satisfied by $j_l(s)$ in (b). Another often useful basis is provided by the spherical Hankel function $h_l(s)=j_l(s)+iy_l(s)$ and its conjugate $h_l^*(s)$.

(e) Show that the recurrence relations for $j_l(s)$ and $y_l(s)$ have the solutions below and that $h_l(s)$ is given by:

$$j_l(s)=(-s)^l\left(\frac{1}{s}\frac{d}{ds}\right)^l\frac{\sin s}{s},\qquad y_l(s)=-(-s)^l\left(\frac{1}{s}\frac{d}{ds}\right)^l\frac{\cos s}{s}\implies h_l(x)=-i(-x)^l\left(\frac{1}{x}\frac{d}{dx}\right)^l\frac{e^{ix}}{x}.$$

Show that $j_l(s)$ is smooth for all real $s$ and that $y_l(s)$ and $h_l(s)$ are smooth everywhere except $s=0$.

(f) Show that for $|s|$ very large, $h_l(s)$ has the asymptotic behavior:

$$h_l(s)=(-i)^{l+1}\frac{e^{is}}{s}\left(1+\mathcal{O}\left(\frac{1}{s^2}\right)\right)$$

(g) Using the completeness of the Legendre polynomials and with $k=|\vec{k}|$, $r=|\vec{x}|$, and $kr\cos\theta=\vec{k}\cdot\vec{x}$,

show that a plane-wave can be decomposed as:

$$e^{i\vec{k}\cdot\vec{x}} = \sum_{l=0}^{\infty}(2l+1)i^l j_l(kr)P_l(\cos\theta).$$

(h) Using the orthonormality of plane waves and the above, show that the $j_l(s)$ satisfy:

$$\int_0^\infty j_l(kr)j_l(kr')k^2dk = \frac{\pi}{2rr'}\delta(r-r').$$

(i) Let $|\psi\rangle$ be an eigenstate everywhere in space except a small ball around the origin of the free Hamiltonian operator $\hat{H}_0 = \frac{\hat{\vec{p}}^2}{2m}$ with eigenvalue $\frac{\hbar^2k^2}{2m}$. Show that for some constants $A_{lm}$ and $B_{lm}$:

$$\psi(\vec{x}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}(A_{lm}j_l(kr) + B_{lm}y_l(kr))Y_{lm}(\theta,\varphi).$$

Explain that if $\psi(\vec{x})$ is an eigenstate of $\hat{H}_0$ everywhere then all of the $B_{lm} = 0$.

## 1.8 Spinor Theory - Pauli Equation

The natural next theory to study after the scalar (spin-0) theory with spin space dimension $D = 1$ is the spinor (spin-1/2) theory with spin space dimension $D = 2$. While the spin-0 theory is pretty good for describing the behavior of electrons in some experiments, it turns out that electrons are spin-1/2 particles (we will study how this is demonstrated experimentally in the homework). Because the wavefunction of a spin-1/2 particle has 2 components, we write $\psi(t,\vec{x})$ as:

$$\psi(t,\vec{x}) = \psi_+(t,\vec{x})\begin{pmatrix}1\\0\end{pmatrix} + \psi_-(t,\vec{x})\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}\psi_+\\\psi_-\end{pmatrix} \tag{1.8.1}$$

where the vector with the upper component 1 corresponds to $|\frac{1}{2},\frac{1}{2}\rangle$, the vector in the spin space with $\hat{S}_z$ eigenvalue of $\hbar/2$, and the vector with lower component 1 corresponds to $|\frac{1}{2},-\frac{1}{2}\rangle$, the vector in the spin space with $\hat{S}_z$ eigenvalue of $-\hbar/2$, and $\psi_\pm$ correspond to the components of $\psi$ in each of these two directions in the spin space. From Theorem 1.11, immediately in this notation the spin matrices for spin-1/2 are given by:

$$\hat{S}_a = \frac{\hbar}{2}\sigma_a, \qquad \text{where:} \qquad \sigma_x = \begin{pmatrix}0&1\\1&0\end{pmatrix}, \qquad \sigma_y = \begin{pmatrix}0&-i\\i&0\end{pmatrix}, \qquad \sigma_z = \begin{pmatrix}1&0\\0&-1\end{pmatrix} \tag{1.8.2}$$

are called the Pauli matrices. Every quantum mechanic has these matrices memorized! By direct calculation, they satisfy several very nice properties. In particular:

$$\sigma_a\sigma_b - \sigma_b\sigma_a = 2i\epsilon_{ab}{}^c\sigma_c \qquad \sigma_a\sigma_b + \sigma_b\sigma_a = 2\delta_{ab}\mathbb{1}. \tag{1.8.3}$$

The first of these had to be true by definition of the spin operators but the second is non-trivial and highly convenient. Together, these allow any product of Pauli matrices to be reduced to a linear combination as:

$$\sigma_a\sigma_b = \delta_{ab}\mathbb{1} + i\epsilon_{ab}{}^c\sigma_c. \tag{1.8.4}$$

Now let us go through the commutators in (1.6.2) again to find the Hamiltonian for a spin-1/2 particle, mirroring what we did in the scalar case:

$$[\hat{\vec{x}},\hat{\vec{G}}] = 0 \implies \hat{\vec{G}} = \hat{\vec{G}}(t,\hat{\vec{x}},\hat{\vec{S}}). \tag{1.8.5}$$

Because $\hat{\vec{G}}$ commutes with $\hat{\vec{x}}$ it still cannot be a function of $\hat{\vec{p}}$. However, now that we have spin degrees of freedom, each of the vector components of $\hat{\vec{G}}$ is a $2\times 2$ matrix on the spin space. Next commuting with the momentum:

$$[\hat{\vec{p}},\hat{\vec{G}}] = im\mathbb{1} = m[\hat{\vec{x}},\hat{\vec{p}}] \implies [\hat{\vec{p}},\hat{\vec{G}} + m\hat{\vec{x}}] = 0 \implies \hat{\vec{G}} = -m\hat{\vec{x}} + \hat{\vec{F}}(t,\hat{\vec{S}}). \tag{1.8.6}$$

Now let us use parity symmetry. Parity by definition inverts space and so a Dirac delta at $\vec{x}$ becomes a Dirac delta at $-\vec{x}$:

$$\hat{\mathcal{P}}|\vec{x}\rangle \overset{\text{def}}{=} |-\vec{x}\rangle \implies \hat{\mathcal{P}} = \int_{\mathbb{S}}|-\vec{x}\rangle\langle\vec{x}|d\mathbb{S}. \tag{1.8.7}$$

Because this is supposed to be a symmetry transformation (and time is not reversed), this better be unitary by Wigner's Theorem. A short calculation shows that it is, that it is self-adjoint, and that it is its own inverse

(this should not be surprising as inverting space twice brings everything back to its original orientation):

$$\hat{\mathcal{P}}^\dagger = \hat{\mathcal{P}}, \qquad \hat{\mathcal{P}}^\dagger\hat{\mathcal{P}} = \hat{\mathcal{P}}^2 = \mathbb{1}. \tag{1.8.8}$$

This tells us how parity acts on the Hilbert space. It is also possible to act parity on the operators. Consider an orthonormal basis $|n\rangle$ and act an operator $\hat{\mathcal{O}}$ on it. By definition, $\hat{\mathcal{O}}|n\rangle$ must be in the Hilbert space and so can be written as a linear combination of the $|n\rangle$. Therefore, there exist coefficients $\mathcal{O}_{mn}$ so that:

$$\hat{\mathcal{O}}|n\rangle = \sum_{mn} \mathcal{O}_{mn}|m\rangle \implies \hat{\mathcal{O}} = \sum_{mn} \mathcal{O}_{mn}|m\rangle\langle n|. \tag{1.8.9}$$

Given how the vectors transform under parity, we can calculate how operators transform under parity if their components are held fixed, as:

$$\hat{\mathcal{O}}' = \sum_{mn} \mathcal{O}_{mn}|m'\rangle\langle n'| = \sum_{mn} \mathcal{O}_{mn}\hat{\mathcal{P}}|m\rangle\langle n|\hat{\mathcal{P}}^\dagger = \hat{\mathcal{P}}\hat{\mathcal{O}}\hat{\mathcal{P}}^\dagger. \tag{1.8.10}$$

From this, with a little algebra, we can see unsurprisingly that the $\hat{\vec{x}}$ and $\hat{\vec{p}}$ operators are inverted by parity:

$$\hat{\mathcal{P}}\hat{\vec{x}}\hat{\mathcal{P}}^\dagger = -\hat{\vec{x}}, \qquad \hat{\mathcal{P}}\hat{\vec{p}}\hat{\mathcal{P}}^\dagger = -\hat{\vec{p}} \tag{1.8.11}$$

exactly matching their classical parity properties as proper vectors. Just like in classical mechanics, the angular momentum is a pseudovector:

$$\hat{\mathcal{P}}\hat{L}_a\hat{\mathcal{P}}^\dagger = \hat{\mathcal{P}}\epsilon_{ab}{}^c\hat{x}^b\hat{p}_c\hat{\mathcal{P}}^\dagger = \epsilon_{ab}{}^c\hat{\mathcal{P}}\hat{x}^b\hat{\mathcal{P}}^\dagger\hat{\mathcal{P}}\hat{p}_c\hat{\mathcal{P}}^\dagger = \epsilon_{ab}{}^c(-\hat{x}^b)(-\hat{p}_c) = \hat{L}_a. \tag{1.8.12}$$

Because orbital angular momentum is a pseudovector, spin angular momentum must be as well (angular momentum conservation conserves only the total $\vec{\mathrm{S}} + \vec{\mathrm{L}}$; for this to hold under parity reversal we need $\vec{\mathrm{S}}$ to also stay unchanged). The fact that spin is parity even allows us to define the parity transformation to act trivially on the spin space. Since $\vec{x}$ is a proper vector, so is the velocity $\vec{v}$. Therefore, under parity a velocity boost goes to a boost by the same amount in the opposite direction which means that the generator of velocity boosts is also a proper vector in a theory with parity symmetry. So:

$$\hat{\mathcal{P}}\hat{\vec{S}}\hat{\mathcal{P}}^\dagger = \hat{\vec{S}}, \qquad \hat{\mathcal{P}}\hat{\vec{v}}\hat{\mathcal{P}}^\dagger = -\hat{\vec{v}}, \qquad \hat{\mathcal{P}}\hat{\vec{G}}\hat{\mathcal{P}}^\dagger = -\hat{\vec{G}}. \tag{1.8.13}$$

Now applying this to (1.8.6):

$$\hat{\mathcal{P}}\hat{\vec{G}}\hat{\mathcal{P}}^\dagger = \hat{\mathcal{P}}(-m\hat{\vec{x}} + \hat{\vec{F}}(t,\hat{\vec{S}}))\hat{\mathcal{P}}^\dagger \implies -\hat{\vec{G}} = m\hat{\vec{x}} + \hat{\vec{F}}(t,\hat{\vec{S}}). \tag{1.8.14}$$

By matching (1.8.14) to (1.8.6) we see that $\hat{\vec{F}} = 0$. Continuing with the final commutator from (1.6.2):

$$[\hat{\vec{v}},\hat{\vec{G}}] = i\hbar\mathbb{1} \implies [\hat{\vec{v}},-m\hat{\vec{x}}] = [\hat{\vec{x}},\hat{\vec{p}}] \implies [\hat{\vec{x}},\hat{\vec{p}} - m\hat{\vec{v}}] = 0 \implies \hat{\vec{p}} - m\hat{\vec{v}} = \hat{\vec{f}}(t,\hat{\vec{x}},\hat{\vec{S}}) \tag{1.8.15}$$

for some $\hat{\vec{f}}(t,\hat{\vec{x}},\hat{\vec{S}})$ which is again a $2\times 2$ matrix on the spin space.

To this point we haven't actually used that we are working with spin-1/2 (and so the results to this point apply to any spin, up to changing the matrix size of $\hat{\vec{f}}$) but now spin-1/2 will be critical. The space of $2\times 2$ complex matrices is 4 dimensional as a complex vector space. Because the identity matrix and the 3 Pauli matrices are linearly independent, it automatically follows that any $2\times 2$ complex matrix can be written as a linear combination of them. Further, because the identity and the Pauli matrices are self-adjoint, any self-adjoint $2\times 2$ complex matrix can be written as a linear combination of them with real coefficients. This means that for spin-1/2 we don't have to consider arbitrary functions of the spin and can instead take without loss of generality:

$$\hat{f}_a(t,\hat{\vec{x}},\hat{\vec{S}}) = \hat{K}_a(t,\hat{\vec{x}})\mathbb{1} - m\hat{Y}_{ab}(t,\hat{\vec{x}})\hat{S}^b \tag{1.8.16}$$

where $\hat{\vec{K}}$ is an operator vector and $\hat{Y}$ an operator matrix, neither acting on spin space. Then (1.8.15) becomes:

$$\frac{\hat{\vec{p}} - \hat{\vec{K}}}{m} + \hat{Y}\cdot\hat{\vec{S}} = \hat{\vec{v}} = \frac{1}{i\hbar}[\hat{\vec{x}},\hat{H}]. \tag{1.8.17}$$

This is a pretty complicated restrint on $\hat{H}$, but we can solve it in steps. We know how to produce part of it from the spin-0 theory, and so define $\hat{H}_0$ just as in (1.6.10):

$$\hat{H}_0 = \frac{(\hat{\vec{p}} - \hat{\vec{K}})^2}{2m} \implies \hat{Y}^{ab}\hat{\vec{S}}_b = \frac{1}{i\hbar}[\hat{x}^a, \hat{H} - \hat{H}_0]. \tag{1.8.18}$$

Next we need to find a term with produces the left hand side of (1.8.18). It has no dependence on $\hat{\vec{p}}$ and is

the result of commuting with $\hat{\vec{x}}$. Commuting with $\hat{\vec{x}}$ kills one power of $\hat{\vec{p}}$ and so the term will be some linear function of $\hat{\vec{p}}$. Consider, for some $3\times 3$ matrix of operators with components with $\hat{\mathcal{O}}_{ab}(t,\hat{\vec{x}})$:

$$\hat{H}_1 = \frac{1}{2}\hat{p}^a\hat{\mathcal{O}}_{ab}(t,\hat{\vec{x}})\hat{S}^b + \frac{1}{2}\hat{S}^b\hat{\mathcal{O}}^\dagger_{ab}(t,\hat{\vec{x}})\hat{p}^a \implies \frac{1}{i\hbar}[\hat{x}_a,\hat{H}_1] = \frac{1}{2}\hat{\mathcal{O}}_{ab}\hat{S}^b + \frac{1}{2}\hat{\mathcal{O}}^\dagger_{ab}\hat{S}^b \tag{1.8.19}$$

where $\hat{\mathcal{O}}^\dagger_{ab}$ means the operator adjoint of the operator $\hat{\mathcal{O}}_{ab}$ and no transpose is taken on the $ab$ components. By identifying:

$$\hat{Y}_{ab} = \frac{\hat{\mathcal{O}}_{ab}+\hat{\mathcal{O}}^\dagger_{ab}}{2} \implies 0 = \frac{1}{i\hbar}[\hat{\vec{x}},\hat{H}-\hat{H}_0-\hat{H}_1] \implies \hat{H}-\hat{H}_0-\hat{H}_1 = \hat{H}_2(t,\hat{\vec{x}},\hat{\vec{S}}) \tag{1.8.20}$$

for some $\hat{H}_2(t,\hat{\vec{x}},\hat{\vec{S}})$. We can use our spin-1/2 trick from before to expand $\hat{H}_2$ as:

$$\hat{H}_2(t,\hat{\vec{x}},\hat{\vec{S}}) = \hat{U}(t,\hat{\vec{x}}) + \hat{\vec{\nu}}(t,\hat{\vec{x}})\cdot\hat{\vec{S}} \tag{1.8.21}$$

for some self-adjoint scalar operator $\hat{U}$ and self-adjoint vector operator $\hat{\vec{\nu}}$. Put together, we have that the Galilean symmetric spin-1/2 Hamiltonian operator takes the form:

$$\hat{H} = \frac{(\hat{\vec{p}}-\hat{\vec{K}}(t,\hat{\vec{x}}))^2}{2m} + \hat{U}(t,\hat{\vec{x}}) + \frac{1}{2}\hat{p}^a\hat{\mathcal{O}}_{ab}(t,\hat{\vec{x}})\hat{S}^b + \frac{1}{2}\hat{\mathcal{O}}^\dagger_{ab}(t,\hat{\vec{x}})\hat{p}^a\hat{S}^b + \hat{\vec{\nu}}(t,\hat{\vec{x}})\cdot\hat{\vec{S}}. \tag{1.8.22}$$

We can simplify our Hamiltonian operator a bit further by expanding $\hat{\mathcal{O}}_{ab}$ in terms of its self-adjoint and anti-self-adjoint parts. By (1.8.20), $\hat{Y}_{ab}(t,\hat{\vec{x}})$ is its self-adjoint part, and so let $\hat{Z}_{ab}(t,\hat{\vec{x}})$ be its anti-self-adjoint part:

$$\hat{Z}_{ab} = \frac{\hat{\mathcal{O}}_{ab}-\hat{\mathcal{O}}^\dagger_{ab}}{2i} \implies \hat{\mathcal{O}}_{ab} = \hat{Y}_{ab} + i\hat{Z}_{ab}, \qquad \hat{\mathcal{O}}^\dagger_{ab} = \hat{Y}_{ab} - i\hat{Z}_{ab} \tag{1.8.23}$$

$$\implies \hat{H} = \frac{(\hat{\vec{p}}-\hat{\vec{K}}(t,\hat{\vec{x}}))^2}{2m} + \hat{U}(t,\hat{\vec{x}}) + \mathrm{Sym}(\hat{p}^a,\hat{Y}_{ab})\hat{S}^b + \frac{i}{2}[\hat{p}_a,\hat{Z}^{ab}]\hat{S}_b + \hat{\vec{\nu}}(t,\hat{\vec{x}})\cdot\hat{\vec{S}}. \tag{1.8.24}$$

Now simplifying the commutator, we have:

$$[\hat{p}_a,\hat{Z}^{ab}] = -i\hbar\frac{\partial\hat{Z}^{ab}(t,\hat{\vec{x}})}{\partial\hat{x}^a} \implies \hat{H} = \frac{(\hat{\vec{p}}-\hat{\vec{K}}(t,\hat{\vec{x}}))^2}{2m} + \hat{U}(t,\hat{\vec{x}}) + \mathrm{Sym}(\hat{p}^a,\hat{Y}_{ab})\hat{S}^b + \left(\hat{\nu}^a + \frac{\hbar}{2}\frac{\partial\hat{Z}^{ba}}{\partial\hat{x}^b}\right)\hat{S}_a. \tag{1.8.25}$$

That is, the term $\hat{Z}^{ab}$ contributes can be completely absorbed into a redefinition of $\hat{\vec{\nu}}$. Similarly, in anticipation of gauge invariance considerations, it would be nicer to have $\hat{p}^a-\hat{K}^a$ in the $\hat{Y}_{ab}$ term than just $\hat{p}^a$. This also just amounts to shifting the definition of $\hat{\vec{\nu}}$ since the extra term $\hat{K}_a\hat{Y}^{ab}\hat{S}_b$ is of the same form as $\hat{\vec{\nu}}\cdot\hat{\vec{S}}$. Thus, after these choices, we have that for some real scalar field $U(t,\vec{x})$, vector fields $\vec{K}(t,\vec{x})$ and $\vec{\nu}(t,\vec{x})$, and matrix field $\hat{Y}(t,\vec{x})$, the Hamiltonian operator takes the form:

$$\hat{H} = \frac{(\hat{\vec{p}}-\hat{\vec{K}}(t,\hat{\vec{x}}))^2}{2m} + \hat{U}(t,\hat{\vec{x}}) + \mathrm{Sym}(\hat{p}^a-\hat{K}^a,\hat{Y}_{ab}(t,\hat{\vec{x}}))\hat{S}^b + \hat{\vec{\nu}}(t,\hat{\vec{x}})\cdot\hat{\vec{S}}. \tag{1.8.26}$$

Let's see what parity transforming our Hamiltonian operator does. For a free particle, we know the Hamiltonian operator will have no nontrivial external fields available to interact with. In general though, $U,\vec{K},\vec{\nu},\hat{Y}$ may originate from interactions, and so may transform under parity based on what fields they are comprised of. Parity symmetry then will require that the Hamiltonian operator takes the same form in terms of their transformed values:

$$\hat{\mathcal{P}}\hat{H}\hat{\mathcal{P}}^\dagger = \frac{(-\hat{\vec{p}}-\hat{\mathcal{P}}\hat{\vec{K}}\hat{\mathcal{P}}^\dagger)^2}{2m} + \hat{\mathcal{P}}\hat{U}\hat{\mathcal{P}}^\dagger + \frac{1}{2}\mathrm{Sym}(-\hat{p}^a-\hat{\mathcal{P}}\hat{K}^a\hat{\mathcal{P}}^\dagger,\hat{\mathcal{P}}\hat{Y}_{ab}\hat{\mathcal{P}}^\dagger)\hat{S}^b + \hat{\mathcal{P}}\hat{\vec{\nu}}\hat{\mathcal{P}}^\dagger\cdot\hat{\vec{S}} \tag{1.8.27}$$

$$\overset{!}{=} \hat{H}' = \frac{(\hat{\vec{p}}-\hat{\vec{K}}')^2}{2m} + \hat{U}' + \mathrm{Sym}(\hat{p}^a-\hat{K}'^a,\hat{Y}'_{ab})\hat{S}^b + \hat{\vec{\nu}}'\cdot\hat{\vec{S}}. \tag{1.8.28}$$

$$\implies \hat{\mathcal{P}}\hat{\vec{K}}\hat{\mathcal{P}}^\dagger = -\hat{\vec{K}}', \qquad \hat{\mathcal{P}}\hat{U}\hat{\mathcal{P}}^\dagger = \hat{U}', \qquad \hat{\mathcal{P}}\hat{Y}_{ab}\hat{\mathcal{P}}^\dagger = -\hat{Y}'_{ab}, \qquad \hat{\mathcal{P}}\hat{\vec{\nu}}\hat{\mathcal{P}}^\dagger = \hat{\vec{\nu}}'. \tag{1.8.29}$$

That is, $U$ is a proper scalar, $\vec{K}$ a proper vector, $\vec{\nu}$ a pseudovector, and $\hat{Y}$ a pseudotensor.

If the spin-1/2 particle is free, the Hamiltonian will necessarily have translational and rotational symmetries, causing all of these to be uniform constants. A uniform constant $U$ can be removed by gauge transformation. The vectors $\vec{K}$ and $\vec{\nu}$ become 0 and $Y_{ab}$ becomes proportional to $\delta_{ab}$ by rotational symmetry.

$Y_{ab}$ then becomes 0 due to its parity transformation property. So, a free spin-1/2 particle has $\hat{H} = \frac{\hat{\vec{p}}^2}{2m}$, just like a scalar particle. The case of a spin-1/2 particle interacting with electromagnetism is naturally more interesting. The logic of making local U(1) gauge transformations as in (1.6.50) is completely unchanged from the scalar case, so we define the gauge covariant derivative in exactly the same way and still identify $\vec{K} = q\vec{A}$ and $U = q\phi$. Indeed, $\phi$ is a proper 3-scalar and $\vec{A}$ a is a proper 3-vector, matching our parity requirements.

For $\hat{Y}$ and $\vec{\nu}$, we can use familiar effective theory logic. To maintain gauge invariance, we need them to only be functions of $\vec{E}$ and $\vec{B}$. At leading order, we need not consider derivatives of $\vec{E}$ and $\vec{B}$, as those contribute at shorter length scales. Out of $\vec{E}$ and $\vec{B}$ we can form three 3-vectors, $\vec{E}$, $\vec{B}$ and $\vec{E} \times \vec{B}$, so $\vec{\nu}$ must be a linear combination of these with coefficients that are rotational scalars made out of $\vec{E}$ and $\vec{B}$. The available rotational scalars are their inner products $|\vec{E}|^2, \vec{E} \cdot \vec{B}, |\vec{B}|^2$. We also need Galilean boost invariance, which by Orig. (3.3.75) only permits the pseudovector $\vec{B}$, proper scalar $|\vec{B}|^2$, and pseudoscalar $\vec{E} \cdot \vec{B}$. Because $\vec{\nu}$ must be a pseudovector, we have that for some function $D_0(|\vec{B}|^2, (\vec{E} \cdot \vec{B})^2)$, $\vec{\nu} = D_0\vec{B}$. The leading order behavior of $D_0$ for weak fields is then for it to just be some constant $-\gamma$, so that $\vec{\nu} = -\gamma\vec{B}$ with corrections that are at least cubing order in the electromagnetic fields. This parameter $\gamma$ is called the gyromagnetic ratio of the particle.

Next for $\hat{Y}$. Due to Galilean boost invariance, the only allowed dependence on $\vec{E}$ is through $\vec{E} \cdot \vec{B}$. The rank two tensors that can be formed then are $\delta_{ab}$, $\epsilon_{abc}B^c$, and $B_aB_b$, all of which are proper tensors. Because $\hat{Y}$ must be a pseudotensor, this means for some functions $D_n(|\vec{B}|^2, (\vec{E} \cdot \vec{B})^2)$, $Y_{ab} = (\vec{E} \cdot \vec{B})(D_1\delta_{ab} + D_2\epsilon_{abc}B^c + D_3B_aB_b)$. At leading order for weak fields each $D_n$ will be some constant, and so the leading $D_1$ term in $\hat{Y}$ is quadratic in the electromagnetic field, the leading $D_2$ term cubic, and the leading $D_3$ term quartic. The overall leading term is thus $Y_{ab} = \beta(\vec{E} \cdot \vec{B})\delta_{ab}$ for some constant $\beta$. So, up to corrections which are at least cubic in the electromagnetic field, the Hamiltonian operator is:

$$\hat{H} = \frac{(\hat{\vec{p}} - q\hat{\vec{A}})^2}{2m} + q\hat{\phi} + \beta \text{Sym}(\hat{\vec{p}} \cdot \hat{\vec{S}}, \hat{\vec{E}} \cdot \hat{\vec{B}}) - \gamma\hat{\vec{B}} \cdot \hat{\vec{S}}. \tag{1.8.30}$$

There is another symmetry which is natural to impose on our Hamiltonian operator. Although we never explicitly demanded it, all of the electromagnetic, gravitational, and particle theories of the Orig. automatically had time reversal symmetry. It is a genuine experimental question whether we should expect such symmetry in our universe. Time reversal symmetry is genuinely observed to hold in experiments where nonrelativistic quantum mechanics is appropriate. Therefore we take:

**Assumption :: GQMV: Time Reversal Symmetry**

Isolated systems have time reversal symmetry.

The Hamiltonian operator (1.8.30) is not automatically time reversal symmetric. Let $\hat{\mathcal{T}}$ be the time reversal pseudo-operator. That is, we write $|\psi'\rangle = \hat{\mathcal{T}}|\psi\rangle$ for the time reversed wavefunction. By problem 1.40, the position operator is time reversal even while the momentum operator is time reversal odd. Immediately it follows that the orbital angular momentum operator $\hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}}$ is time reversal odd, and so the spin operator must be as well (since only the total angular momentum has a conservation law):

$$\hat{\mathcal{T}}\hat{\vec{x}}\hat{\mathcal{T}}^{-1} = \hat{\vec{x}}, \qquad \hat{\mathcal{T}}\hat{\vec{p}}\hat{\mathcal{T}}^{-1} = -\hat{\vec{p}} \implies \hat{\mathcal{T}}\hat{\vec{L}}\hat{\mathcal{T}}^{-1} = -\hat{\vec{L}} \implies \hat{\mathcal{T}}\hat{\vec{S}}\hat{\mathcal{T}}^{-1} = -\hat{\vec{S}}. \tag{1.8.31}$$

Let's see explicitly what time reversal does to the wavefunction in components. For some unitary $2 \times 2$ matrix $\mathtt{t}$ on the spin space we must have:

$$\psi'(t, \vec{x}) = \mathtt{t}\psi^*(-t, \vec{x}). \tag{1.8.32}$$

Now computing the spin:

$$\vec{S}' = \langle\psi'|\hat{\vec{S}}|\psi'\rangle = \langle\psi'^T|\hat{\vec{S}}^T|\psi'^T\rangle = \langle\psi|(\mathtt{t}^\dagger\hat{\vec{S}}\mathtt{t})^*|\psi\rangle \overset{!}{=} -\vec{S} = -\langle\psi|\hat{\vec{S}}|\psi\rangle \implies \mathtt{t}^\dagger\sigma^a\mathtt{t} = -(\sigma^a)^*. \tag{1.8.33}$$

By expanding $\mathtt{t}$ with general components and checking these three conditions, it follows that $\mathtt{t}$ must be proportional to $\sigma_2$. The proportionality must be an overall phase so that $\mathtt{t}$ is unitary. By convention, we take $\mathtt{t} = i\sigma_2$ so that it has real coefficients. that is:

$$\psi'(t, \vec{x}) = i\sigma_2\psi^*(-t, \vec{x}). \tag{1.8.34}$$

Interestingly, if you time reverse the spin-1/2 wavefunction twice, you do not get back to the original state

vector. Letting $\psi''$ be the time reversal of $\psi'$, instead you find:

$$\psi''(t,\vec{x}) = -\psi(t,\vec{x}). \tag{1.8.35}$$

This is perfectly okay, since because $\psi''$ and $\psi$ different only by an overall phase, they produce all of the same operator summaries and represent the same ray in Hilbert space.

By Orig. problem 4.83, the vector potential and magnetic field are time reversal odd while the scalar potential and electric field are time reversal even. Thus, under time reversal, $\hat{H}$ transforms as:

$$\hat{H}' = \frac{(-\hat{\vec{p}} + q\hat{\vec{A}})^2}{2m} + q\hat{\phi} + \beta\text{Sym}((-\hat{\vec{p}})\cdot(-\hat{\vec{S}}), \hat{\vec{E}}\cdot(-\hat{\vec{B}})) - \gamma(-\hat{\vec{B}})\cdot(-\hat{\vec{S}}) \stackrel{!}{=} \hat{H} \implies \beta = 0. \tag{1.8.36}$$

The $\beta$ term is time reversal odd, and so time reversal symmetry requires $\beta = 0$. Finally then:

**Theorem 1.12: The Galilean Spinor Schrödinger Equation**

Up to corrections which are cubic in the electromagnetic fields, the general parity and time reversal symmetric Hamiltonian operator for a Spin-1/2 Galilean quantum wave of mass $m$, charge $q$, and gyromagnetic ratio $\gamma$ interacting with electromagnetism is:

$$\hat{H} = \frac{(\hat{\vec{p}} - q\hat{\vec{A}})^2}{2m} + q\hat{\phi} - \gamma\hat{\vec{B}}\cdot\hat{\vec{S}}. \tag{1.8.37}$$

After much ado, when we include all symmetries we only get one new term for spin-1/2. Plainly, the Schrödinger equation in position space now reads:

$$i\hbar\mathscr{D}_t\psi = -\frac{\hbar^2}{2m}\vec{\mathscr{D}}^2\psi - \gamma\vec{B}\cdot\hat{\vec{S}}\psi. \tag{1.8.38}$$

Let's see how this extra term affects the motion of the center of mass of the wavefunction, following the same steps as those that led to (1.6.26). Most of the calculation is completely unchanged, and the extra term just tacks on at the end:

$$\frac{d\langle\hat{\vec{x}}\rangle}{dt} = \frac{1}{i\hbar}\langle[\hat{\vec{x}},\hat{H}]\rangle = \frac{\langle\hat{\vec{p}}\rangle - q\langle\hat{\vec{A}}\rangle}{m} \implies m\frac{d^2\langle\hat{\vec{x}}\rangle}{dt^2} = \frac{1}{i\hbar}\langle[\hat{\vec{p}} - q\hat{\vec{A}},\hat{H}]\rangle - q\langle\frac{\partial\hat{\vec{A}}}{\partial t}\rangle \tag{1.8.39}$$

$$\implies m\frac{d^2\langle\hat{\vec{x}}\rangle}{dt^2} = q\langle\vec{E}(t,\hat{\vec{x}})\rangle + \frac{q}{2}\langle\Big(\hat{\vec{v}}\times\vec{B}(t,\hat{\vec{x}}) - \vec{B}(t,\hat{\vec{x}})\times\hat{\vec{v}}\Big)\rangle + \gamma\langle\nabla(\vec{B}(t,\hat{\vec{x}})\cdot\hat{\vec{S}})\rangle. \tag{1.8.40}$$

If this new term is very small in a situation which is already approximately classical, then this will lead to the same classical equation of motion as the scalar particle. This new term will be small as long as the magnetic field does not vary too much over the scale of the size of the wave (size meaning the width of the region in which it is mainly localized). Because the two spin eigenvectors of $\vec{B}\cdot\hat{\vec{S}}$ have opposite eigenvalues, the new term will have equal and opposite effects on the two components of $\psi$ projected into the directions aligned and anti-aligned with $\vec{B}$. So the presence of a magnetic field gradient causes these two components to be pulled away from each other (which is known as the Stern-Gerlach effect). If $\vec{B}$ changes persistently in approximately the same direction, this will cause an initial wave-packet to separate into two wave-packets each with spin polarized according to being either aligned or anti-aligned with that change, which leads immediately to interesting experimental consequences. In particular, because it rips the wave-packet apart, it can be used to break the validity of the classical limit.

The calculation leading to the Schrödinger Lagrangian density (1.6.49) carries through just the same also with this new term, so that for a spin-1/2 particle:

$$\mathcal{L} = \int_{\mathbb{S}}\left(i\hbar\psi^\dagger\mathscr{D}_t\psi - \frac{\hbar^2}{2m}|\vec{\mathscr{D}}\psi|^2 + \gamma\psi^\dagger\vec{B}\cdot\hat{\vec{S}}\psi\right)d\mathbb{S}. \tag{1.8.41}$$

From this Lagrangian density, we have an unchanged charge density but the new term does affect the current density: Therefore the charge density is unchanged but the current density picks up a new term:

$$\rho = -\frac{\delta\mathcal{L}_{\text{int}}}{\delta\phi} = q|\psi|^2, \qquad \vec{J} = \frac{\delta\mathcal{L}_{\text{int}}}{\delta\vec{A}} = \frac{q\hbar}{m}\text{Im}(\psi^\dagger\vec{\mathscr{D}}\psi) + \gamma\nabla\times(\psi^\dagger\hat{\vec{S}}\psi). \tag{1.8.42}$$

The new current density is again automatically gauge invariant as the phase added by gauge transformation cancels between $\psi$ and $\psi^\dagger$. The continuity equation follows easily using (1.8.38) because the divergence of

any curl is 0:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{J} = \frac{\partial}{\partial t} q|\psi|^2 + \nabla \cdot \left( \frac{\hbar q}{m} \mathrm{Im}(\psi^\dagger \vec{\mathscr{D}} \psi) + \gamma \nabla \times (\psi^\dagger \hat{\vec{S}} \psi) \right) \tag{1.8.43}$$

$$= \frac{q}{i\hbar} \psi^* \left( i\hbar \mathscr{D}_t \psi + \frac{\hbar^2}{2m} \mathscr{D}^2 \psi \right) - \frac{q}{i\hbar} \left( -i\hbar \mathscr{D}_t \psi^* + \frac{\hbar^2}{2m} \mathscr{D}^2 \psi^* \right) \psi + 0 \tag{1.8.44}$$

$$= \frac{q}{i\hbar} \psi^* \left( -\gamma \vec{B} \cdot \hat{\vec{S}} \right) \psi - \frac{q}{i\hbar} \psi^* (-\gamma \vec{B} \cdot \hat{\vec{S}}) \psi = 0. \tag{1.8.45}$$

Because the charge density is the same as in the scalar theory, (1.6.89) still holds for the electric dipole moment. For the magnetic dipole moment, we have a new contribution (using some integration by parts):

$$\vec{\mu} = \frac{1}{2} \int_{\mathbb{S}} \vec{x} \times \vec{J} d\mathbb{S} = \frac{1}{2} \int_{\mathbb{S}} \vec{x} \times \left( \frac{q\hbar}{2m} \mathrm{Im}(\psi^\dagger \vec{\mathscr{D}} \psi) + \gamma \nabla \times (\psi^\dagger \hat{\vec{S}} \psi) \right) d\mathbb{S} \tag{1.8.46}$$

$$= \frac{q}{2m} \langle \hat{\vec{x}} \times (-i\hbar \hat{\vec{\mathscr{D}}}) \rangle + \frac{\gamma}{2} \int_{\mathbb{S}} \left( \nabla(\psi^\dagger \hat{S}_a \psi) x^a - x^a \nabla_a (\psi^\dagger \hat{\vec{S}} \psi) \right) d\mathbb{S} \tag{1.8.47}$$

$$= \frac{q}{2m} \vec{\mathrm{L}} + \gamma \langle \hat{\vec{S}} \rangle = \frac{q}{2m} \vec{\mathrm{L}} + \gamma \vec{\mathrm{S}}. \tag{1.8.48}$$

The new contribution to the magnetic dipole moment is proportional to the spin of the particle! If the gyromagnetic ratio $\gamma = \frac{q}{2m}$, then $\mu$ is proportional to the total angular momenutm $\vec{\mathrm{L}} + \vec{\mathrm{S}}$. So, we define the $g$-factor:

$$g = \frac{2m\gamma}{q} \implies \vec{\mu} = \frac{q}{2m} (\vec{\mathrm{L}} + g\vec{\mathrm{S}}). \tag{1.8.49}$$

Particles for which $\mu \propto \vec{\mathrm{J}}$ have $g = 1$. However, for different species of spin-1/2 particles $g$ or equivalently $\gamma$ can take any value and is unrestricted by symmetries. There are particles with $q = 0$ and $\gamma \neq 0$, such as the neutron, and so the $g$ factor is not well-defined for every species of particle.

The Hamiltonian operator (1.8.37) can be written nicely using the Pauli matrices and the $g$-factor. Using (1.8.3): In terms of the $g$ factor, the Hamiltonian operator can also be written nicely using the Pauli-matrices.

$$(\sigma^a(\hat{p}_a - q\hat{A}_a))^2 = (\delta^{ab}\mathbb{1} + i\epsilon^{abc}\sigma_c)(\hat{p}_a - q\hat{A}_a)(\hat{p}_b - q\hat{A}_b) = (\hat{\vec{p}} - q\hat{\vec{A}})^2 + \frac{i}{2}\epsilon^{abc}\sigma_c[\hat{p}_a - q\hat{A}_a, \hat{p}_b - q\hat{A}_b] \tag{1.8.50}$$

$$= (\hat{\vec{p}} - q\hat{\vec{A}})^2 + \frac{i}{2}\epsilon^{abc}\sigma_c i\hbar q(\partial_a A_b - \partial_b A_a) = (\hat{\vec{p}} - q\hat{\vec{A}})^2 - 2q\vec{B} \cdot \hat{\vec{S}}. \tag{1.8.51}$$

Returning this to our Hamiltonian operator, we have that it can be written in Pauli form as:

$$\hat{H} = \frac{(\hat{\vec{\sigma}} \cdot (\hat{\vec{p}} - q\hat{\vec{A}}))^2}{2m} + q\hat{\phi} - \frac{q(g-2)}{2m} \hat{\vec{B}} \cdot \hat{\vec{S}}. \tag{1.8.52}$$

This form is particularly nice if the particle of interest has $g = 2$ or very close to it. Historically, for reasons we will partly understand when we study relativistic electrons, the electron was originally believed to have a value of $g = 2$ exactly, motivating this form. When $g = 2$, this is known as the Pauli equation. It turns out that due to higher order interactions between the electron and the electromagnetic field (and with other fields we haven't yet discussed), the $g$ value of the electron gets renormalized so that it instead has the true, experimentally measured value [37]:

$$g_e = 2.00231930436118(27). \tag{1.8.53}$$

## Exercises

**Problem 1.59: Rabi Flopping.** Consider a spin-1/2 particle which is approximately localized in an approximately uniform but time-dependent magnetic field $\vec{B}(t)$.

(a) Show that the relevant Hamiltonian operator for the system's spin evolution is:

$$\hat{H} = -\gamma \hat{\vec{B}} \cdot \hat{\vec{S}}.$$

Show that due to the approximate localization of the particle, we can describe the evolution of the wavefunction's spin degrees of freedom by components $\psi_+(t)$ and $\psi_-(t)$ which represent the total wavefunction of the particle at the localized location in each of the two spin polarization directions.

(b) Consider the magnetic field to have a constant component along the $z$ axis and a rotating piece in the $xy$ plane so that for some constants $B_0, B_1, \omega$:

$$\vec{B}(t) = (B_1 \cos \omega t, -B_1 \sin \omega t, B_0).$$

Show that he Schrödinger equation of motion for the wavefunction's components becomes:

$$\dot{\psi}_+ = \frac{i\gamma B_0}{2}\psi_+ + \frac{i\gamma B_1}{2}e^{i\omega t}\psi_-, \qquad \dot{\psi}_- = \frac{i\gamma B_1}{2}e^{-i\omega t}\psi_+ - \frac{i\gamma B_0}{2}\psi_-.$$

(c) Define a new wavefunction $|\phi\rangle$ by rotating the system about the $z$ axis at angular velocity $\omega$. That is, with $\hat{U}(\theta)$ the representation of an angle $\theta$ $z$-axis rotation on a the spin-1/2 Hilbert space:

$$|\phi\rangle = \hat{U}(\omega t)|\psi\rangle$$

Show that the equations of motion for $\phi_+$ and $\phi_-$ become:

$$\dot{\phi}_+ = \frac{i}{2}(\gamma B_0 - \omega)\phi_+ + \frac{i\gamma B_1}{2}\phi_-, \qquad \dot{\phi}_- = \frac{i\gamma B_1}{2}\phi_+ - \frac{i}{2}(\gamma B_0 - \omega)\phi_-.$$

(d) Solve the equations of motion for $|\phi\rangle$ and return them to the definition of $|\psi\rangle$ to show that:

$$\psi_+(t) = e^{\frac{i}{2}\omega t}\left(\left(\cos\left(\frac{\Omega t}{2}\right) - i\sin\delta\sin\left(\frac{\Omega t}{2}\right)\right)\psi_+(0) + i\cos\delta\sin\left(\frac{\Omega t}{2}\right)\psi_-(0)\right)$$

$$\psi_-(t) = e^{-\frac{i}{2}\omega t}\left(i\cos\delta\sin\left(\frac{\Omega t}{2}\right)\psi_+(0) + \left(\cos\left(\frac{\Omega t}{2}\right) + i\sin\delta\sin\left(\frac{\Omega t}{2}\right)\right)\psi_-(0)\right)$$

where $\omega - \gamma B_0 = \Omega\sin\delta$, $\gamma B_1 = \Omega\cos\delta$.

(e) Suppose the particle begins initially in the "spin-down" state ($|\psi_-(0)| = 1$). Let $I(t)$ be the projection intensity of $|\psi(t)\rangle$ onto the "spin-up" state. Show that:

$$I(t) = |\psi_+(t)|^2 = \frac{\gamma^2 B_1^2}{\gamma^2 B_1^2 + (\omega - \gamma B_0)^2}\sin^2\left(\frac{t}{2}\sqrt{\gamma^2 B_1^2 + (\omega - \gamma B_0)^2}\right).$$

Information can be stored in the spin state of the spin-1/2 particle. By keeping a spin-1/2 particle in a background uniform magnetic field (given by $B_0$), it stably precesses about the direction of that field "saving" its state. By then turning on an oscillatory field (given by $B_1$) we can change the spin state of the particle. Show that by choosing the frequency $\omega$ of the oscillatory field correctly, for an arbitrarily weak applied field $B_1$, we can fully flip the spin state of the particle (change spin-down to spin-up or vice versus). Find the characteristic time $\tau$ necessary to completely flip the spin.

**Problem 1.60: ⋆ Stern-Gerlach Experiment.** Consider a spin-1/2 particle experiencing an interior quadrupole magnetic field:

$$\vec{B} = (0, -My, Mz)$$

for a constant $M$. (This magnetic field may be produced in the interior region of two appropriately spaced current carrying coils which are oppositely oriented.) Assume further that the particle is charge neutral ($q = 0$) but has a gyromagnetic ratio ($\gamma \neq 0$).

(a) For the moment, assume the particle is localized as a wave packet and that it remains localized as it evolves. Show that the center of the particle $\vec{x}(t)$ and the spin angular momentum $\mathrm{S}(t)$ of the particle satisfy the equations of motion:

$$m\ddot{\vec{x}} = \gamma\nabla(\vec{B}(t,\vec{x})\cdot\vec{\mathrm{S}}(t)), \qquad \frac{d\vec{\mathrm{S}}}{dt} = \gamma\vec{\mathrm{S}}(t)\times\vec{B}(t,\vec{x}(t)).$$

Suppose further the particle is initially moving in the $x$ direction with velocity $v$ beginning from an emitter at the origin and travels from where it is emitted to a screen a distance $\ell$ away at which it is detected (through some interaction with the screen). Show that to leading order in $M$, the particle hits the screen at $z$ coordinate:

$$z = \frac{\gamma M\ell^2}{2mv^2}\mathrm{S}_z(0) + \mathcal{O}(M^2).$$

(b) Suppose we now have a beam of particles of the same mechanical properties $(m, q, \gamma)$ emitted as above with different initial spin directions but the same initial velocity. Show that if all initial spin directions are present in the beam, then the beam of particles will produce collisions on the detector in an interval of $z$ coordinates which has length $d$ given to leading order in $M$ by:

$$d \approx \frac{\gamma M\hbar\ell^2}{2mv^2}.$$

Show that the single particle projective intensity per unit $z$ direction length on the screen is given by:

$$\mathcal{I}(z) = \delta\left(z - \frac{\gamma M\hbar\ell^2}{4mv^2}\cos\theta\right)$$

where $\mathrm{S}_z(0) = \frac{\hbar}{2}\cos\theta$. Show that if the initial spins in the beam are uniformly distributed the resulting average intensity per unit length on the screen is given by a uniform distribution from $-\frac{d}{2}$ to $\frac{d}{2}$.

(c) There is a critical flaw in the preceding analysis. It is the result for classical spinning particles as derived from the classical limit of the Schrödinger equation, assuming the particles actually remain as well-packed little wave-packets during their evolution. However, due to the $\langle\nabla(\hat{\vec{S}}\cdot\vec{B})\rangle$ term in (1.8.40) the different spin polarizations are dynamically separated during the actual evolution of the wave. Let $\vec{x}_+$ and $\vec{x}_-$ be the summary positions gotten from using the components $\psi_+$ and $\psi_-$ as wavefunctions respectively (and normalizing them appropriately). Show that to leading order in $M$, $\vec{x}_+$ and $\vec{x}_-$ separate in the $z$ direction so that by the time they have reached the screen they have a separation of $d$ as given above. Therefore, if that separation is large enough to be experimentally meaningful to us, the full wavefunction cannot be said to be in the classical limit. Using this, show that instead the classical limit can be valid for each of the two polarizations separately and that the resulting more accurate estimate of the projective intensity per unit length on the screen for a beam of $N$ particles with the same spin is given by:

$$\mathcal{I}(z) = N\cos^2\left(\frac{\theta}{2}\right)\delta\left(z - \frac{\gamma M\hbar\ell^2}{4mv^2}\right) + N\sin^2\left(\frac{\theta}{2}\right)\delta\left(z + \frac{\gamma M\hbar\ell^2}{4mv^2}\right).$$

Using this show that if the initial spins in the beam are uniformly distributed the resulting average intensity per unit length on the screen is given by:

$$\mathbb{E}[\mathcal{I}] = \frac{N}{2}\delta\left(z - \frac{d}{2}\right) + \frac{N}{2}\delta\left(z + \frac{d}{2}\right).$$

Thus rather than a uniformly dense line segment of collisions on the screen (which would be expected from pure classical mechanics), a better approximation of the Schrödinger equation suggests one should expect two bright spots with clear separation between them on the screen. The two bright spots in reality will certainly not be point-like Dirac deltas and there will be some intensity of signal inbetween them if one uses the full Schrödinger equation (which can be confirmed with numerical simulations or more detailed analytical calculations). But, the prediction of two brighter regions near $\pm\frac{d}{2}$ with a reduced signal between them is correct.

Historically, this experiment was first done with silver atoms, which are charge neutral. Silver atoms have 47 protons, 47 electrons, and stable forms have either 60 or 62 neutrons. The lowest energy configuration of the silver atom involves 46 of the electrons coming in pairs of oppositely oriented spins and forming a spherically symmetric charge density around the atomic nucleus. This causes the total magnetic moment of a silver atom to be very nearly the magnetic moment of the single unpaired electron and so to act as spin-1/2 with the electron's value of $\gamma$ (protons and neutrons are also spin-1/2 and do have nonzero magnetic moments, but they are much smaller than the electron magnetic moment). The fact that such a beam splits into two bright spots acts as a confirmation that in fact the electron is spin-1/2. (For generic spins, the number of spots formed on the screen is $2s+1$; think about why.)

**Problem 1.61: Stern-Gerlach Polarizer.** The Stern-Gerlach experimental setup from the previous question can be used like a "spin polarizer", in the sense that it separates the beam of incoming atoms of various spin directions into two beams, each of which has a specific spin-direction. By keeping only one of the separated beams, we then have ideally a fully (or in reality a mostly) polarized output beam. When we use a Stern-Gerlach magnet setup as in the previous question with the beam separation occurring in the direction $\vec{n}$ and keep for future use only the separated beam which has its spin aligned in the $\vec{n}$ direction, we call it a Stern-Gerlach polarizer in the direction $\vec{n}$. The mathematical structure of this polarizer is nearly identical to the mathematical structure of polarizers of light.

(a) Suppose that $|\psi_{\text{in}}\rangle$ is the (spatially localized, and so given simply by two complex numbers) wavefunction of a spin-1/2 particle entering a Stern-Gerlach polarizer in the direction $\vec{n}$. Show that the wavefunction of the polarizer output $|\psi_{\text{out}}\rangle$ is given by:

$$|\psi_{\text{out}}\rangle = e^{i\alpha}\hat{J}|\psi_{\text{in}}\rangle$$

where $\alpha$ is some phase accumulated while traveling through the polarizer and $\hat{J}$, called the Jones matrix of the polarizer, is given by:

$$\hat{J} = \frac{\mathbb{1} + \vec{n} \cdot \hat{\vec{\sigma}}}{2}.$$

(b) Suppose there is an angle $\theta$ between the spin directions of $|\psi_{\text{in}}\rangle$ and $|\psi_{\text{out}}\rangle$. Let $\mathcal{I}_{\text{in}} = \|\psi_{\text{in}}\|^2$ and $\mathcal{I}_{\text{out}} = \|\psi_{\text{out}}\|^2$ be the projective intensities of the two (not necessarily normalized) beams. Show that:

$$\mathcal{I}_{\text{out}} = \mathcal{I}_{\text{in}} \cos^2 \frac{\theta}{2}.$$

Compare this result to Malus' law from Orig. 4.39.

(c) Suppose we pass the incoming beam through a sequence of three Stern-Gerlach polarizers with Jones matrices $\hat{J}_1, \hat{J}_2, \hat{J}_3$ in the order 1, 2, 3. Show that the resulting wavefunction of the final beam is:

$$|\psi_{\text{out}}\rangle = \hat{J}_3 \hat{J}_2 \hat{J}_1 |\psi_{\text{in}}\rangle.$$

**Problem 1.62: Partially Polarized Particle Beams and Density Matrices.**

(a) Let $G$ be a quantity of interest that depends on the spin of a particle so that it takes the form $G = \langle\psi|\hat{G}|\psi\rangle$ where $\hat{G}$ is a self-adjoint operator on the spin space. Show that equivalently:

$$G = \text{Tr}(\hat{G}\hat{\rho}), \quad \mathcal{I} = \text{Tr}(\hat{\rho}) \quad \text{where:} \quad \hat{\rho} = |\psi\rangle\langle\psi|$$

is a non-negative self-adjoint operator called the density matrix corresponding to $|\psi\rangle$.

(b) Now consider a particle beam source which does not produce particles with exactly predictable spins. Instead, due to our incomplete knowledge we model the source as producing wavefunctions $|\psi_1\rangle$, $|\psi_2\rangle$, ... each with probabilities $p_1$, $p_2$, ... Each particle emitted from the source will have some independently selected $|\psi_1\rangle$ from this distribution. Explain then that as is typical in experiment by sampling many waves from the source one can measure the expected value of the previous mentioned quantity $G$ by:

$$\mathbb{E}[G] = \sum_n p_n \langle\psi_n|\hat{G}|\psi_n\rangle.$$

As well, show that one can **NOT** in general identify a wavefunction $|\psi_0\rangle$ so that $\mathbb{E}[G] \stackrel{!}{=} \langle\psi_0|\hat{G}|\psi_0\rangle$ and thus that there is no way to define an effective wavefunction for a probabilistic source. However, show that one can in general define the density matrix $\hat{\rho}$ so that:

$$\hat{\rho} = \sum_n p_n |\psi_n\rangle\langle\psi_n|, \qquad \mathbb{E}[G] = \text{Tr}(\hat{G}\hat{\rho}).$$

As well, show that in addition to reducing to the original expression given for the density matrix in the case that only one value of $|\psi\rangle$ is possible, this general density matrix is also self-adjoint, has nonnegative eigenvalues, and has trace given by the expected intensity of the beam. Thus, while the wavefunction description cannot be generalized to the case of a probabilistic source, the density matrix description can. Many sources are probabilistic due to them having very complicated internal dynamics and little if any bias toward any particular directions of polarizations. Thus, for modeling real equipment the density matrix formalism is often necessary.

(c) Show that if the beam from such a source passes through a sequence of 3 linear components with Jones matrix $\hat{J}_1, \hat{J}_2, \hat{J}_3$, the density matrix afterward is given by:

$$\hat{\rho}' = \hat{J}_3 \hat{J}_2 \hat{J}_1 \hat{\rho} \hat{J}_1^\dagger \hat{J}_2^\dagger \hat{J}_3^\dagger.$$

(d) Show that there exists an effective wavefunction $|\psi_0\rangle$ for a given density matrix (meaning $\hat{\rho} = |\psi_0\rangle\langle\psi_0|$) if and only if $\text{Tr}(\hat{\rho}^2) = (\text{Tr}(\hat{\rho}))^2$ or, if the beam is of spin-1/2 particles, equivalently if and only if $\det(\hat{\rho}) = 0$. Such a source is called completely polarized.

(e) Consider an unpolarized spin-1/2 beam source, i.e. one which produces all polarizations uniformly. Show that the corresponding density matrix is proportional to the identity matrix. Show that an unpolarized beam produces the same expected values as a source which takes spin-up and spin-down along any particular axis each with probability 1/2.

(f) For a spin-1/2 beam, using that $\hat{\rho}$ is a $2 \times 2$ Hermitian matrix, show that for real Stokes' parameters $S_0, S_1, S_2, S_3$ the density matrix can be written exactly as in Orig. problem 4.40(i). Show that $S_0$ is the intensity of the beam and that the vector $\frac{\hbar}{2}(S_1, S_2, S_3)$ gives the total expected spin angular momentum of the beam. Let $d$, the "degree of polarization" be defined also exactly as in Orig. problem 4.40(i). Show that $0 \leq d \leq 1$ with $d = 0$ if and only if the source is unpolarized and $d = 1$ if and

only if the source is completely polarized. Show that the density matrix can be written as the sum of a completely unpolarized part of intensity $\mathcal{I}_u$ and a completely polarized part of intensity $\mathcal{I}_p$ and polarization (normalized wavefunction) $|\psi\rangle$ as:

$$\hat{\rho} = \mathcal{I}_u \frac{\mathbb{1}}{2} + \mathcal{I}_p |\psi\rangle\langle\psi|$$

and find the relationship between $\mathcal{I}_u, \mathcal{I}_p$, the explicit component expression for $|\psi\rangle$, and the four Stokes parameters. As well, show that the total expected intensity $\mathcal{I}$ is given by $\mathcal{I} = \mathcal{I}_u + \mathcal{I}_p$ and that the degree of polarization can be reexpressed as:

$$d = \frac{\mathcal{I}_p}{\mathcal{I}_p + \mathcal{I}_u}$$

and thus can be understood as the fraction of the total intensity which is due to the polarized beam.

(g) Repeat Orig. problem 4.41 in its entirety, but applied to spin-1/2 waves rather than electromagnetic waves.

**Problem 1.63: Pauli Transport Equations.** Consider the Lagrangian density (1.8.41) together with that of Maxwell's equations, analogously to problem 1.41. Show that spin-1/2 wave has mass density $\rho_m$, momentum density $\vec{\wp}$, energy density $\mathcal{E}$, total stress tensor $\sigma^{ab} = \sigma^{ba}$, and energy current density $\vec{\mathcal{I}}$ given by:

$$\rho_m = m|\psi|^2, \qquad \vec{\wp} = \hbar \mathrm{Im}(\psi^\dagger \vec{\mathscr{D}}\psi) + \frac{1}{2}\nabla \times (\psi^\dagger \hat{\vec{S}}\psi), \qquad \mathcal{E} = \frac{\hbar^2}{2m}|\vec{\mathscr{D}}\psi|^2 - \gamma \psi^\dagger \vec{B}\cdot\hat{\vec{S}}\psi,$$

$$\sigma^{ab} = \frac{\hbar^2}{2m}\mathrm{Re}(\psi^\dagger \mathscr{D}^{(a}\mathscr{D}^{b)}\psi - \mathscr{D}^a\psi^\dagger\mathscr{D}^b\psi) + \gamma\psi^\dagger(\vec{B}\cdot\hat{\vec{S}}\delta^{ab} - B^{(a}\hat{S}^{b)})\psi - \frac{\hbar}{m}\nabla_c \mathrm{Im}(\psi^\dagger \hat{S}_d \epsilon^{cd(a}\mathscr{D}^{b)}\psi),$$

$$\mathcal{I}_a = \frac{\hbar^3}{2m^2}\mathrm{Im}(\mathscr{D}^b\psi^\dagger\mathscr{D}_b\mathscr{D}_a\psi) - \frac{\gamma\hbar}{m}\mathrm{Im}(\psi^\dagger\vec{B}\cdot\hat{\vec{S}}\mathscr{D}_a\psi) - \gamma\psi^\dagger(\vec{E}\times\hat{\vec{S}})_a\psi$$

in the sense that, using (1.8.42), these satisfy the same nonrelativistic transport equations as in problem 1.41.

**Problem 1.64: Schwinger Model Revisited.** Consider a spin-1/2 particle which is effectively localized in position space, so that its spin state can be represented with a single spinor $|\psi(t)\rangle$. Call the component of this spinor aligned with the $z$-direction $a(t)$ and the component anti-aligned $b(t)$.

(a) Show that the Poisson brackets of the components of the spinor are (hint: just use (1.2.27)):

$$\{a,a\} = \{a,b\} = \{a,b^*\} = \{b,b\} = 0, \qquad \{a,a*\} = \{b,b^*\} = -\frac{i}{\hbar}.$$

(b) Show that the spinor's spin angular momentum of the spinor are (hint: just use the Pauli matrices):

$$S_x = \frac{\hbar}{2}(a^*b + b^*a), \qquad S_y = \frac{\hbar}{2i}(a^*b - b^*a), \qquad S_z = \frac{\hbar}{2}(a^*a - b^*b).$$

Compare these to the expressions for the angular momentum operators in Problem 1.46. Explain that by "quantizing" this system, meaning taking it as a classical system and defining a Hilbert space with operators $\hat{a}$ and $\hat{b}$ whose operator summaries have the correct Poisson brackets, we arrive at the Schwinger model of angular momentum.

(c) The Schwinger model thinking is extremely useful for creating multitudes of representations of generic Lie groups. It is also an important hint at how quantum field theory will work. Suppose you have a Lie algebra of interest with generators $\tau_a$ and structure constants $f_{ab}{}^c$:

$$\{\tau_a, \tau_b\} = f_{ab}{}^c \tau_c.$$

If $\hat{\tau}_a$ are some self-adjoint operators on some Hilbert space satisfying:

$$[\hat{\tau}_a, \hat{\tau}_b] = i\hbar f_{ab}{}^c \hat{\tau}_c$$

then by Theorem 1.5 the operator summaries $\langle\hat{\tau}_a\rangle$ form a representation of the algebra. Now consider the components of the wavefunction as $\psi^a$ (and of its conjugate as $\psi^*_a$. Let $(\hat{\tau}_a)^b{}_c$ be the components of $\tau_a$ as a matrix. Show:

$$\{\psi^a, \psi^b\} = 0, \qquad \{\psi^a, \psi^*_b\} = -\frac{i}{\hbar}\delta^a_b, \qquad \tau_a \overset{\text{def}}{=} \langle\hat{\tau}_a\rangle = (\hat{\tau}_a)^b{}_c \psi^*_b \psi^c.$$

Now define a quantum system with wavefunction $|\Psi\rangle$ and operators $\hat{\psi}^a$, $\hat{\psi}^\dagger_a$ so that $\langle\hat{\psi}^a\rangle$ and $\langle\hat{\psi}^\dagger_a\rangle$ satisfy the Poisson-brackets satisfied previously by $\psi^a$. Explain that these operators $\hat{\psi}^a$ for each $a$ act like an

independent set of creation and annihilation operators for a quantum harmonic oscillator. Show that this Hilbert space consists of linear combinations of states $|n_1, ..., n_D\rangle$ where there is an $n_k \in \mathbb{N}$ for each component of $\psi^a$ and where each $n_k$ represents the excitation number corresponding to the harmonic oscillator of $\hat{\psi}^k$. Show that the $\tau_a$ on this Hilbert space are represented by the operators $\hat{T}_a$ as:

$$\hat{T}_a = (\hat{\tau}_a)^b{}_c \hat{\psi}^*_b \hat{\psi}^c.$$

(d) Suppose there is a subspace of the Hilbert space of $|\Psi\rangle$ which is closed under the unitary transformations $e^{-\frac{i}{\hbar}\theta^a \hat{T}_a}$ for real parameters $\theta^a$. Explain that this subspace is itself a representation of the original Lie group of interest. Thus, this construction allows us to take an initial representation of our Lie group (which had Lie algebra generating operators $\hat{\tau}_a$) and produce a bigger Hilbert space which may contain many different representations (in the case of SU(2) it contains all representations).

## 1.9 Scattering Theory

Even more so than in classical mechanics and electromagnetism, scattering experiments are immensely important in quantum mechanics. In scattering we are interested in the question: given an incoming state from the distant past, what will its outgoing state be in the distant future? This may seem bizarre as certainly all experiments take place over finite time. But, if we wait for a time much longer than the time it takes our system to stabilize it is effectively like waiting forever. More rigorously, we could treat scattering as the lowest order approximation in a series expansion in powers of $t_{\text{evolve}}/t_{\text{wait}}$ where $t_{\text{evolve}}$ is a typical characteristic timescale of the Hamiltonian of our system and $t_{\text{wait}}$ is how long we let it evolve before making measurements. For many quantum systems, $t_{\text{evolve}}$ is orders of magnitude smaller than a microsecond and $t_{\text{wait}}$ is a macroscopic human relevant time scale so their ratio will be extremely small that we can effectively treat as 0. Because the interaction happens over the small time $t_{\text{evolve}}$, in scattering we treat the quantum system as being essentially "free" initially, interacting for a little while, and then being free again in the end.

More formally, for scattering we consider a Hamiltonian operator:

$$\hat{H} = \hat{H}_0 + \hat{H}_1 \tag{1.9.1}$$

where $\hat{H}_0$ is the "free-part" (corresponding to the Hamiltonian operator when the system of interest is isolated) and $\hat{H}_1$ is called "interacting-part" (corresponding to everything due to not being isolated). We consider initial and final states which are approximately free. This is made precise by requiring that at very early times (as $t \to -\infty$) and at very late times (as $t \to \infty$), $\hat{H}_1$ should approximately annihilate the state:

$$\lim_{t \to \pm\infty} \|\hat{H}_1 |\psi(t)\rangle\|^2 = 0. \tag{1.9.2}$$

In practice, $\hat{H}_1$ will usually depend on external fields (like electromagnetic fields or atomic potentials) which are only nonzero in a finite region of space and the incoming/outgoing states will be wavepackets coming in from far away or going out to far away, so there is very little (asymptotically 0) overlap between the regions in which $\hat{H}_1$ and $|\psi\rangle$ are mutually nonvanishing.

You might worry that the assumption that we can treat our system as asymptotically not interacting will miss important physics. It can! Check the assumptions of scattering theory before blindly applying it. Scattering theory is excellent at perturbatively giving an approximation to the full theory as a power series in parameters that control the interaction strength. However, sometimes the physical properties of a system may have some nonperturbative behavior that cannot be captured by a power series, such as $e^{-\frac{1}{\lambda}}$ as $\lambda \to 0$. Scattering theory, by treating the interaction as an eventually vanishing deviation from the free system, is essentially intrinsically perturbative and it is just a mathematical fact that perturbation theory doesn't always tell you everything. There may be genuine interacting states of your system no matter how long you wait that need some effective description as a composite object, and such things are missed by perturbation theory. There are many advanced tools for how to treat these naively missed pieces within scattering theory and try to develop something more complete, but we will only touch on some of them briefly.

Because of (1.9.2), scattering essentially separates time into three sections. First, it has the very distant past of the experiment in which the incoming state is approximately free (a beam of freely propagating wavepackets). Next, it has the present middle in which that beam is close to whatever target it is fired at and all the interesting interactions happen. Finally, it has the distant future in which the state is approximately free again and which ultimately becomes a spray of free particles onto a detector surrounding the experiment. Because the states we are considering are nonzero very far away, they will not be square-integrable and thus

not in the proper Hilbert space. Instead, the asymptotic incoming and outgoing states will look like plane waves and only be part of the rigged Hilbert space.

### 1.9.1 Interaction Picture

Because we consider incoming states which asymptotically evolve according to $\hat{H}_0$, it is useful to decompose those states into eigenstates of $\hat{H}_0$. Under the time evolution generated by $\hat{H}_0$ in the Schrödinger equation, an eigenstate of $\hat{H}_0$ will remain that same eigenstate, up to a phase, for all time. Consequently, it is useful to define a way of "dividing out" this phase, so that asymptotically the object we consider becomes genuinely constant. We do this by introducing the interaction picture. Consider a system with state $|\psi(t)\rangle$ evolving under the time independent Hamiltonian operator $\hat{H}$:

$$i\hbar\frac{d|\psi\rangle}{dt} = \hat{H}|\psi\rangle \implies |\psi(t)\rangle = e^{-\frac{i}{\hbar}\hat{H}t}|\psi(0)\rangle. \tag{1.9.3}$$

As well, any operator summary (of an operator $\hat{\mathcal{O}}$) can be computed in time as:

$$\mathcal{O}(t) = \langle\psi(t)|\hat{\mathcal{O}}|\psi(t)\rangle = \langle\psi(0)|e^{\frac{i}{\hbar}\hat{H}t}\hat{\mathcal{O}}e^{-\frac{i}{\hbar}\hat{H}t}|\psi(0)\rangle. \tag{1.9.4}$$

We define the interaction picture state $|\psi_I(t)\rangle$, relative to an arbitrary time $t = 0$, by:

$$\boxed{|\psi_I(t)\rangle = e^{\frac{i}{\hbar}\hat{H}_0 t}e^{-\frac{i}{\hbar}\hat{H}t}|\psi(0)\rangle.} \tag{1.9.5}$$

Using (1.9.1), the interaction picture state evolves according to:

$$i\hbar\frac{d|\psi_I\rangle}{dt} = i\hbar\left(\frac{i}{\hbar}\hat{H}_0 e^{\frac{i}{\hbar}\hat{H}_0 t}e^{-\frac{i}{\hbar}\hat{H}t}|\psi(0)\rangle + e^{\frac{i}{\hbar}\hat{H}_0 t}\left(-\frac{i}{\hbar}\hat{H}\right)e^{-\frac{i}{\hbar}\hat{H}t}|\psi(0)\rangle\right) \tag{1.9.6}$$

$$= -\hat{H}_0|\psi_I(t)\rangle + e^{\frac{i}{\hbar}\hat{H}_0 t}\left(\hat{H}_0 + \hat{H}_1\right)e^{-\frac{i}{\hbar}\hat{H}_0 t}e^{\frac{i}{\hbar}\hat{H}_0 t}e^{-\frac{i}{\hbar}\hat{H}t}|\psi(0)\rangle \tag{1.9.7}$$

$$= e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{H}_1 e^{-\frac{i}{\hbar}\hat{H}_0 t}|\psi_I(t)\rangle. \tag{1.9.8}$$

So, in terms of an interaction picture Hamiltonian operator $\hat{H}_I(t)$, we have the interaction picture Schrödinger equation:

$$\boxed{i\hbar\frac{d|\psi_I\rangle}{dt} = \hat{H}_I(t)|\psi_I\rangle, \qquad \hat{H}_I(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{H}_1 e^{-\frac{i}{\hbar}\hat{H}_0 t}.} \tag{1.9.9}$$

We can see that the interaction picture has the advantage that the state only evolves "due to the interaction". That is, if $\hat{H}_1 = 0$ then the state does not evolve at all. Because we consider past and future states on which $\hat{H}_1$ asymptotically vanishes, this means that as $t \to \pm\infty$, $|\psi_I(t)\rangle$ becomes asymptotically constant.

Inverting (1.9.5) and returning it to (1.9.4), we find that operator summaries can be written in terms of interaction picture operators as:

$$\boxed{\mathcal{O}(t) = \langle\psi_I(t)|\hat{\mathcal{O}}_I(t)|\psi_I(t)\rangle, \qquad \hat{\mathcal{O}}_I(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{\mathcal{O}}e^{-\frac{i}{\hbar}\hat{H}_0 t}.} \tag{1.9.10}$$

That is, in the interaction picture the operators evolve in time according to the free Hamiltonian operator $\hat{H}_0$ only while the state evolves in time according to the interaction part.

Since we consider states satisfying (1.9.2), $|\psi_I(t)\rangle$ asymptotically becomes constant. The value that it takes in the distant past is called $|\psi_{\text{in}}\rangle$ and the value it takes in the distant future is called $|\psi_{\text{out}}\rangle$. Precisely:

$$|\psi_{\text{in}}\rangle = \lim_{t\to-\infty}|\psi_I(t)\rangle = \hat{\Omega}_-|\psi(0)\rangle \qquad \text{where:} \qquad \hat{\Omega}_- = \lim_{t\to-\infty} e^{\frac{i}{\hbar}\hat{H}_0 t}e^{-\frac{i}{\hbar}\hat{H}t} \tag{1.9.11}$$

is the "backward operator" as it takes the state $|\psi(0)\rangle$ in the interacting present and pulls it all the way back into the free past, and:

$$|\psi_{\text{out}}\rangle = \lim_{t\to\infty}|\psi_I(t)\rangle = \hat{\Omega}_+|\psi(0)\rangle \qquad \text{where:} \qquad \hat{\Omega}_+ = \lim_{t\to\infty} e^{\frac{i}{\hbar}\hat{H}_0 t}e^{-\frac{i}{\hbar}\hat{H}t} \tag{1.9.12}$$

is the "forward operator" as it takes the state in the interacting present and pushes it all the way forward into the free future.

We can see automatically from their definitions that the forward and backward operators are unitary[14]:

$$\hat{\Omega}_{\pm}{}^{-1} = \lim_{t\to\pm\infty} e^{\frac{i}{\hbar}\hat{H}t}e^{-\frac{i}{\hbar}\hat{H}_0 t} = \left(\lim_{t\to\pm\infty} e^{\frac{i}{\hbar}\hat{H}_0 t}e^{-\frac{i}{\hbar}\hat{H}t}\right)^{\dagger} = \hat{\Omega}_{\pm}^{\dagger}. \tag{1.9.13}$$

Combining (1.9.11) and (1.9.12), we can see that the incoming and outgoing state are related by the Scattering Operator $\hat{S}$ or "S Matrix":

$$\hat{S} = \hat{\Omega}_+\hat{\Omega}_-^{\dagger} \implies \boxed{|\psi_{\text{out}}\rangle = \hat{S}|\psi_{\text{in}}\rangle.} \tag{1.9.14}$$

It follows from this definition that the S matrix is unitary:

$$\hat{S}^{-1} = \left(\hat{\Omega}_-^{\dagger}\right)^{-1}\hat{\Omega}_+^{-1} = \hat{\Omega}_-\hat{\Omega}_+^{\dagger} = \left(\hat{\Omega}_+\hat{\Omega}_-^{\dagger}\right)^{\dagger} = \hat{S}^{\dagger}. \tag{1.9.15}$$

The unitarity of the S matrix is very important in proving several other important results. The S Matrix pulls a state from the free past and pushes it into the free future and the whole theoretical side of the business of scattering amounts to computing it.

### 1.9.2 Dyson Series

The most direct way to compute the S matrix is to simply solve the interaction picture Schrödinger equation perturbatively. Because (1.9.9) is linear, we know that $|\psi_I(t)\rangle$ is related to $|\psi_I(t_0)\rangle$ by some interaction picture time evolution operator $\hat{U}(t,t_0)$ as:

$$|\psi_I(t)\rangle = \hat{U}(t,t_0)|\psi_I(t_0)\rangle. \tag{1.9.16}$$

Taking the limits that $t\to\infty$ and $t_0\to-\infty$ and using (1.9.14) we have:

$$\hat{S} = \lim_{\substack{t\to\infty\\ t_0\to-\infty}} \hat{U}(t,t_0). \tag{1.9.17}$$

By inserting (1.9.16) into (1.9.9) we find:

$$i\hbar\frac{d\hat{U}(t,t_0)}{dt} = \hat{H}_I(t)\hat{U}(t,t_0), \qquad \hat{U}(t_0,t_0) = \mathbb{1}. \tag{1.9.18}$$

This equation is extremely analogous to Orig. (7.5.69) and we will treat it in exactly the same way. Integrating both sides of the first of these with respect to $t$ from $t_0$ to $t$ and using the second:

$$\hat{U}(t,t_0) = \mathbb{1} - \frac{i}{\hbar}\int_{t_0}^{t}\hat{H}_I(t')\hat{U}(t',t_0)dt'. \tag{1.9.19}$$

Equation (1.9.19) is not obviously helpful as $\hat{U}$ is expressed in terms of itself. But, by inserting another copy of (1.9.19) back into itself on the right hand side we see:

$$\hat{U}(t,t_0) = \mathbb{1} - \frac{i}{\hbar}\int_{t_0}^{t}\hat{H}_I(t')dt' - \frac{1}{\hbar^2}\int_{t_0}^{t}\int_{t_0}^{t'}\hat{H}_I(t')\hat{H}(t'')\hat{U}(t'',t_0)dt''dt'. \tag{1.9.20}$$

We can again circuitously insert equation (1.9.19) back into this again and again and after several iterations we arrive at:

$$\hat{U}(t,t_0) = \mathbb{1} - \frac{i}{\hbar}\int_{t_0}^{t}\hat{H}_I(t')dt' + ... + \left(\frac{-i}{\hbar}\right)^N\int_{t_0}^{t}...\int_{t_0}^{t_2}\hat{H}_I(t_N)...\hat{H}_I(t_1)\hat{U}(t_1,\hat{t}_0)dt_1...dt_N. \tag{1.9.21}$$

As $N$ gets larger, $t_N$ spends an increasing portion of its integral close to $t_0$ and so as $N\to\infty$ we have $\hat{U}(t_1,t_0)\to\hat{U}(t_0,t_0) = \mathbb{1}$. Thus, in the $N\to\infty$ limit we have:

$$\boxed{\hat{U}(t,t_0) = \sum_{n=0}^{\infty}\left(\frac{-i}{\hbar}\right)^n\int_{t_0}^{t}\overset{n}{...}\int_{t_0}^{t_2}\hat{H}_I(t_n)...\hat{H}_I(t_1)dt_1...dt_n.} \tag{1.9.22}$$

This series expansion is called the Dyson series. Cutting off the series after a finite number of terms at least gives an approximation for $\hat{U}(t,t_0)$ and thus can be used then to obtain an approximation for $\hat{S}$. Equation (1.9.22) is often the most useful equation for perturbatively computing the time evolution operator $\hat{U}$. However, it can be simplified in terms of notation (though made less useful for calculation), by introducing the time

[14] Permitting some lack of care on the commutation of limits. This issue is actually extremely technically subtle, but the details are not necessary for any of our analysis.

ordering transformation $\mathscr{T}$ : which is **NOT** an operator. It is a notation for how to write a product of time dependent operators. That is, for 2 operators $\hat{A}(t_1)$ and $\hat{B}(t_2)$:

$$\mathscr{T}: \hat{A}(t_1)\hat{B}(t_2) = \begin{cases} \hat{A}(t_1)\hat{B}(t_2), & t_1 > t_2 \\ \frac{\hat{A}(t_1)\hat{B}(t_2)+\hat{B}(t_2)\hat{A}(t_1)}{2}, & t_1 = t_2 \\ \hat{B}(t_2)\hat{A}(t_1), & t_1 < t_2 \end{cases} \tag{1.9.23}$$

and similarly so for products of 3 or more operators. Then, following exactly the same steps as from Orig. (7.5.73) to (7.5.80), we have that:

$$\hat{U}(t,t_0) = \mathscr{T}: e^{-\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}_I(t')dt'} \implies \boxed{\hat{S} = \mathscr{T}: e^{-\frac{i}{\hbar}\int_{-\infty}^{\infty}\hat{H}_I(t)dt}.} \tag{1.9.24}$$

This formula is true but usually not very useful, at least in ordinary quantum mechanics. However, it does have applications that we will see later.

### 1.9.3 Lippmann-Schwinger Equation

There is another other way of computing the S matrix which has the advantage of taking the $t \to \pm\infty$ limits early rather than at the end. Let $\{|\alpha_0\rangle\}$ be an of orthonormal basis of eigenvectors of the scattering states of the free Hamiltonian operator $\hat{H}_0$. That is:

$$\hat{H}_0|\alpha_0\rangle = E_\alpha|\alpha_0\rangle, \qquad \int |\alpha_0\rangle\langle\alpha_0|d\alpha = \mathbb{1}, \qquad \langle\alpha_0|\beta_0\rangle = \delta(\alpha-\beta) \tag{1.9.25}$$

where the $\mathbb{1}$ operator is understood as the identity operator on the scattering subspace of the rigged Hilbert space (meaning we exclude all the "bound states" which decay too quickly to reach a far away detector). Here we will use $\alpha, \beta, \gamma$ as labels on these states. Because often these labels will be continuous (such as labeling the outgoing wavenumbers of plane waves), we write "sums" over them as integrals. Because we assume we know how to diagonalize $\hat{H}_0$, the scattering operator $\hat{S}$ can be written as:

$$\hat{S} = \int |\beta_0\rangle\langle\beta_0|d\beta\hat{S}\int|\alpha_0\rangle\langle\alpha_0|d\alpha = \int |\beta_0\rangle\langle\beta_0|\hat{S}|\alpha_0\rangle\langle\alpha_0|d\alpha d\beta. \tag{1.9.26}$$

So, all we need to fully specify the S matrix are the "S matrix elements":

$$S_{\beta\alpha} = \langle\beta_0|\hat{S}|\alpha_0\rangle. \tag{1.9.27}$$

Suppose we begin our system with the incoming state as an eigenstate of $\hat{H}_0$, $|\psi_{\text{in}}\rangle = |\alpha_0\rangle$. Then, the present state is given by $|\psi(0)\rangle = \hat{\Omega}_-^\dagger|\alpha_0\rangle$. If we now act the full Hamiltonian operator $\hat{H}$ on this state:

$$\hat{H}|\psi(0)\rangle = \hat{H}\lim_{t\to-\infty} e^{\frac{i}{\hbar}\hat{H}t}e^{-\frac{i}{\hbar}\hat{H}_0t}|\alpha_0\rangle = \lim_{t\to-\infty} e^{\frac{i}{\hbar}\hat{H}t}\left(\hat{H}_0+\hat{H}_1\right)e^{-\frac{i}{\hbar}\hat{H}_0t}|\alpha_0\rangle \tag{1.9.28}$$

$$= E_\alpha|\psi(0)\rangle + \left(\lim_{t\to-\infty} e^{\frac{i}{\hbar}\hat{H}t}\hat{H}_1e^{-\frac{i}{\hbar}\hat{H}t}\right)|\psi(0)\rangle. \tag{1.9.29}$$

This second term is the expression for the interaction term of the Hamiltonian operator in the "active picture" going into the distant past. Because of (1.9.2), it must be 0, which is a significant restriction on how $\hat{H}$ can be divided into $\hat{H}_0$ and $\hat{H}_1$. Automatically $|\psi(0)\rangle$ is a (generalized) eigenstate of the total Hamiltonian operator $\hat{H}$ and it has the same eigenvalue as $|\alpha_0\rangle$. Thus, at the very least requiring the interaction to asymptotically vanish demands that $\hat{H}$ and $\hat{H}_0$ have the same spectrum of scattering states. Since we have pushed our incoming eigenstate $|\alpha_0\rangle$ of $\hat{H}_0$ into the present, we call the resulting eigenstate of $\hat{H}$ $|\alpha_+\rangle$:

$$\boxed{|\alpha_+\rangle = \hat{\Omega}_-^\dagger|\alpha_0\rangle, \qquad \hat{H}|\alpha_+\rangle = E_\alpha|\alpha_+\rangle.} \tag{1.9.30}$$

Similarly, suppose instead that our system ends up as an outgoing state which is an eigenstate of $\hat{H}_0$, $|\psi_{\text{out}}\rangle = |\alpha_0\rangle$. Then, the corresponding present state would be $|\psi(0)\rangle = \hat{\Omega}_+^\dagger|\alpha_0\rangle$. And, if we act the full Hamiltonian operator on this state:

$$\hat{H}|\psi(0)\rangle = \hat{H}\lim_{t\to\infty} e^{\frac{i}{\hbar}\hat{H}t}e^{-\frac{i}{\hbar}\hat{H}_0t}|\alpha_0\rangle = \lim_{t\to\infty} e^{\frac{i}{\hbar}\hat{H}t}\left(\hat{H}_0+\hat{H}_1\right)e^{-\frac{i}{\hbar}\hat{H}_0t}|\alpha_0\rangle \tag{1.9.31}$$

$$= E_\alpha|\psi(0)\rangle + \left(\lim_{t\to\infty} e^{\frac{i}{\hbar}\hat{H}t}\hat{H}_1e^{-\frac{i}{\hbar}\hat{H}t}\right)|\psi(0)\rangle. \tag{1.9.32}$$

For the same reason as in (1.9.29), the second term vanishes from (1.9.2) so that $\hat{\Omega}_+^\dagger|\alpha_0\rangle$ is also an eigenstate of $\hat{H}$. Since we have pushed our outgoing eigenstate $|\alpha_0\rangle$ of $\hat{H}_0$ back into the past, we call the resulting

eigenstate of $\hat{H}$ $|\alpha_-\rangle$:

$$\boxed{|\alpha_-\rangle = \hat{\Omega}_+^\dagger|\alpha_0\rangle, \qquad \hat{H}|\alpha_-\rangle = E_\alpha|\alpha_-\rangle.} \tag{1.9.33}$$

The plus and minus states give a convenient formula for the S matrix elements:

$$S_{\beta\alpha} = \langle\beta_0|\hat{S}|\alpha_0\rangle = \langle\beta_0|\hat{\Omega}_+\hat{\Omega}_-^\dagger|\alpha_0\rangle \implies \boxed{S_{\beta\alpha} = \langle\beta_-|\alpha_+\rangle.} \tag{1.9.34}$$

Thus, finding how to compute $|\alpha_\pm\rangle$ will be our next goal.

Using (1.9.30) and (1.9.33):

$$(\hat{H}_0 + \hat{H}_1)|\alpha_\pm\rangle = E_\alpha|\alpha_\pm\rangle \implies (E_\alpha - \hat{H}_0)|\alpha_\pm\rangle = \hat{H}_1|\alpha_\pm\rangle. \tag{1.9.35}$$

$E_\alpha - \hat{H}_0$ is shorthand for $E_\alpha\mathbb{1} - \hat{H}_0$ where $\mathbb{1}$ is the identity operator. The operator $E_\alpha - \hat{H}_0$ is not invertible because it annihilates any eigenstate of $\hat{H}_0$ with eigenvalue $E_\alpha$ (and we know at least one such state, $|\alpha_0\rangle$, exists). But, because $\hat{H}_0$ is self-adjoint and so has real eigenvalues only, $\hat{E}_\alpha - \hat{H}_0 + i\varepsilon$ annihilates no state in the rigged Hilbert space for any real $\varepsilon \neq 0$ and thus it is invertible. So, letting $\varepsilon$ be a positive real number:

$$(E_\alpha - \hat{H}_0 \pm i\varepsilon)|\alpha_\pm\rangle = \hat{H}_1|\alpha_\pm\rangle \pm i\varepsilon|\alpha_\pm\rangle \tag{1.9.36}$$

$$\implies |\alpha_\pm\rangle = (E_\alpha - \hat{H}_0 \pm i\varepsilon)^{-1}\hat{H}_1|\alpha_\pm\rangle \pm i\varepsilon(E_\alpha - \hat{H}_0 \pm i\varepsilon)^{-1}|\alpha_\pm\rangle. \tag{1.9.37}$$

Now use the integral fact that for a complex number $z$ with positive real part:

$$\int_0^\infty e^{-zx}dx = \frac{1}{z}. \tag{1.9.38}$$

The operator $\varepsilon \mp i(E_\alpha - \hat{H}_0)$ has complex eigenvalues with positive real part only and so (recalling how an operator exponential is defined and inserting the completeness relation over eigenstates of $\hat{H}_0$):

$$\int_0^\infty e^{-(\varepsilon\mp i(E_\alpha - \hat{H}_0))x}dx = \frac{\pm i}{E_\alpha - \hat{H}_0 \pm i\varepsilon}. \tag{1.9.39}$$

Now returning this to (1.9.37) and taking the $\varepsilon \to 0^+$ limit (as it is valid for all positive $\varepsilon$):

$$|\alpha_\pm\rangle = \lim_{\varepsilon\to 0^+}\frac{1}{E_\alpha - \hat{H}_0 \pm i\varepsilon}\hat{H}_1|\alpha_\pm\rangle + \lim_{\varepsilon\to 0^+}\varepsilon\int_0^\infty e^{-(\varepsilon\mp i(E_\alpha - \hat{H}_0))x}dx|\alpha_\pm\rangle. \tag{1.9.40}$$

You need to be careful handling the limit on the first term because at $\varepsilon = 0$ the operator $E_\alpha - \hat{H}_0 \pm i\varepsilon$ loses invertibility. The second term is more straightforward using another integral fact. Consider a function $f(x)$ which has a finite limit as $x \to \infty$:

$$\lim_{\varepsilon\to 0^+}\int_0^\infty \varepsilon e^{-\varepsilon x}f(x)dx \overset{(u=\varepsilon x)}{=} \lim_{\varepsilon\to 0^+}\int_0^\infty e^{-u}f\left(\frac{u}{\varepsilon}\right)du = \int_0^\infty e^{-u}\lim_{x\to\infty}f(x)du = \lim_{x\to\infty}f(x). \tag{1.9.41}$$

Then:

$$\lim_{\varepsilon\to 0^+}\varepsilon\int_0^\infty e^{-(\varepsilon\mp i(E_\alpha - \hat{H}_0))x}|\alpha_\pm\rangle dx = \lim_{\varepsilon\to 0^+}\varepsilon\int_0^\infty e^{-(\varepsilon\mp i(E_\alpha - \hat{H}_0))x}\int|\beta_0\rangle\langle\beta_0|d\beta|\alpha_\pm\rangle dx \tag{1.9.42}$$

$$= \lim_{x\to\infty}\int e^{\pm i(E_\alpha - E_\beta)x}|\beta_0\rangle\langle\beta_0|\alpha_\pm\rangle d\beta \tag{1.9.43}$$

$$= \lim_{t\to\pm\infty}e^{-\frac{i}{\hbar}\hat{H}_0 t}e^{\frac{i}{\hbar}\hat{H}t}|\alpha_\pm\rangle = \hat{\Omega}_\mp|\alpha_\pm\rangle = |\alpha_0\rangle. \tag{1.9.44}$$

Returning this to (1.9.40) gives us the Lippmann-Schwinger equation:

$$\boxed{|\alpha_\pm\rangle = |\alpha_0\rangle + \lim_{\varepsilon\to 0^+}\frac{1}{E_\alpha - \hat{H}_0 \pm i\varepsilon}\hat{H}_1|\alpha_\pm\rangle.} \tag{1.9.45}$$

Similarly to $\hat{U}$ in (1.9.19), $|\alpha_\pm\rangle$ appears on both sides of the Lippmann-Schwinger equation and like in (1.9.19), we can solve it perturbatively in powers of $\hat{H}_1$ by iteratively inserting the equation into its own right hand side over and over. The resulting series is known as the **Born series**.

There is another useful (and equivalent) way to solve the Lippmann-Schwinger equation by using a Green's Operator (which is just the operator way of writing a Green's function, also called a resolvent operator). Define, for a complex number $z$:

$$\hat{G}(z) = (z - \hat{H}_0)^{-1}, \qquad \hat{R}(z) = (z - \hat{H})^{-1}. \tag{1.9.46}$$

$\hat{G}$ is the Green's operator of the unperturbed Hamiltonian operator $\hat{H}_0$ and $\hat{R}$ is the Green's operator of the

full Hamiltonian operator $\hat{H}$. As long as $z$ has non-zero imaginary part these inverses exists because the eigenvalues of $\hat{H}_0$ and $\hat{H}$ are purely real. We will often be interested in the behavior of these operators as the complex number $z$ limits toward being real (surprisingly perhaps, they have sensible limiting behavior). Consequently, we define for a real number $E$:

$$\hat{G}_\pm(E) = \lim_{\varepsilon\to 0^+} \hat{G}(E \pm i\varepsilon), \qquad \hat{R}_\pm(E) = \lim_{\varepsilon\to 0^+} \hat{R}(E \pm i\varepsilon). \tag{1.9.47}$$

Soon we will understand precisely why $\hat{G}_+$ cannot possibly equal $\hat{G}_-$ (and identically so for $\hat{R}_\pm$). From the definition of $\hat{R}(z)$:

$$(z - \hat{H}_0 - \hat{H}_1)\hat{R}(z) = \mathbb{1} \implies \hat{G}^{-1}(z)\hat{R}(z) - \hat{H}_1\hat{R}(z) = \mathbb{1} \implies \hat{R}(z) = \hat{G}(z) + \hat{G}(z)\hat{H}_1\hat{R}(z) \tag{1.9.48}$$

Thus:

$$(\mathbb{1} - \hat{G}(z)\hat{H}_1)(\mathbb{1} + \hat{R}(z)\hat{H}_1) = \mathbb{1}. \tag{1.9.49}$$

Using this, by solving for $|\alpha_0\rangle$, the Lippmann-Schwinger equation solution can be written in terms of $\hat{R}(z)$:

$$|\alpha_0\rangle = (\mathbb{1} - \hat{G}_\pm(E_\alpha))\hat{H}_1|\alpha_\pm\rangle = \left(\mathbb{1} + \hat{R}_\pm(E_\alpha)\hat{H}_1\right)^{-1}|\alpha_\pm\rangle \tag{1.9.50}$$

$$\implies |\alpha_\pm\rangle = |\alpha_0\rangle + \hat{R}_\pm(E_\alpha)\hat{H}_1|\alpha_0\rangle = |\alpha_0\rangle + \lim_{\varepsilon\to 0^+} \frac{1}{E_\alpha - \hat{H} \pm i\varepsilon}\hat{H}_1|\alpha_0\rangle. \tag{1.9.51}$$

Iteratively applying the right side of (1.9.48), we have that $\hat{R}(z)$ is given by:

$$\hat{R} = \hat{G} + \hat{G}\hat{H}_1\hat{G} + \hat{G}\hat{H}_1\hat{G}\hat{H}_1\hat{G} + \hat{G}\hat{H}_1\hat{G}\hat{H}_1\hat{G}\hat{H}_1\hat{G} + ... \tag{1.9.52}$$

Truncating this series after a finite number of terms gives an approximation of the full Green's operator.

### 1.9.4 Transition Matrix

Both the S matrix and the Lippmann-Schwinger equation can be nicely understood in terms of the transition matrix. The transition operator (or transition matrix) $\hat{T}$ is defined to satisfy:

$$\hat{T}|\alpha_0\rangle = \hat{H}_1|\alpha_+\rangle \tag{1.9.53}$$

and the**Transition Amplitudes** $T_{\beta\alpha}$ are defined by:

$$T_{\beta\alpha} = \langle\beta_0|\hat{T}|\alpha_0\rangle = \langle\beta_0|\hat{H}_1|\alpha_+\rangle. \tag{1.9.54}$$

In terms of the transition operator, the plus Lippmann-Schwinger becomes:

$$|\alpha_+\rangle = |\alpha_0\rangle + \lim_{\varepsilon\to 0^+} \frac{1}{E_\alpha - \hat{H}_0 + i\varepsilon}\hat{T}|\alpha_0\rangle. \tag{1.9.55}$$

With this, we can find a nice formula for the S matrix. From (1.9.51):

$$\langle\alpha_-| = \langle\alpha_0| + \langle\alpha_0|\hat{H}_1 \lim_{\varepsilon\to 0^+} \frac{1}{E_\alpha - \hat{H} + i\varepsilon}. \tag{1.9.56}$$

Then, using that $|\alpha_+\rangle$ is an eigenstate of $\hat{H}$:

$$S_{\beta\alpha} = \langle\beta_-|\alpha_+\rangle = \left(\langle\beta_0| + \langle\beta_0|\hat{H}_1 \lim_{\varepsilon\to 0^+} \frac{1}{E_\beta - \hat{H} + i\varepsilon}\right)|\alpha_+\rangle \tag{1.9.57}$$

$$= \langle\beta_0|\alpha_+\rangle + \lim_{\varepsilon\to 0^+} \frac{1}{E_\beta - E_\alpha + i\varepsilon}\langle\beta_0|\hat{H}_1|\alpha_+\rangle \tag{1.9.58}$$

$$= \langle\beta_0|\left(|\alpha_0\rangle + \lim_{\varepsilon\to 0^+} \frac{1}{E_\alpha - \hat{H}_0 + i\varepsilon}\hat{T}|\alpha_0\rangle\right) + \lim_{\varepsilon\to 0^+} \frac{1}{E_\beta - E_\alpha + i\varepsilon}\langle\beta_0|\hat{T}|\alpha_0\rangle \tag{1.9.59}$$

$$= \delta(\alpha - \beta) + T_{\beta\alpha} \lim_{\varepsilon\to 0^+} \left(\frac{1}{E_\alpha - E_\beta + i\varepsilon} - \frac{1}{E_\alpha - E_\beta - i\varepsilon}\right) \tag{1.9.60}$$

The above limit is a little tricky. Let:

$$f(x) = \lim_{\varepsilon\to 0^+} \left(\frac{1}{x + i\varepsilon} - \frac{1}{x - i\varepsilon}\right) = \lim_{\varepsilon\to 0^+} \frac{-2i\varepsilon}{x^2 + \varepsilon^2} \tag{1.9.61}$$

and consider a smooth function $g(x)$. Then:

$$\int_{-\infty}^{\infty} f(x)g(x)dx = \lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\frac{-2i\varepsilon}{x^2+\varepsilon^2}g(x)dx \overset{x=\varepsilon u}{=} \lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\frac{-2i}{u^2+1}g(\varepsilon u)du \tag{1.9.62}$$

$$= \int_{-\infty}^{\infty}\frac{-2i}{u^2+1}g(0)du = -2\pi i g(0) = \int_{-\infty}^{\infty}(-2\pi i\delta(x))g(x)dx. \tag{1.9.63}$$

Because this holds for all smooth $g(x)$, we have:

$$\lim_{\varepsilon\to 0^+}\left(\frac{1}{x+i\varepsilon}-\frac{1}{x-i\varepsilon}\right) = -2\pi i\delta(x). \tag{1.9.64}$$

This kind of limit of $\frac{1}{x+i\varepsilon}$ will be repeatedly useful, so it is also worth looking at how the sum (rather than difference) of these two behaves. Define the so-called "Cauchy Principal Value distribution of $\frac{1}{x}$", written $\mathcal{P}\frac{1}{x}$ by, for any smooth $g(x)$:

$$\lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty} g(x)\mathcal{P}\frac{1}{x}dx = \lim_{\varepsilon\to 0^+}\int_{\varepsilon}^{\infty}\frac{g(x)-g(-x)}{x}dx. \tag{1.9.65}$$

Plainly this principal value distribution under integration is a linear functional on healthy functions, just like the Dirac delta. As well, by just changing $x\to -x$ in the second term below and using the evenness of the resulting integral:

$$\lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\left(\frac{1}{x+i\varepsilon}+\frac{1}{x-i\varepsilon}\right)g(x)dx = 2\lim_{\varepsilon\to 0^+}\int_{\varepsilon}^{\infty}\frac{g(x)-g(-x)}{x}dx = 2\int_{-\infty}^{\infty}g(x)\mathcal{P}\frac{1}{x}dx. \tag{1.9.66}$$

Combining this with (1.9.64), we have the useful identity:

$$\lim_{\varepsilon\to 0^+}\frac{1}{x+i\varepsilon} = \mathcal{P}\frac{1}{x} - i\pi\delta(x). \tag{1.9.67}$$

Applying (1.9.64) to (1.9.60), the S matrix can be written in terms of the transition matrix as:

$$\boxed{S_{\beta\alpha} = \delta(\alpha-\beta) - 2\pi i T_{\beta\alpha}\delta(E_\alpha - E_\beta).} \tag{1.9.68}$$

So, the S matrix can be thought of as decomposing into a part in which nothing happens (the $\delta(\alpha-\beta)$ term) and a term which contains the scattering (the transition matrix term). Further, the scattering term conserves energy[15] between the incoming and outgoing states due to the $\delta(E_\alpha - E_\beta)$. One can see immediately that if $\hat{H}_1 = 0$, then $T_{\beta\alpha} = 0$ and $S_{\beta\alpha} = \delta(\alpha-\beta)$ and so all scattering is due to the interaction part $\hat{H}_1$, as it must have been. Because the S matrix can be written in terms of the transition matrix and the transition matrix elements can be written as in (1.9.54) involving only the $|\alpha_+\rangle$ and not $|\alpha_-\rangle$, we only ever really need to study the + solutions to the Lippmann-Schwinger equation.

Like the Lippmann-Schwinger equation or the Green's operator, the transition matrix can be computed perturbatively. Using (1.9.53) and (1.9.54):

$$\hat{T}|\alpha_0\rangle = \hat{H}_1|\alpha_0\rangle + \hat{H}_1\hat{G}_+(E_\alpha)\hat{T}|\alpha_0\rangle. \tag{1.9.69}$$

Reinserting this formula into the expression for $\hat{T}$ on the right hand side repeatedly gives:

$$\hat{T}|\alpha_0\rangle = (\hat{H}_1 + \hat{H}_1\hat{G}_+\hat{H}_1 + \hat{H}_1\hat{G}_+\hat{H}_1\hat{G}_+\hat{H}_1 + ...)|\alpha_0\rangle \tag{1.9.70}$$

where each copy of $\hat{G}_+$ is evaluated at $E_\alpha$. Because the argument of $\hat{G}_+$ depends on $\alpha$, it is not possible to simply equate the term in parentheses with $\hat{T}$. But, it remains true that:

$$T_{\beta\alpha} = \langle\beta_0|(\hat{H}_1 + \hat{H}_1\hat{G}_+\hat{H}_1 + \hat{H}_1\hat{G}_+\hat{H}_1\hat{G}_+\hat{H}_1 + ...)|\alpha_0\rangle. \tag{1.9.71}$$

Using (1.9.64) we can identify the precise nature of the difference between $\hat{G}_+$ and $\hat{G}_-$:

$$\hat{G}_+(E) - \hat{G}_-(E) = \lim_{\varepsilon\to 0^+}\left(\frac{1}{E+i\varepsilon-\hat{H}_0}-\frac{1}{E-i\varepsilon-\hat{H}_0}\right) = -2\pi i\delta(E-\hat{H}_0) \tag{1.9.72}$$

(with an analogous identity for $R_\pm$ and $\hat{H}$). So $\hat{G}_+$ and $\hat{G}_-$ differ by an operator which annihilates any eigenstate of $\hat{H}_0$ that does not have eigenvalue $E$. As well, neither of $\hat{G}_\pm$ are self-adjoint, instead they are

[15] The energy conservation occurs because we did not allow bound states to form. If we modified the completeness relation among scattering states to include bound states, then the outgoing state's energy may be different than that of the incoming state.

each other's adjoint:

$$\hat{G}_+^\dagger(E) = \lim_{\varepsilon\to 0^+} \frac{1}{E - \hat{H}_0 - i\varepsilon} = \hat{G}_-(E). \tag{1.9.73}$$

### 1.9.5 Transition Rates and Fermi's Master Equation

The transition matrix is essential in characterizing the time averaged rate at which physical processes happen. Suppose at time $t_0$ our system begins in an eigenstate $|\iota\rangle$ of $\hat{H}_0$ (when represented in the interaction picture) and is subject to the perturbation $\hat{H}_1$ until a final time $t$. Let $P_{\iota\to f}$ be the transition intensity for transitions to a different eigenstate $|f\rangle$ of $\hat{H}_0$ at the time $t$. That is:

$$P_{\iota\to f} = |\langle f|\psi(t)\rangle|^2. \tag{1.9.74}$$

Using the interaction picture and that $|f\rangle$ is an eigenstate of $\hat{H}_0$:

$$P_{\iota\to f} = |\langle f|e^{\frac{i}{\hbar}\hat{H}_0 t}|\psi_I(t)\rangle|^2 = |\langle f|\psi_I(t)\rangle|^2 = |\langle f|\hat{U}(t,t_0)|\iota\rangle|^2 \tag{1.9.75}$$

The instantaneous transition rate is defined as the present rate at which transitions occur (and so the $t$ derivative of the transition intensity). Using the interaction picture Schrödinger equation, we can simplify it significantly:

$$\frac{dP_{\iota\to f}}{dt} = \frac{d}{dt}(\langle f|\hat{U}|\iota\rangle\langle\iota|\hat{U}^\dagger|f\rangle) = \langle f|\frac{d\hat{U}}{dt}|\iota\rangle\langle\iota|\hat{U}^\dagger|f\rangle + \langle f|\hat{U}|\iota\rangle\langle\iota|\frac{d\hat{U}^\dagger}{dt}|f\rangle \tag{1.9.76}$$

$$= -\frac{i}{\hbar}\langle f|\hat{H}_I\hat{U}|\iota\rangle\langle\iota|\hat{U}^\dagger|f\rangle + \frac{i}{\hbar}\langle f|\hat{U}|\iota\rangle\langle\iota|\hat{U}^\dagger\hat{H}_I|f\rangle = \frac{2}{\hbar}\mathrm{Im}\left(\langle f|\hat{H}_I(t)\hat{U}(t,t_0)|\iota\rangle\langle\iota|\hat{U}^\dagger(t,t_0)|f\rangle\right). \tag{1.9.77}$$

The long time transition rate, also simply called the transition rate $\Gamma_{\iota\to f}$ is defined as the limit of $\frac{dP_{\iota\to f}}{dt}$ after we have allowed the process to go on for arbitrarily long. (The transition rate is the answer to the question, "At what rate will the process be happening now, if we started it a very long time ago?") Let $\tau = t - t_0$ be the duration for which the interaction has occurred. Keeping $t$ to be in the finite present while taking $\tau \to \infty$, requires that $t_0$ is changed in the limit. Thus:

$$\Gamma_{\iota\to f} = \lim_{\tau\to\infty}\frac{dP_{\iota\to f}}{dt} = \lim_{\tau\to\infty}\frac{2}{\hbar}\mathrm{Im}\left(\langle f|\hat{H}_I(t)\hat{U}(t,t-\tau)|\iota\rangle\langle\iota|\hat{U}^\dagger(t,t-\tau)|f\rangle\right). \tag{1.9.78}$$

The only $\tau$ dependence is in the time evolution operator. By using (1.9.5) and (1.9.16): the definition of the interaction picture wavefunction:

$$|\psi_I(t)\rangle = e^{\frac{i}{\hbar}\hat{H}_0 t}e^{-\frac{i}{\hbar}\hat{H}t}e^{\frac{i}{\hbar}\hat{H}t_0}e^{-\frac{i}{\hbar}\hat{H}_0 t_0}|\psi_I(t_0)\rangle \implies \hat{U}(t,t_0) = e^{\frac{i}{\hbar}\hat{H}_0 t}e^{-\frac{i}{\hbar}\hat{H}(t-t_0)}e^{-\frac{i}{\hbar}\hat{H}_0 t_0}. \tag{1.9.79}$$

Then, the $\tau$ limit is actually quite simple:

$$\lim_{\tau\to\infty}\hat{U}(t,t-\tau) = \lim_{\tau\to\infty}e^{\frac{i}{\hbar}\hat{H}_0 t}e^{-\frac{i}{\hbar}\hat{H}\tau}e^{\frac{i}{\hbar}\hat{H}_0\tau}e^{-\frac{i}{\hbar}\hat{H}_0 t} = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{\Omega}_-^\dagger e^{-\frac{i}{\hbar}\hat{H}_0 t}. \tag{1.9.80}$$

Returning this to the transition rate and using the definition of the interaction Hamiltonian:

$$\Gamma_{\iota\to f} = \frac{2}{\hbar}\mathrm{Im}\left(\langle f|e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{H}_1 e^{-\frac{i}{\hbar}\hat{H}_0 t}e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{\Omega}_-^\dagger e^{-\frac{i}{\hbar}\hat{H}_0 t}|\iota\rangle\langle\iota|e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{\Omega}_- e^{-\frac{i}{\hbar}\hat{H}_0 t}|f\rangle\right) \tag{1.9.81}$$

$$= \frac{2}{\hbar}\mathrm{Im}\left(\langle f|\hat{H}_1\hat{\Omega}_-^\dagger|\iota\rangle\langle\iota|\hat{\Omega}_-|f\rangle\right) \tag{1.9.82}$$

By using the definition of the plus states, (1.9.54), and (1.9.55):

$$\langle f|\hat{H}_1\hat{\Omega}_-^\dagger|\iota\rangle = \langle f|\hat{H}_1|\iota_+\rangle = \langle f|\hat{T}|\iota\rangle = T_{f\iota} \tag{1.9.83}$$

$$\langle f|\hat{\Omega}_-^\dagger|\iota\rangle = \langle f|\iota_+\rangle = \langle f|\iota\rangle + \lim_{\varepsilon\to 0^+}\langle f|\frac{1}{E_\iota - \hat{H}_0 + i\varepsilon}\hat{T}|\iota\rangle = \lim_{\varepsilon\to 0^+}\frac{T_{f\iota}}{E_\iota - E_f + i\varepsilon} \tag{1.9.84}$$

where in the last step we used that we are considering a final eigenstate $|f\rangle \neq |\iota\rangle$. Using these, we can simplify the transition rate to arrive at **Fermi's Golden Rule**:

$$\Gamma_{\iota\to f} = \frac{2}{\hbar}\lim_{\varepsilon\to 0^+}\mathrm{Im}\left(T_{f\iota}\frac{T_{f\iota}^*}{E_\iota - E_f - i\varepsilon}\right) = \frac{2}{\hbar}\lim_{\varepsilon\to 0^+}\frac{|T_{f\iota}|^2}{2i}\left(\frac{1}{E_\iota - E_f - i\varepsilon} - \frac{1}{E_\iota - E_f + i\varepsilon}\right) \tag{1.9.85}$$

$$\implies \boxed{\Gamma_{\iota\to f} = \frac{2\pi}{\hbar}|T_{f\iota}|^2\delta(E_f - E_\iota).} \tag{1.9.86}$$

Because the microscopic timescales of quantum systems are typically so fast, $\Gamma_{\iota\to f}$ often serves as an extremely good approximation to $\frac{dP_{\iota\to f}}{dt}$ on the experimentally relevant timescale.

The total decay rate $\Gamma_\iota$ for the state $|\iota\rangle$ is defined as the long term value of the rate of change of $|\langle\iota|\psi(t)\rangle|^2$. That is:

$$\Gamma_\iota = \lim_{\tau\to\infty} \frac{dP_{\iota\to\iota}}{dt}. \tag{1.9.87}$$

Due to the unitarity of time evolution (and inserting a complete set of eigenstates $|n\rangle$ of $\hat{H}_0$, where a sum versus an integral is used as appropriate for the system of interest):

$$0 = \frac{d}{dt}\langle\psi_I|\psi_I\rangle = \frac{d}{dt}\sum_n \langle\psi_I|n\rangle\langle n|\psi_I\rangle = \frac{d}{dt}\left(P_{\iota\to\iota} + \sum_{f\neq\iota} P_{\iota\to f}\right) \implies \boxed{\Gamma_\iota = -\sum_{f\neq\iota}\Gamma_{\iota\to f}} \tag{1.9.88}$$

where in the last step we have taken the $\tau\to\infty$ limit.

It is often useful to consider separately the magnitude and phase of each component of the interaction wavefunction (in the basis of free eigenstates). So, define $\rho_n$ and $\varphi_n$ as real quantities so that:

$$\langle n|\psi_I(t)\rangle = \sqrt{\rho_n(t)}e^{i\varphi_n(t)} \implies \rho_n = |\langle n|\psi_I(t)\rangle|^2. \tag{1.9.89}$$

We found in section 1.2 that the $\rho_n$ are adiabatic invariants and so evolve much more slowly in time than the phases $\varphi_n$. As well, the operator summary of the free Hamiltonian and anything that commutes with it only depends on the wavefunction through the $\rho_n$. Therefore, often the $\rho_n$ are all that is accessible in a direct way experimentally as the phases are generally much more temperamental (in some experiments, phase differences may be quite well-behaved, so this argument is only one which is often useful but not always). Using the definition of the time evolution operator, we can compute an exact formula for the time evolution of $\rho_n$, valid for all $t$ and $t_0$:

$$\rho_n(t) = |\langle n|\hat{U}(t,t_0)|\psi_I(t_0)\rangle|^2 = \sum_{ml}\langle\psi_I(t_0)|m\rangle\langle m|\hat{U}^\dagger|n\rangle\langle n|\hat{U}|l\rangle\langle l|\psi_I(t_0)\rangle \tag{1.9.90}$$

$$= \sum_{ml}\langle m|\hat{U}^\dagger|n\rangle\langle n|\hat{U}|l\rangle\sqrt{\rho_m(t_0)\rho_l(t_0)}e^{i(\varphi_l(t_0)-\varphi_m(t_0))}. \tag{1.9.91}$$

Every experiment is performed in a background of weakly coupled junk throughout the universe (for example, the electromagnetic field is everywhere and has infinitely many harmonic oscillator modes because it can oscillate with any frequency; those modes interact weakly with basically everything). Consequently, it is typically accurate to assume that the experiment is prepared in such a way that the initial phases of the system are completely random with respect to our preparation and to use that the phases evolve very quickly in time, effectively randomly so due to their weak bombardment from an inescapable background of degrees of freedom. The $\rho_n$ are saved from this fate specifically due to their adiabatic invariance. Using this, we can define a sensible notion of a phase-averaged behavior of this equation. If $\varphi_n(t_0)$ is uniformly randomly distributed and each phase is independent, then averaging over such phases immediately gives:

$$\mathbb{E}\left[e^{i(\varphi_l(t_0)-\varphi_m(t_0))}\right] = \delta_{lm}. \tag{1.9.92}$$

When it is reasonable in regard to experimental knowledge to treat the phases of a quantum system as effectively random (essentially due to them having unknown uniformly distributed initial conditions and oscillating extremely rapidly and subject to uncontrolled interactions with the environment), the system is said to be **decoherent**. Macroscopically large systems are essentially guaranteed to be decoherent because of the number of background degrees of freedom necessarily affecting them. Under the decoherence approximation then, (1.9.91) becomes:

$$\rho_n(t) \approx \sum_{ml}\langle m|\hat{U}^\dagger|n\rangle\langle n|\hat{U}|l\rangle\sqrt{\rho_m(t_0)\rho_l(t_0)}\delta_{ml} = \sum_m |\langle n|\hat{U}|m\rangle|^2\rho_m(t_0) = \sum_m P_{m\to n}\rho_m(t_0). \tag{1.9.93}$$

Note that the decoherence approximation involves both a time averaging (over the fast scale of the phase evolution) and an ensemble averaging (over the many realizations of phase initial conditions). Differentiating with respect to time gives:

$$\frac{d\rho_n}{dt} \approx \sum_m \frac{dP_{m\to n}}{dt}(t,t_0)\rho_m(t_0) = \sum_{m\neq n}\frac{dP_{m\to n}}{dt}(t,t_0)(\rho_m(t_0)-\rho_n(t_0)). \tag{1.9.94}$$

We now use the adiabatic invariance of $\rho_n$ to further approximate this. The explicit duration dependence

of $\frac{dP_{m\to n}}{dt}(t,t_0)$ is determined by the microscopic physics and so by the microscopic timescale $\tau_{\text{micro}}$ of the quantum system. That timescale is very small compared to the experimental resolution timescale $\tau_{\text{exp}}$ on which we will apply this equation, and so by taking $t-t_0$ to be of order $\tau_{\text{exp}}$, we can approximately replace $\frac{dP_{m\to n}}{dt}(t,t_0)$ by its long duration value of $\Gamma_{m\to n}$. However, due to the adiabatic invariance theorem, the timescale $\tau_{\text{adiabatic}}$ on which the $\rho_n$ change significantly is of order $\frac{\tau_{\text{exp}}^2}{\tau_{\text{micro}}}$, and so if $t-t_0$ is of order $\tau_{\text{exp}}$ then $\rho_n(t_0)$ is approximately $\rho_n(t)$ as there is not a significant enough time difference between $t$ and $t_0$ for $\rho_n$ to have changed appreciably. This hierarchy of timescales is valid precisely only because the $\rho_n$ are the adiabatic invariants of $\hat{H}_0$. Under this approximation, (1.9.94) produces a closed differential equation for the $\rho_n(t)$:

$$\boxed{\frac{d\rho_n(t)}{dt} = \sum_{m\neq n} \Gamma_{m\to n}(\rho_m(t) - \rho_n(t)).} \tag{1.9.95}$$

This "equation" (which is only truly an often very good approximation valid for a decoherent system on long time scales) is called the **Fermi Master Equation**. The Fermi master equation is only valid because there is a systematic information loss in treating the phases as decoherent. In general in a quantum system an enormity of the information in the system state is stored in the phases, and phase information is necessary for the existence of unitary, invertible time evolution. Decoherence occurs due to the assumption of a weak coupling to a large number of background degrees of freedom and a time averaging, with essentially identical thinking to how the canonical ensemble is achieved in statistical mechanics from those same factors.

### 1.9.6 Analytic Properties of the S Matrix

By using (1.9.51) and (1.9.53), we can define a generalization of the transition operator $\hat{T}$ to a function of the energy eigenvalue $\hat{T}(E)$ in terms of the Green's operator of the full system $\hat{R}_+(E)$:

$$\hat{T}(E) = \hat{H}_1 + \hat{H}_1\hat{R}_+(E)\hat{H}_1 \implies \hat{T}|\alpha_0\rangle = \hat{T}(E_\alpha)|\alpha_0\rangle \tag{1.9.96}$$

Because of this formula, the behavior of the matrix elements of $\hat{T}(E)$ (and thus equivalently, the matrix elements of the S matrix) as a function of $E$ can reveal information about the bound states of the full Hamiltonian operator $\hat{H}$. The matrix elements of $\hat{R}_+$ corresponding to any square-integrable eigenstate $|\psi\rangle$ of $\hat{H}$ with energy eigenvalue $E_\psi$ develop poles at that energy. One can see this by considering the operator summary of $\hat{R}_+$ in such a state:

$$\langle\psi|\hat{R}_+(E)|\psi\rangle = \lim_{\varepsilon\to 0^+}\langle\psi|\frac{1}{E-\hat{H}+i\varepsilon}|\psi\rangle = \lim_{\varepsilon\to 0+}\frac{\langle\psi|\psi\rangle}{E-E_\psi+i\varepsilon} = \frac{1}{E-E_\psi} \tag{1.9.97}$$

where the last equality is okay if the expression is only evaluated numerically and not used as a distribution to be integrated against. Consequently, through its dependence on $\hat{R}_+(E)$, the matrix elements of $\hat{T}(E)$ develop poles at values of $E$ corresponding to the eigenvalues of normalizable eigenstates of the full Hamiltonian. The normalizable eigenstates are precisely states which decay sufficiently quickly at large distances to be square-integrable, i.e. bound states. Therefore, bound states of $\hat{H}$ appear as poles in the transition amplitudes (even though we originally set out when doing scattering to study unbound states!) Moreover, the only place in (1.9.96) that poles may appear is in $\hat{R}_+$, and so **each bound state of the full theory manifests as a real energy pole in the transition amplitudes, and each real energy pole in the transition amplitudes indicates a bound state in the full theory.**

I was careful to indicate that bound states correspond specially to poles of the matrix elements of $\hat{T}(E)$ which occur at real values of $E$. It is also insightful to study the behavior of $\hat{T}(E)$ as a function of $E$ treated as a complex number. Because the Hamiltonian operator $\hat{H}$ is self-adjoint, all of its eigenvalues corresponding to states in the rigged Hilbert space are guaranteed to be real. However, it may be possible (and often is) to construct "states", called virtual states, which are not actually in the Hilbert space (but are in an appropriately larger set containing the Hilbert space) which have complex eigenvalues under $\hat{H}$. For example, if $\hat{H}$ were the simple free particle Hamiltonian $\frac{\hat{p}^2}{2m}$ in one spatial dimension, the function $\psi(x) = e^{i\kappa x}$ is an "eigenstate" of this operator for every complex value of $\kappa$, however if $\kappa$ has any nonzero imaginary part then this function blows up at either large positive or large negative $x$, and so is not even in the rigged Hilbert space. Consequently, such a $\psi$ for general complex $\kappa$ is called a virtual state. Virtual states are sometimes useful because while no state in the Hilbert space may actually be a virtual state, a real state in the Hilbert space may be (at least for finite times in some regions of space) well-approximated by a virtual state and virtual states have very simple time evolution. This may sound odd but the situation is actually familiar

from electromagnetism where we know the electromagnetic fields on all space can be expressed in terms of a linear combination of plane waves but in certain regions of space (such as in a box) may be very accurately approximated by spatially or temporally decaying solutions. Because a virtual state $|\psi(t)\rangle$ is like an energy eigenstate with complex energy $\widetilde{E}_0$, its time evolution is given by:

$$i\hbar\frac{d|\psi\rangle}{dt}=\hat{H}|\psi\rangle \implies \frac{d|\psi\rangle}{dt}=-\frac{i}{\hbar}\widetilde{E}_0|\psi\rangle \implies |\psi(t)\rangle=e^{-\frac{i}{\hbar}\widetilde{E}_0 t}|\psi(0)\rangle=e^{-\frac{i}{\hbar}\mathrm{Re}(\widetilde{E}_0)t}e^{\frac{\mathrm{Im}(\widetilde{E}_0)}{\hbar}t}|\psi(0)\rangle. \quad (1.9.98)$$

We see that the real part of $\widetilde{E}_0$ simply gives oscillations in time in line with the behavior of real states while the imaginary part of $\widetilde{E}_0$ gives either exponential growth or decay (depending on its sign) which is very much unlike the behavior of real states. It is useful to introduce the symbol $E_0$ so the real part so that $\frac{E_0}{\hbar}$ is the real angular frequency and $\Gamma$ as the decay rate of the virtual state:

$$E_0=\mathrm{Re}(\widetilde{E}_0), \qquad \Gamma=-\frac{2\mathrm{Im}(\widetilde{E}_0)}{\hbar} \implies |\psi(t)\rangle=e^{-\frac{i}{\hbar}E_0 t}e^{-\frac{\Gamma}{2}t}|\psi(0)\rangle. \quad (1.9.99)$$

Because we are discussing a virtual state, it necessarily has infinite norm. However, we could consider a cutoff (like only integrating over a finite region of space or summing over a finite number of terms, depending on the system) controlled by a parameter $\Lambda$ so that when $\Lambda\to\infty$ the full norm is taken but at finite $\Lambda$ a finite result is achieved ($\Lambda$ could be the volume of the integration region, for example). Then:

$$\|\psi(t)\|_\Lambda^2 \text{ is finite and } \lim_{\Lambda\to\infty}\|\psi(t)\|_\Lambda^2=\infty. \quad (1.9.100)$$

We can see that the complexity of $\widetilde{E}_0$ causes the norm of $|\psi(t)\rangle$ to change:

$$\|\psi(t)\|_\Lambda^2=e^{-\Gamma t}\|\psi(0)\|_\Lambda^2 \quad (1.9.101)$$

which is why $\Gamma$ (when positive) is identified with the decay rate of the virtual state. Virtual states with positive $\Gamma$ are for many purposes more interesting than virtual states with negative $\Gamma$, because they are often much better approximations of true physical behavior. The reason why is because for positive $\Gamma$ the virtual state decays and so the fact that a real state can only be approximated locally and for a finite time is compatible with the fact that after a finite time the virtual state is negligible and so the other components of the wavefunction dominant while for negative $\Gamma$, the virtual state explodes with time and so necessarily becomes a very poor approximation of the physical state as the part approximated by the virtual state comes to dominate any other components. (However, negative $\Gamma$ virtual states are interesting for studying instabilities in systems.) Going forward we will assume $\Gamma>0$ for the applications desired.

When viewed as a decay rate, $\Gamma$ can also be related to a "lifetime" $\tau$ of the state. Let $p(t)$ be defined by:

$$p(t)=1-\frac{\|\psi(t)\|_\Lambda^2}{\|\psi(0)\|_\Lambda^2}=1-e^{-\Gamma t} \quad (1.9.102)$$

be the fraction of the state's magnitude which has not decayed after a time $t$. The instantaneous rate of decay of the state is then $\frac{dp}{dt}$ and the average decay time (meaning average time at which a given amount of state magnitude has decayed), called the lifetime, is:

$$\tau=\int_0^\infty t\frac{dp}{dt}dt=\frac{1}{\Gamma}. \quad (1.9.103)$$

The half-life of the state, meaning the time $\tau_{1/2}$ after which half of the magnitude of the state has decayed is:

$$e^{-\Gamma\tau_{1/2}}\overset{!}{=}\frac{1}{2} \implies \tau_{1/2}=\tau\ln 2=\frac{\ln 2}{\Gamma}. \quad (1.9.104)$$

Because a virtual state corresponds to a "complex eigenvalue eigenstate" of $\hat{H}$ in an extended state space, it similarly corresponds to a complex pole of $\hat{R}_+(E)$ and $\hat{T}(E)$. We can see this through an analogous calculation to (1.9.97):

$$\frac{\langle\psi|R_+(E)|\psi\rangle_\Lambda}{\|\psi\|_\Lambda^2}=\lim_{\varepsilon\to 0^+}\frac{1}{\|\psi\|_\Lambda^2}\langle\psi|\frac{1}{E-\hat{H}+i\varepsilon}|\psi\rangle_\Lambda=\frac{1}{E-E_0+\frac{i\hbar\Gamma}{2}}. \quad (1.9.105)$$

Therefore, by exactly the same thinking as in the bound states, **each virtual state of the full theory manifests as a complex energy pole in the transition amplitudes, and each complex energy pole in the transition amplitudes indicates a virtual state in the full theory.**

Many physical observable quantities of interest (such as scattering cross-sections or Fermi's Golden Rule transition rates) depend on the magnitude-squared of transition amplitudes. Suppose $\hat{H}$ has a virtual or bound state so that there is a pole in a particular transition amplitude $T(E)$ (some particular matrix element

$T_{fi}$) when viewed as a function of energy $E$ at a complex value $E_0 - \frac{i\hbar\Gamma}{2}$, exactly as above. Then because of (1.9.105), for energies in the neighborhood of that value, the transition amplitude will take the form:

$$T(E) = \frac{A}{E - E_0 + \frac{i\hbar\Gamma}{2}} + B(E) \tag{1.9.106}$$

where $A$ is some constant and $B(E)$ is a complex function which is finite at $E = E_0 - \frac{i\hbar\Gamma}{2}$. Suppose that $\Gamma$ is sufficiently small that $B(E)$ is finite and smooth as a function of $E$ around $E = E_0$ with a radius larger than $\frac{\hbar\Gamma}{2}$ (the state, if virtual, is thus assumed not to decay *too* quickly). Consequently, the magnitude-squared of the transition amplitude becomes:

$$|T(E)|^2 = \frac{|A|^2 - \hbar\Gamma \mathrm{Im}(A^* B(E_0))}{(E - E_0)^2 + \frac{\hbar^2\Gamma^2}{4}} + |B(E_0)|^2 + \mathcal{O}(E - E_0). \tag{1.9.107}$$

We can see that when the full theory has a virtual or bound state, there is a local maximum in the energy dependence of the observable (which depends on $E$ through $|T(E)|^2$) at $E = E_0$. The local maximum value of $|T(E)|^2$ is:

$$M_{\max} = 4\frac{|A|^2 - \hbar\Gamma \mathrm{Im}(A^* B(E_0))}{\hbar^2\Gamma^2} + |B(E_0)|^2 + \mathcal{O}(E - E_0). \tag{1.9.108}$$

For sufficiently small $\Gamma$, this first term dominates the rest so that nearby the maximum we may approximate:

$$|T(E)|^2 \approx \frac{\frac{\hbar^2\Gamma^2}{4}}{(E - E_0)^2 + \frac{\hbar^2\Gamma^2}{4}} M_{\max}. \tag{1.9.109}$$

This gives the characteristic profile of the shape of peaks in scattering observables as a function of energy. This profile is known as the **Breit-Wigner distribution**. Conversely, peaks in scattering observables (which are directly experimentally measurable) imply the existence of virtual or bound states in the interacting system. Notice that for a virtual state, the peak is of finite height (and is simply called called a resonance) whereas for a bound state ($\Gamma \to 0$) the peak is of infinite height (and is also called a resonance, but more specifically a pole) as $M_{\max} \to \infty$. We can give another interpretation to the decay rate in terms of the profile of the resonance. The Full-Width-at-Half-Maximum (FWHM) of a profile is the length $b - a$ of the smallest interval $[a, b]$ which contains the maximum of the profile and so that at each of the endpoints $a, b$ the profile is at half of its maximum value. From (1.9.109), that interval is simply $[E_0 - \frac{\hbar\Gamma}{2}, E_0 + \frac{\hbar\Gamma}{2}]$ so that the FWHM for the Breit-Wigner profile is:

$$\mathrm{FWHM} = \hbar\Gamma. \tag{1.9.110}$$

For further understanding and an example of how a physical state may reasonably be approximated by a virtual state, I like section 16.6 of [2].

## Exercises

**Problem 1.65: Nonrelativistic Single Scalar Particle Scattering.** Consider a nonrelativistic scalar particle scattering in $d$ spatial dimensions so that its free Hamiltonian operator is $\hat{H}_0 = \frac{\hat{\vec{p}}^2}{2m}$ and so that $\hat{H}_1$ acts on wavefunctions in position space as:

$$\langle \vec{x}|\hat{H}_1|\psi\rangle = V(\vec{x}, -i\nabla)\psi(\vec{x})$$

for some differential operator $V(\vec{x}, -i\nabla)$ which is written by convention so that all of the explicit $\vec{x}$ dependence occurs to the left of explicit $\nabla$ dependence (explain why this can always be done).

(a) Explain that the scattering eigenstates of $\hat{H}_0$ are planewaves. By convention, we normalize them slightly differently than we usual do for momentum eigenstates to shave off some explicit factors of $2\pi$, as:

$$\langle \vec{x}|\alpha_0\rangle = e^{i\vec{x}\cdot\vec{k}}.$$

Explain that with this normalization and (1.9.25), we need the label $\alpha = \frac{\vec{k}}{2\pi}$. Show that the corresponding energy eigenvalue of $|\alpha_0\rangle$ is $E_k = \frac{\hbar^2 k^2}{2m}$ where $k^2 = |\vec{k}|^2$.

(b) Let $\psi_{\vec{k}\pm}(\vec{x}) = \langle \vec{x}|\alpha_\pm\rangle$ be the wavefunction solution to the Lippmann-Schwinger equation. Show that the Lippmann-Schwinger equation becomes:

$$\psi_{\vec{k}\pm}(\vec{x}) = e^{i\vec{k}\cdot\vec{x}} + \frac{2m}{\hbar^2}\int G_\pm(k, \vec{x}, \vec{x}')V(\hat{\vec{x}}', -i\nabla')\psi_{\vec{k}\pm}(\vec{x}')d^d\vec{x}'$$

$$\text{where:} \qquad G_\pm(k,\vec{x},\vec{x}') = \frac{\hbar^2}{2m}\langle\vec{x}|\hat{G}_\pm(E_k)|\vec{x}'\rangle = \lim_{\varepsilon\to 0^+}\int \frac{e^{i\vec{k}'\cdot(\vec{x}-\vec{x}')}}{k^2-k'^2\pm i\varepsilon}\frac{d^d\vec{k}'}{(2\pi)^d}$$

are (up to a convenient normalization) Green's functions of the unperturbed Hamiltonian operator.

(c) Show that $G_+^*(k,\vec{x},\vec{x}') = G_-(k,\vec{x},\vec{x}')$, that $G_\pm(k,\vec{x},\vec{x}') = G_\pm(k,\vec{x}',\vec{x})$.

(d) Let $T_{\vec{k}'\vec{k}} = \langle\beta_0|\hat{T}|\alpha_0\rangle$ with $\alpha = \frac{\vec{k}}{2\pi}$ and $\beta = \frac{\vec{k}'}{2\pi}$. Show that (1.9.54) becomes:

$$T_{\vec{k}'\vec{k}} = \int e^{-i\vec{k}'\cdot\vec{x}}V(\vec{x},-i\nabla)\psi_{\vec{k}+}(\vec{x})d^d\vec{x}.$$

(e) Let $S_{\vec{k}'\vec{k}}$ be defined from $\hat{S}$ the same way $T_{\vec{k}'\vec{k}}$ is defined from $\hat{T}$. Show that (1.9.68) and the unitarity of the S matrix become:

$$S_{\vec{k}'\vec{k}} = (2\pi)^d\delta(\vec{k}-\vec{k}') + \frac{2\pi m}{i\hbar^2|\vec{k}|}T_{\vec{k}'\vec{k}}\delta(|\vec{k}'|-|\vec{k}|), \qquad \int S^*_{\vec{k}''\vec{k}'}S_{\vec{k}''\vec{k}}d^4\vec{k}'' = (2\pi)^{2d}\delta(\vec{k}-\vec{k}').$$

**Problem 1.66: 1 dimensional Scattering.** Consider the same setup as the previous problem, but in 1 spatial dimension, and recall the results of that problem.

(a) Show using partial fraction decomposition and contour integration that:

$$G_+(k,x,x') = \frac{1}{2k}\lim_{\varepsilon\to0^+}\int_{-\infty}^{\infty}\frac{e^{ik'(x-x')}}{k+i\varepsilon-k'}\frac{dk'}{2\pi} + \frac{1}{2k}\lim_{\varepsilon\to0^+}\int_{-\infty}^{\infty}\frac{e^{ik'(x-x')}}{k+i\varepsilon+k'}\frac{dk'}{2\pi} = \frac{e^{i|k(x-x')|}}{2i|k|}.$$

(b) Consider an initially right moving wave (explain why this means $k>0$). Show, using the Lippmann-Schwinger equation, that asymptotically for very large $x$:

$$\psi_{k+}(x) = t(k)e^{ikx} \qquad \text{where:} \qquad t(k) = 1 + \frac{m}{i\hbar^2|k|}\int_{-\infty}^{\infty}e^{-ik\cdot x'}V(x',-i\nabla')\psi_{k+}(x')dx'$$

is called the transmission coefficient of the system, as it gives how strongly the wave transmits in the long run out into the same direction it was originally going. Verify that this formula for the transmission coefficient remains valid for left moving waves.

(c) Consider an initially right moving wave. Show, using the Lippmann-Schwinger equation, that asymptotically for very large $-x$:

$$\psi_{k+}(x) = e^{ikx} + r(k)e^{-ikx} \qquad \text{where:} \qquad r(k) = \frac{m}{i\hbar^2|k|}\int_{-\infty}^{\infty}e^{ikx'}V(x',-i\nabla')\psi_{k+}(x')dx'$$

is called the reflection coefficient of the system, as it gives how strongly part of the wave is reflected toward the opposite of its initial direction. Verify that this formula for the reflection coefficient remains valid for left moving waves.

(d) Show that the reflection and transmission coefficients are related to the T and S matrices by:

$$t(k) = 1 + \frac{m}{i\hbar^2|k|}T_{kk}, \qquad r(k) = \frac{m}{i\hbar^2|k|}T_{-kk}, \qquad S_{k'k} = (2\pi)t(k)\delta(k'-k) + 2\pi r(k)\delta(k'+k).$$

Everything in 1 dimensional scattering is determined by the reflection and transmission coefficients.

(e) Show that unitarity of the S matrix is equivalent to the conditions:

$$|t(k)|^2 + |r(k)|^2 = 1, \qquad t^*(-k)r(k) + r^*(-k)t(k) = 0.$$

**Problem 1.67: 3 dimensional Scattering.** Consider the same setup as the previous problem, but in 3 spatial dimensions, and recall the results of the previous two problems.

(a) Show, by completing the angular integrals in the same way as Orig. (3.9.17), using partial fraction decomposition, and contour integration that:

$$G_+(k,\vec{x},\vec{x}') = \lim_{\varepsilon\to0^+}\frac{1}{2\pi^2|\vec{x}-\vec{x}'|}\int_0^\infty\frac{k'\sin(|\vec{x}-\vec{x}'|k')}{k^2-k'^2+i\varepsilon}dk' = -\frac{e^{ik|\vec{x}-\vec{x}'|}}{4\pi|\vec{x}-\vec{x}'|}.$$

(b) Show that for very large $r=|\vec{x}|$ the solution to the Lippmann-Schwinger equation becomes:

$$\psi_{\vec{k}+}(\vec{x}) = e^{i\vec{k}\cdot x} + f(\vec{k},k\vec{\imath}_r)\frac{e^{ikr}}{r} + \mathcal{O}\left(\frac{1}{r^2}\right)$$

$$\text{where:} \qquad f(\vec{k},\vec{k}') = -\frac{m}{2\pi\hbar^2}T_{\vec{k}'\vec{k}} = -\frac{m}{2\pi\hbar^2}\int e^{-i\vec{k}'\cdot\vec{x}}V(\vec{x},-i\nabla)\psi_{\vec{k}+}(\vec{x})d^3\vec{x}$$

is called the **scattering amplitude** of the process.

(c) Show that the scattering amplitude is related to the S matrix by:

$$S_{\vec{k}'\vec{k}} = (2\pi)^3\delta(\vec{k}-\vec{k}') + \frac{4\pi^2 i}{|\vec{k}|}f(\vec{k},\vec{k}')\delta(|\vec{k}|-|\vec{k}'|).$$

Since they determine everything, 3 dimensional scattering is all about scattering amplitudes.

(d) Let $\Omega$ indicate the unit sphere, $\vec{\iota}(\theta,\varphi)$ indicate a unit direction on $\Omega$, and $d\Omega$ be the usual area element on the sphere. Explain that $f(\vec{k},\vec{k}')$ is only nonzero when $|\vec{k}| = |\vec{k}'|$ and in the case unitarity of the S matrix requires the so-called **Optical Theorem**:

$$f(\vec{k},\vec{k}') - f^*(\vec{k}',\vec{k}) = \frac{i|\vec{k}|}{2\pi}\int_\Omega f^*(\vec{k}',|\vec{k}|\vec{\iota})f(\vec{k},|\vec{k}|\vec{\iota})d\Omega.$$

(e) Just as originally in Orig. problem 3.56, the scattering cross section $\frac{d\sigma}{d\Omega}$ is defined as the ratio of the outgoing current per unit solid angle $J_{\text{out}}$ on the detector sphere to the incoming current per unit cross sectional area $J_{\text{in}}$ of the source beam, i.e. $\frac{d\sigma}{d\Omega} = \frac{J_{\text{out}}}{J_{\text{in}}}$. Show, using (1.6.83), (with $\vec{\iota}_k = \frac{\vec{k}}{|\vec{k}|}$):

$$J_{\text{in}} = \frac{q\hbar}{m}\vec{\iota}_k \cdot \text{Im}(\psi_{\text{in}}^*\vec{\mathscr{D}}\psi_{\text{in}}), \qquad \psi_{\text{in}} = e^{i\vec{k}\cdot x},$$

$$J_{\text{out}} = \frac{q\hbar}{m}\lim_{r\to\infty} r^2\vec{\iota}_r \cdot \text{Im}(\psi_{\text{out}}^*\vec{\mathscr{D}}\psi_{\text{out}}) \qquad \psi_{\text{out}} = f(\vec{k},\vec{k}')\frac{e^{ikr}}{r} + \mathcal{O}\left(\frac{1}{r^2}\right)$$

$$\implies \frac{d\sigma}{d\Omega} = |f(\vec{k},\vec{k}')|^2.$$

Thus, the differential cross section is very cleanly given in theory-land in terms of the scattering amplitude. $J_{\text{in}}$ is set by the design of the particle beam and $J_{\text{out}}$ is measured in the particle detector. Therefore, the differential cross section is directly experimentally observable and an excellent comparison point between theory and experiment, just as in classical mechanics and in electromagnetism.

**Problem 1.68: Scattering through a $\delta$-potential.** Consider a nonrelativistic scalar quantum particle of mass $m$ in one spatial dimension experiencing an attractive Dirac delta potential so that:

$$\hat{H} = \frac{\hat{p}^2}{2m} - \alpha\delta(\hat{x})$$

(a) Show that the plus Lippmann-Schwinger equation for this system becomes:

$$\psi_{k+}(x) = e^{ikx} + \frac{im\alpha}{\hbar^2|k|}e^{i|kx|}\psi_{k+}(0).$$

(b) Show that for $k > 0$ the solutions to the plus Lippmann-Schwinger equation give:

$$\psi_{k+}(x) = \begin{cases} e^{ikx} + r(k)e^{-ikx}, & x < 0 \\ t(k)e^{ikx}, & x > 0 \end{cases} \qquad \text{where:} \qquad r(k) = \frac{im\alpha}{\hbar^2 k - im\alpha}, \qquad t(k) = \frac{\hbar^2 k}{\hbar^2 k - im\alpha}.$$

Show as well by considering negative $k$ that $r$ and $t$ are both even functions of $k$.

(c) Show that the unitarity conditions in problem 1.66 are satisfied by $r$ and $t$.

(d) Show that the S-matrix (look at $t(k)$) has a pole at $k = \frac{im\alpha}{\hbar^2}$ and thus energy $E = \frac{\hbar^2k^2}{2m} = -\frac{m\alpha^2}{2\hbar^2}$, indicating the presence of a single bound state for this system. Show by explicitly solving the eigenvalue problem for $\hat{H}$ that indeed it does have a single bound state and that it is given by:

$$\psi(x) = \sqrt{\frac{m\alpha}{\hbar^2}}e^{-\frac{m\alpha|x|}{\hbar^2}}.$$

**Problem 1.69: Born Approximation and Perturbative Unitarity.** Consider the total Hamiltonian operator to be of the form $\hat{H} = \hat{H}_0 + \lambda\hat{H}_1$ in 3 spatial dimensions and take the perturbative expansion of the scattering amplitude in powers of some parameter $\lambda$ which characterizes the strength of the interaction:

$$f(\vec{k},\vec{k}') = \lambda f^{(1)}(\vec{k},\vec{k}') + \lambda^2 f^{(2)}(\vec{k},\vec{k}') + \mathcal{O}(\lambda^3).$$

(a) Show that the corrections to $f$ are related to the those of the Lippmann-Schwinger equation by:

$$f^{(n)}(\vec{k},\vec{k}') = -\frac{m}{2\pi\hbar^2}\int e^{-i\vec{k}'\cdot\vec{x}}V(\vec{x},-i\nabla)\psi_{\vec{k}+}^{(n-1)}(\vec{x})d^3\vec{x}.$$

with $V$ defined as in problem 1.65.

(b) Show that the optical theorem relates scattering amplitude corrections at different orders by:

$$f^{(n)}(\vec{k},\vec{k}') - f^{(n)*}(\vec{k}',\vec{k}) = \frac{i|\vec{k}|}{2\pi}\int_\Omega \sum_{l=1}^{n-1} f^{(l)*}(\vec{k}',|\vec{k}|\iota) f^{(n-l)}(\vec{k},|\vec{k}|\iota) d\Omega.$$

Explain how this formula, known as the perturbative optical theorem, can be used to provide a helpful consistency check on calculations of the scattering amplitude.

**Problem 1.70: Forward Scattering and the Optical Theorem.** Forward scattering is the limit that the outgoing wave is aligned with the incoming beam. That is, $\vec{k}' \to \vec{k}$ in $f(\vec{k},\vec{k}')$. Show, using the optical theorem that the total cross section for scattering of an incoming beam of wavenumber $\vec{k}$ is given in terms of the imaginary part of the forward scattering amplitude by:

$$\sigma = \int_\Omega \frac{d\sigma}{d\Omega} d\Omega = \frac{4\pi}{|\vec{k}|}\text{Im}(f(\vec{k},\vec{k})).$$

**Problem 1.71: First Born Scattering.** Consider a scalar particle scattering in 3 spatial dimensions with the same conventions as problem 1.65. Let $\vec{q} = \vec{k}' - \vec{k}$ be the "wavenumber transfer".

(a) Show that at lowest order in perturbation theory/in the first Born approximation:

$$f^{(1)}(\vec{k},\vec{k}') = -\frac{m}{2\pi\hbar^2}\widetilde{V}(\vec{k}'-\vec{k},\vec{k}) \qquad \text{where:} \qquad \widetilde{V}(\vec{q},\vec{k}) = \int e^{-i\vec{q}\cdot\vec{x}} V(\vec{x},\vec{k}) d^3\vec{x}$$

is the Fourier transform of $V$ with respect to its first argument, normalized without any $2\pi$'s.

(b) Suppose that $V$ is spherically symmetric so that it is a function of $r$, $\nabla^2$ only, i.e. $V = V(r,\sqrt{-\nabla^2})$. Show that the Fourier transform can be reduced to a single integral (where $k = |\vec{k}|$):

$$\widetilde{V}(\vec{q},\vec{k}) = \frac{4\pi}{|\vec{q}|}\int_0^\infty r V(r,k)\sin(|\vec{q}|r) dr.$$

(c) Show, recalling that $\vec{k}' = k\vec{n}(\theta,\varphi)$ with $\theta$ the scattering angle, that:

$$|\vec{q}| = 2k\sin\frac{\theta}{2}$$

and show that to lowest order in perturbation theory the differential cross section for scattering of an energy $E = \frac{\hbar^2k^2}{2m}$ per particle incoming beam by a spherically symmetric potential with can be written:

$$\frac{d\sigma}{d\Omega} = \frac{1}{4E^2\sin^2\frac{\theta}{2}}\left(\int_0^\infty r V(r,k)\sin\left(2kr\sin\frac{\theta}{2}\right)k dr\right)^2 + \text{H.O.T.}$$

**Problem 1.72: First look at the Yukawa Potential.** The effective potential between some particles is described by the so-called Yukawa potential, which takes the form:

$$V(r) = -\frac{A}{r}e^{-\kappa r}$$

for constants $A$ and $\kappa > 0$. This is like an exponentially suppressed Coulomb potential. We will derive it from more fundamental considerations in the coming chapters.

(a) Show that:

$$\int_0^\infty r V(r)\sin\left(2kr\sin\frac{\theta}{2}\right)k dr = -\frac{2Ak^2\sin\frac{\theta}{2}}{\kappa^2 + 4k^2\sin^2\frac{\theta}{2}}.$$

(b) From this, show that the differential and total cross sections at lowest order are given by:

$$\frac{d\sigma}{d\Omega} = \left(\frac{A}{4}\right)^2 \frac{1}{\left(\frac{\hbar^2\kappa^2}{8m} + E\sin^2\frac{\theta}{2}\right)^2} + \text{H.O.T.}, \qquad \sigma = \frac{2\pi m A^2}{\hbar^2\kappa^2}\frac{1}{E + \frac{\hbar^2\kappa^2}{8m}} + \text{H.O.T.}$$

By fitting data to these functional forms for different $E$ and $\theta$, one can measure $A$ and $\kappa$.

**Problem 1.73: ⋆ Rutherford Scattering Revisited.** As you have become well aware, an electron interacting with a single proton at distances much further away than the proton size far from the experiences the perturbation:

$$V(r) = -\frac{q_e^2}{4\pi\epsilon_0 r}.$$

Because the Coulomb potential decays very slowly with distance, we will have to treat the large $r$ behavior of integrals with slight care. Explain that the Coulomb potential can be treated as the $\kappa \to 0^+$ limit of the Yukawa potential, and from this show that the differential cross section for Coulomb scattering at lowest order in perturbation theory is given by:

$$\frac{d\sigma}{d\Omega} = \left(\frac{q_e^2}{16\pi\epsilon_0 E}\right)^2 \frac{1}{\sin^4\frac{\theta}{2}}.$$

Note that this exactly matches the classical Coulomb cross section from Orig. problem 3.58. In fact, surprisingly even though we derived this at leading order in perturbation theory, it is the exact quantum result [35].

**Problem 1.74: Higher Energy Scattering and the Eikonal Formula.** Consider 3 dimensional scalar particle scattering with the same conventions as problem 1.65 and with $V = U(\vec{x})$.

(a) Using that $|\psi_+\rangle$ is an eigenvector of $\hat{H}$ with eigenvalue $\frac{\hbar^2 k^2}{2m}$ and decomposing as $\psi_+(\vec{x}) = \sqrt{n(\vec{x})}e^{\frac{i}{\hbar}S(\vec{x})}$ for real $n$, $S$ akin to problem 1.30. Show that $n$ and $S$ exactly satisfy:

$$\frac{|\nabla S|^2}{2m} + U(\vec{x}) - \frac{\hbar^2\nabla^2\sqrt{n}}{2m\sqrt{n}} = \frac{\hbar^2 k^2}{2m}, \qquad \nabla\cdot(n\nabla S) = 0.$$

(b) Now begin to consider the high energy (large $k$) behavior of the scattering. Show that $n$ and $S$ can be given series expansions in $\frac{1}{k^2}$, with $\vec{b} = \vec{x} - \vec{\iota}_k\vec{\iota}_k\cdot\vec{x}$ called the "impact parameter" and $z = \vec{\iota}_k\cdot\vec{x}$, as:

$$n = 1 + \mathcal{O}\left(\frac{1}{k^2}\right), \quad S = \hbar\vec{k}\cdot\vec{x} + \frac{1}{k}S_1(\vec{x}) + \mathcal{O}\left(\frac{1}{k^3}\right) \quad \text{where:} \quad S_1(\vec{x}) = -\frac{m}{\hbar}\int_{-\infty}^{z} U(\vec{b}+s\vec{\iota}_k)ds.$$

(c) Show that in the high energy limit, for finitely large scattering angle $\theta$ the amplitude $f$ goes to 0 due to its integral expression oscillating wildly. Show that if the wavenumber transfer $\vec{q}$ is kept finite then at high energies $\theta$ is small and:

$$f(\vec{k},\vec{k}+\vec{q}) \approx \frac{k}{2\pi i}\int e^{-i\vec{q}\cdot\vec{b}}\left(e^{i\chi(\vec{b},\vec{k})}-1\right)d^2\vec{b} \qquad \text{where:} \qquad \chi(\vec{b},\vec{k}) = -\frac{1}{\hbar}\int_{-\infty}^{\infty} U\left(\vec{b}+\frac{\hbar\vec{k}}{m}t\right)dt$$

is called the eikonal phase. Explain how $\hbar\chi$ is exactly the high energy limit of the interaction picture classical on-shell action for a particle with initial momentum $\hbar\vec{k}$.

**Problem 1.75: Spherically Symmetric Scattering.** Consider 3 dimensional spherically symmetric scalar particle scattering with the same conventions as problems 1.65 and 1.71(b).

(a) Demonstrate that the scattering amplitude $f(\vec{k},\vec{k}')$ is a function only of $k = |\vec{k}|$ and $\vec{\iota}_k\cdot\vec{\iota}_{k'}$ where $\vec{\iota}_k = \frac{\vec{k}}{k}$ and $\vec{\iota}_{k'} = \frac{\vec{k}'}{k}$. From the completeness of the Legendre polynomials, show that for some $f_l(k)$:

$$f(\vec{k},\vec{k}') = \sum_{l=0}^{\infty}(2l+1)f_l(k)P_l(\vec{\iota}_k\cdot\vec{\iota}_{k'}).$$

(b) Show, using the spherical harmonic projection theorem from 1.54 and the optical theorem from 1.67(d), that unitarity of the S matrix is equivalent to the $f_l(k)$ satisfying:

$$f_l(k) - f_l^*(k) = 2ik|f_l(k)|^2.$$

(c) Show that $f_l(k)$ satisfy the above if and only if there exist real functions $\delta_l(k)$, called the partial wave phase shifts, so that $kf_l(k) = e^{i\delta_l}\sin\delta_l$. Then, show that $f(\vec{k},\vec{k}')$ and the total cross section are:

$$f(\vec{k},\vec{k}') = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)e^{i\delta_l}\sin\delta_l P_l(\vec{\iota}_k\cdot\vec{\iota}_{k'}) \implies \sigma = \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l.$$

(d) Assume that simply $V = U(r)$. Show, using the projection theorem from 1.54 that $G_+(\vec{k},\vec{x},\vec{x}')$ as in problem 1.65 can be written in terms of spherical harmonics, the spherical Bessel functions, and the spherical Hankel functions as:

$$G_+(k,\vec{x},\vec{x}') = -ik\sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}(\theta,\varphi)Y_{lm}^*(\theta',\varphi')h_l(k\max(r,r'))j_l(k\min(r,r')).$$

To do so, the following integral fact will be useful, for $0 < \alpha < 1$ and real $\beta$:

$$\lim_{\varepsilon \to 0^+} \int_0^\infty \frac{j_l(x) j_l(\alpha x) x^2 dx}{\beta^2 - x^2 + i\varepsilon} dx = \frac{\pi \beta}{2i} h_l(\beta) j_l(\alpha \beta).$$

For extra challenge, prove this integral (using contour integration and the results of problem 1.58).

(e) Show that for spherically symmetric scattering we can take the $|\alpha_0\rangle$, for some unimportant normalization $N$, as below and that the resulting $|\alpha_+\rangle$ states then take the form:

$$\langle \vec{x} | \alpha_0 \rangle = N Y_{lm}(\theta, \varphi) j_l(kr), \qquad \langle \vec{x} | \alpha_+ \rangle = N Y_{lm}(\theta, \varphi) R_l(k, r)$$

for some $R_l(k, r)$. From these and the above form of the Green's function, show that the spherically symmetric Lippmann-Schwinger equation becomes:

$$R_l(k, r) = j_l(kr) - \frac{2imk}{\hbar^2} \int_0^\infty r'^2 U(r') h_l(k \max(r, r')) j_l(k \min(r, r')) R_l(k, r') dr'$$

and that the partial wave phase shifts are given by:

$$e^{\delta_l} \sin \delta_l = -\frac{2mk}{\hbar^2} \int_0^\infty r^2 U(r) j_l(kr) R_l(k, r) dr.$$

(f) Using the above formula for the phase shift, show that the leading behavior of $\delta_l$ at low energies (small $k$) is proportional to $k^{2l+1}$. Thus, low $l$ wave scattering dominates at low energies.

**Problem 1.76: Hard Sphere Scattering.** Consider 3 dimensional scalar particle scattering with the same conventions as the previous problem and with $V$ giving an impenetrable sphere at $r = \mathfrak{r}$ for a constant $\mathfrak{r}$.

(a) Show that the exact scattering radial wavefunction $R_l(r)$ and the partial wave phase shifts are given by:

$$R_l(r) = \frac{\mathrm{Im}(h_l^*(k\mathfrak{r}) h_l(kr))}{i h_l(k\mathfrak{r})}, \qquad \tan \delta_l = \frac{j_l(k\mathfrak{r})}{y_l(k\mathfrak{r})}.$$

(b) Show that in the low energy limit $k \to 0$, the total cross section becomes $\sigma = 4\pi \mathfrak{r}^2$, the full surface area of the sphere. Conceptually explain why this is the result.

(c) Show that in the high energy limit $k \to \infty$, the total cross section becomes $\sigma = 2\pi \mathfrak{r}^2$, twice the classical result of $\pi \mathfrak{r}^2$. There is no regime in which this system gives the classical answer. Quantum interference effects are genuinely fundamentally important in scattering at all energies.

**Problem 1.77: Electric Dipole Scattering.** Consider the scattering of a charged particle $q$ by an electric dipole potential:

$$\phi(\vec{x}) = \frac{\vec{\wp} \cdot \vec{x}}{4\pi \epsilon_0 |\vec{x}|^3}.$$

(a) Show that in the first Born approximation, the scattering amplitude for the charged particle is given by:

$$f^{(1)}(\vec{k}, \vec{k}') = \frac{imq}{2\pi \hbar^2 \epsilon_0} \frac{\vec{\wp} \cdot (\vec{k}' - \vec{k})}{|\vec{k}' - \vec{k}|^2}.$$

(b) Let the plane spanned by the incoming particle's momentum and the electric dipole moment be the $xz$-plane. Choose the $z$ direction as that of the incoming particle's momentum (which has magnitude $\hbar k$). Show that for constants $\wp$ and $\alpha$ the electric dipole moment can be written as:

$$\vec{\wp} = \wp(\sin \alpha, 0, \cos \alpha)$$

and show that to leading order the differential cross section of the charge's scattering becomes:

$$\frac{d\sigma}{d\Omega} = \left( \frac{mq\wp}{4\pi \epsilon_0 \hbar^2 k} \right)^2 \left( \sin^2 \alpha \cos^2 \varphi \frac{\cos^2 \frac{\theta}{2}}{\sin^2 \frac{\theta}{2}} - \sin(2\alpha) \cos \varphi \frac{\cos \frac{\theta}{2}}{\sin \frac{\theta}{2}} + \cos^2 \alpha \right) + \mathcal{O}(q^3).$$

Note that the total cross section is infinite unless the beam and the dipole moment are perfectly aligned.

**Problem 1.78: Magnetic Dipole Scattering and measuring the electron dipole moment.** Consider the scattering of a spin-1/2 charged particle $q$ of gyromagnetic ratio $\gamma$ by a magnetic dipole potential:

$$\vec{A}(\vec{x}) = \frac{\mu_0 \vec{M} \times \vec{x}}{4\pi |\vec{x}|^3}.$$

(a) Explain that the eigenstates of the free Hamiltonian are labeled by $\vec{k}$ and $\pm$ where $\pm$ indicate whether the spin of the planewave is aligned or anti-aligned with a particular axis of choice. For this question, choose that axis to be the direction $\vec{M}$. Explain that the scattering amplitude is then a $2\times 2$ matrix $f_{\pm\pm}(\vec{k},\vec{k}')$ where each $\pm$ indicates the spin alignment of the incoming or outgoing plane wave. Let $\xi_\pm$ be constant unit spinors with spins aligned and anti-aligned with $\vec{M}$. Show that:

$$\psi_{(\vec{k},\pm_{\text{in}})+}(\vec{x}) = e^{i\vec{k}\cdot\vec{x}}\xi_{\pm_{\text{int}}} + \frac{e^{ikr}}{r}\left(f_{\pm_{\text{in}}+_{\text{out}}}\xi_+ + f_{\pm_{\text{in}}-_{\text{out}}}\xi_-\right) + \mathcal{O}\left(\frac{1}{r^2}\right).$$

(b) Show that the magnetic field produced by the given vector potential is:

$$\vec{B}(\vec{x}) = \frac{2}{3}\mu_0\vec{M}\delta(\vec{x}) + \frac{\mu_0}{4\pi r^3}(3\vec{\imath}_r\vec{\imath}_r\cdot\vec{M} - \vec{M}).$$

To do so, recall that $\nabla^2\frac{1}{r} = -4\pi\delta(\vec{x})$. Show that in the first Born approximation, the scattering amplitude for the charged particle is given by:

$$f^{(1)}_{\pm+}\xi_+ + f^{(1)}_{\pm-}\xi_- = \frac{i\mu_0 q}{2\pi\hbar}\frac{\vec{M}\cdot(\vec{k}\times\vec{k}')}{|\vec{k}'-\vec{k}|^2}\xi_\pm - \frac{\mu_0 m\gamma}{2\pi\hbar^2}\left((\vec{\imath}\cdot\hat{\vec{S}}\xi_\pm)\vec{M}\cdot\vec{\imath} - \vec{M}\cdot\hat{\vec{S}}\xi_\pm\right)$$

where $\vec{\imath}$ is the unit vector in the direction of the momentum transfer, i.e. $\vec{k}'-\vec{k} = |\vec{k}'-\vec{k}|\vec{\imath}$.

(c) Let the plane spanned by the incoming particle's momentum and the magnetic dipole moment be the $xz$-plane. Choose $z$ direction as that of the incoming particle;s momentum (with magnitude $\hbar k$). Show that for constants $M$ and $\alpha$ the electric dipole moment can be written:

$$\vec{M} = M(\sin\alpha, 0, \cos\alpha)$$

and show that to leading order in perturbation theory the differential cross section of the charge's scattering (recalling the definition of the $g$-factor) becomes:

$$\frac{d\sigma}{d\Omega} = \left(\frac{\mu_0 qM}{4\pi\hbar}\right)^2\left(\frac{\sin^2\alpha\sin^2\varphi}{\tan^2\frac{\theta}{2}} + \left(\frac{g}{2}\right)^2\left(\cos^2\frac{\theta}{2}\sin^2\varphi + \left(\cos\alpha\cos\frac{\theta}{2}\cos\varphi - \sin\alpha\sin\frac{\theta}{2}\right)^2\right)\right) + \mathcal{O}(M^3).$$

Show that in the limit that the incoming beam is aligned with the magnetic dipole moment of the dipole, the cross section simply becomes:

$$\frac{d\sigma}{d\Omega} = \left(\frac{\mu_0 gqM}{8\pi\hbar}\right)^2\cos^2\frac{\theta}{2} + \mathcal{O}(M^3).$$

Explain how the angular distribution of scattering particles can be used to determine the values of $q$ and of the $g$-factor if the dipole moment strength $M$ of the source dipole is known. Thus, using scattering experiments it is possible to measure $g$ for charged particles, such as the electron.

**Problem 1.79: Scattering with time reversal symmetry.** Consider scattering in a system with time reversal symmetry by a time reversal pseudo-operator $\hat{\mathcal{T}}$ (both the free and full Hamiltonian operator are to have symmetry under this transformation). Suppose the unperturbed eigenstates $|\alpha_0\rangle$ have been chosen to be time reversal symmetric:

$$\hat{\mathcal{T}}|\alpha_0\rangle = |\alpha_0\rangle.$$

(a) Show that the time evolution operator satisfies:

$$\hat{\mathcal{T}}\hat{U}(t,t_0)\hat{\mathcal{T}}^{-1} = \hat{U}^\dagger(-t_0,-t).$$

Show further that this implies its matrix elements satisfy:

$$\langle\beta_0|\hat{U}(t,t_0)|\alpha_0\rangle = \langle\alpha_0|\hat{U}(-t_0,-t)|\beta_0\rangle.$$

(b) Show that the S matrix elements, transition matrix elements, and transition rates satisfy:

$$S_{\beta\alpha} = S_{\alpha\beta}, \qquad T_{\beta\alpha} = T_{\alpha\beta}, \qquad \Gamma_{\alpha\to\beta} = \Gamma_{\beta\to\alpha}.$$

(c) Show that for 3 dimensional scattering, the scattering amplitude becomes a symmetric function $f(\vec{k},\vec{k}') = f(\vec{k}',\vec{k})$ and that the optical theorem becomes:

$$\text{Im}f(\vec{k},\vec{k}') = |\vec{k}|\int_\Omega f^*(\vec{k}',|\vec{k}|\vec{\imath})f(\vec{k},|\vec{k}|\vec{\imath})\frac{d\Omega}{4\pi}.$$

**Problem 1.80: H-theorem for Fermi's master equation.**

(a) Using the Born rule, explain that $\rho_n(t)$ in the Fermi master equation gives the probability that if the

system is ideally measured at the time $t$ it will be approximately in the state $|n\rangle$ after the measurement.

(b) Let $S(t)$ be the entropy of the distribution of measurement results at time $t$:

$$S(t) = -\sum_n \rho_n \ln \rho_n.$$

Show that $S(t)$ satisfies:

$$\frac{dS}{dt} = \sum_n \sum_{m \neq n} \Gamma_{m \to n} (\rho_n - \rho_m) \ln \rho_n.$$

(c) Suppose the system has time reversal symmetry. Show that $S(t)$ satisfies:

$$\frac{dS}{dt} = \frac{1}{2} \sum_n \sum_{m \neq n} \Gamma_{m \to n} (\rho_n - \rho_m)(\ln \rho_n - \ln \rho_m).$$

Show that $\frac{dS}{dt} > 0$ for all distributions which are not uniform and that $\frac{dS}{dt} = 0$ for a uniform distribution. Finally then, conclude that under time symmetric evolution by the Fermi master equation, the equilibrium distribution of $\rho_n$ is necessarily uniform. This is known as Boltzmann's H-theorem. Explain how this is consistent with unitary invertible time evolution because the Fermi master equation involves the consistent ignoring of information by neglecting the phases conjugate to the $\rho_n$.

## 1.10 Another Big Bold Leap

### 1.10.1 Trouble with Multi-particle Statistics

We already know that Galilean Quantum Mechanics cannot be the end of our story, since Galilean symmetry isn't fundamental. However, the wave theory of matter side of what we have done is extremely successful and was able to address the problem of point-particle divergences (the electromagnetic field energy stored in a finite size charge configuration never needs the divergent point limit). As such, the next natural thing to do is to consider a wave theory of matter with Lorentz symmetry, and we will do this. But, there are other important statistical problems with single particle Galilean quantum mechanics that would persist in a Lorentz symmetric wave theory of matter. These problems are statistical in the sense that they all deal with the properties of multi particle systems.

We said, without sufficient justification (with the details delayed until chapter 2), that ideal measurements only let us access the state of the quantum system statistically due to its microscopic nature (the results of this are summarized by the Born rule and the Collapse rule). We need to understand how quantum systems interact in order to demonstrate this properly, so we need to study more than one particle. As well, there exist experimental processes in which the number of particles change, such as when one decays into multiple. There is no way to account for this within classical mechanics and the quantum mechanics we have done so far this chapter does nothing to solve it. With the wave theory, one might believe that such a process could be explained by considering a wave with multiple separated wave-packets and identifying each wave-packet as a distinct particle. However, this has some problems. In particular, there exist processes in which something charge neutral separates into two particles, one of positive charge and one of negative charge. By our expression for the charge density in (1.6.83), the sign of the total charge matches the sign of the charge density everywhere. Therefore, there is no way a charge neutral wave-packet could ever have non-zero charge density in this theory when it splits into different wave-packets. Consequently multiple wave-packets within one wavefunction cannot explain this phenomenon. The only alternative within our framework at this time then is that each particle gets its own wavefunction. But, there is no way to produce more wavefunctions from one within our theory so far. Seemingly the number of wavefunction is always conserved.

Another issue is that experimentally there exist "indistinguishable" particles, such as electrons. All electrons we have ever observed have produced, to all precision we can measure, the *exact* same electric charge, mass, gyromagnetic ratio, and all other physical properties. It seems to us experimentally impossible to tell the difference between any two electrons. This is quite unlike the case in classical physics where we certainly get objects with similar properties, but never apparently exact duplicates. In the wave theory of matter as we have studied it so far, each electron would have a wavefunction. If each electron was represented by an independent function from the others, it would theoretically be allowed to have independent physical properties. It seems like a massive unexplained coincidence then that all electrons produce the same values of

these parameters. One could assume such a massive coincidence has simply occurred, but it is more than a bit suspicious; such an alignment of parameter values generally in the history of science has indicated that there is an underlying pattern to explain it, that there is some underlying mechanism that *requires* all electrons to be identical. So, we suspect there is something deeper going on to explain the existence of indistinguishable particles which is unexplained by the theory we have discussed so far.

There are further problems still with simply treating each particle as having its own wavefunction. Consider a spin-1/2 particle with wavefunction $\psi(t,\vec{x})$. We can define the functions:

$$\psi(t,\vec{x}) = \begin{pmatrix}\psi_+(t,\vec{x})\\ \psi_-(t,\vec{x})\end{pmatrix}, \qquad \phi(t,\vec{x}) \stackrel{\text{def}}{=} \sqrt{|\psi_+|^2+|\psi_-|^2}, \qquad \chi_\pm \stackrel{\text{def}}{=} \frac{\psi_\pm}{\phi}, \qquad \chi \stackrel{\text{def}}{=} \begin{pmatrix}\chi_+\\ \chi_-\end{pmatrix} \tag{1.10.1}$$

$$\implies \psi(t,\vec{x}) = \begin{pmatrix}\chi_+(t,\vec{x})\phi(t,\vec{x})\\ \chi_-(t,\vec{x})\phi(t,\vec{x})\end{pmatrix} = \chi(t,\vec{x})\phi(t,\vec{x}). \tag{1.10.2}$$

Because $|\chi_+|^2+|\chi_-|^2=1$, it is possible to define angles $\alpha(t,\vec{x}),\theta(t,\vec{x}),\varphi(t,\vec{x})$ so that:

$$\chi_+ = e^{i\alpha}e^{-\frac{i}{2}\varphi}\cos\frac{\theta}{2}, \qquad \chi_- = e^{i\alpha}e^{\frac{i}{2}\varphi}\sin\frac{\theta}{2}. \tag{1.10.3}$$

In the classical limit, $|\psi|^2=\phi^2(t,\vec{x})$ becomes localized and so the only values of $\alpha,\varphi,\theta$ which effect experiments are the values near the classical position. That is, in that limit we approximately can treat $\alpha,\varphi,\theta$ as functions of time only with their values being given by evaluating their spatial dependence at $\vec{x}_{\text{cl}}(t)$. Then, the spin angular momentum carried by the wavefunction becomes:

$$\langle\hat{\vec{S}}\rangle = \langle\psi|\hat{\vec{S}}|\psi\rangle = \int_{\mathbb{S}}\psi^\dagger\hat{\vec{S}}\psi d^3\vec{x} \approx \int_{\mathbb{S}}\phi^2(t,\vec{x})\chi^\dagger(t)\hat{\vec{S}}\chi(t)d^3\vec{x} = \int_{\mathbb{S}}|\psi|^2d^3\vec{x}\langle\chi|\hat{\vec{S}}|\chi\rangle = \langle\chi|\hat{\vec{S}}|\chi\rangle. \tag{1.10.4}$$

In particular, a simple calculation gives:

$$\vec{\mathrm{S}} = \langle\chi|\hat{\vec{S}}|\chi\rangle = \frac{\hbar}{2}(\sin\theta\cos\varphi,\sin\theta\sin\varphi,\cos\theta) \implies |\vec{\mathrm{S}}| = \frac{\hbar}{2}. \tag{1.10.5}$$

So a spin-1/2 particle in the classical limit always carries a total spin angular momentum of exactly magnitude $\frac{\hbar}{2}$. Now consider an experiment in which two spin-1/2 particles with wavefunctions essentially in the classical limit interact. After the interaction we measure a component of $\vec{\mathrm{S}}$ for one of the particles using a Stern-Gerlach device. If we repeat this experiment many times with the same setup of the initial particles (let's say we have polarized incoming beams for each) using multiple different orientations of the Stern-Gerlach device for the particle we measure, we can eventually get good statistics to measure $\langle\hat{\vec{S}}\rangle$ for one of the particles. If we have succeeded in our experimental setup in making it reasonably precise enough that every trial can be treated with the same wavefunction for that particle, then $|\langle\hat{\vec{S}}\rangle|$ must be $\frac{\hbar}{2}$. However, it is an experimental fact that it is possible in this setup for $|\langle\hat{\vec{S}}\rangle| \leq \frac{\hbar}{2}$ and this can be pushed all the way down to 0 for the right intermediate interaction. There is no wavefunction for a spin-1/2 particle that could produce such a result, and the imprecision of the experimental setup is not significant enough to explain the discrepancy. Therefore, it cannot simply be correct to describe each particle in our theory by a different and completely separate wavefunction. This can be explained by the outgoing particles after the interaction having a beam which is unpolarized (and thus requiring a density matrix) but then it needs to be understood how the interaction of two polarized beams could produced an unpolarized beam, which simply solving equations of motion for interacting but separate wavefunctions cannot give.

These statistical problems have essentially nothing to do with Galilean versus Lorentz symmetry, and so demand still some further theoretical insight to solve them. While single particle quantum mechanics successfully describes many microscopic features, we need something superior to formulate multi-particle systems. We can solve this by remembering the other fundamental statistical problem we have so far left unsolved: the ultraviolet catastrophe from Orig. section 5.12.

### 1.10.2 Quantum Field Theory to the Rescue

We showed in detail in Orig. section 5.12 how electromagnetism can decomposed into an infinite number of harmonic oscillators and that when those harmonic oscillators are treated with classical statistical mechanics the resulting thermodynamic quantities all diverge and give unphysical results. The classical density of states

of a harmonic oscillator of angular frequency $\omega$ is:

$$g_{\text{cl}}(\text{E}) = \frac{\vartheta(\text{E})}{\hbar_{\text{sm}}\omega} \tag{1.10.6}$$

where $\hbar_{\text{sm}}$ is used here as the normalization constant introduced in the definition of the phase space density in Orig. (5.2.4). This density of states does produce an experimentally accurate spectral radiance $R(T,\omega)$ at large temperatures (compared to $\hbar_{\text{sm}}\omega$) but it produces an experimentally inaccurate one at low temperatures. Instead, we found in Orig. (5.12.112) that the experimental evidence predicts the density of states:

$$g_{\text{Planck}}(\text{E}) = \sum_{n=0}^{\infty} \delta(\text{E} - n\hbar_{\text{sm}}\omega). \tag{1.10.7}$$

We can understand how this density of states can sometimes be approximated by the $g_{\text{cl}}$ by noting that if it is integrated against a function $f(\text{E})$ which varies very slowly over a change in energy of $\hbar_{\text{sm}}\omega$ then on each such interval it can be replaced by its average value over that interval:

$$\int_{-\infty}^{\infty} f(\text{E}) g_{\text{Planck}}(\text{E}) d\text{E} = \sum_{n=0}^{\infty} f(n\hbar_{\text{sm}}\omega) \approx \sum_{n=0}^{\infty} \frac{1}{\hbar_{\text{sm}}\omega} \int_{n\hbar_{\text{sm}}\omega}^{(n+1)\hbar_{\text{sm}}\omega} f(\text{E}) d\text{E} = \int_{-\infty}^{\infty} f(\text{E}) g_{\text{cl}}(\text{E}) d\text{E}. \tag{1.10.8}$$

Therefore, if $\hbar_{\text{sm}}\omega$ is small compared to the scale on which relevant functions change, we can approximate $g_{\text{Planck}} \approx g_{\text{cl}}$ as distributions.

Now that we have studied a bit of quantum mechanics, $g_{\text{Planck}}$ is very suggestive. We know it is closely related to the harmonic oscillator (as electromagnetism is an infinite number of harmonic oscillators). As well, we found in problem 1.33 that the quantum harmonic oscillator written as having Hamiltonian operator $\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2$ has energy eigenvalues $E_n = \hbar\omega\left(n + \frac{1}{2}\right)$. Of course, as always, the zero-point of the energy can be set to be whatever you like by canonical transformation and is unphysical[16], so we could equally well use the Hamiltonian operator $\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 - \frac{\hbar\omega}{2}\mathbb{1}$, which then has energy eigenvalues $E_n = n\hbar\omega$. Consider the spectral density of the quantum harmonic oscillator in this gauge, meaning the function which when integrated over an interval $(a,b]$ gives the number of eigenvalues of the Hamiltonian operator between $a$ and $b$. Because $E_n = n\hbar\omega$, that spectral density is precisely $g_{\text{Planck}}(\text{E})$ where we identify $\hbar_{\text{sm}} = \hbar$, the actual quantum mechanical $\hbar$. That is, **the experimentally correct density of states for electromagnetism of a particular frequency at low temperatures is that of a quantum harmonic oscillator**.

With all of these observations, we make a big bold leap, perhaps even bigger than that from Galilean symmetry to Lorentz symmetry done in Orig. section 3.9, that even more fundamental than the classical field theory of electromagnetism is a **quantum field theory of electromagnetism**. What the hell is quantum field theory? In single particle quantum mechanics (also called first quantization), we replaced our theory of a classical point particle of position $\vec{x}(t)$ with a theory of a complex wavefunction $\psi(t,\vec{x})$, a time dependent function on the space of positions. Our big leap, called quantum field theory (or second quantization) is to now similarly replace the theory of a classical vector potential $A_\mu(t,\vec{x})$ by a theory of a complex wavefunction(al) $\Psi(t)[A]$, a time dependent functional on the space of vector potential configurations. This looks crazy! Motivating reason to do it right now is to explain the observed black-body radiation, and that requires that electromagnetism should be a bunch of quantum harmonic oscillators. However, it turns out that there are countless phenomena producible in experiments that require this theory of quantum electromagnetism. We will give (in my opinion) theoretically more satisfying arguments where quantum field theory comes from in the next chapter. In quantum field theory (QFT), the wavefunctional state vector $|\Psi(t)\rangle$ is our degree of freedom satisfying $\langle\Psi|\Psi\rangle = 1$, and there is some Hamiltonian operator $\hat{H}$ on the Hilbert space of such state vectors, so that the time evolution of the theory is derived from the Lagrangian:

$$\mathcal{L} = i\hbar\langle\Psi|\dot{\Psi}\rangle - \langle\Psi|\hat{H}|\Psi\rangle. \tag{1.10.9}$$

That is, QFT is just quantum mechanics with the Hilbert space being the space of functionals of field configurations of some fields.

For now, for quantum electromagnetism we will proceed by analogy to what has been done this chapter. First we need to recall the results of Orig. section 4.4 and its homework questions, specifically Orig. problem

[16] Some readers may believe that the zero-point of energy is physical and can be observed, for example in the Casimir effect. This is incorrect. It has been understood for a half-century [56] that the Casimir effect can be derived fully without reference to zero-point energy, does not measure zero-point energy, and that in calculations that make it look like it is doing so it is because one is being sloppy. It is measuring the difference between two simultaneously imposed conflicting conventions on setting the zero-point energy and so is a measure of an energy difference.

4.42. Classical electromagnetism is described by the Maxwell Hamiltonian in Orig. theorem 4.4:

$$\mathcal{H} = \int_{\mathbb{S}} \left( \frac{|\vec{P}_A|^2 + |\nabla \times \vec{A}|^2}{2} - \vec{J} \cdot \vec{A} + \phi(\rho + \nabla \cdot \vec{P}_A) \right) d\mathbb{S} \tag{1.10.10}$$

which has $\phi$ as a Lagrange multiplier enforcing Gauss' Law as a first class constraint. In Fourier space this Hamiltonian becomes:

$$\mathcal{H} = \int_{\mathbb{K}} \left( \frac{|\widetilde{\vec{P}}_A(t,\vec{k})|^2 + |i\vec{k} \times \widetilde{\vec{A}}(t,\vec{k})|^2}{2} - \widetilde{\vec{J}}^*(t,\vec{k}) \cdot \widetilde{\vec{A}}(t,\vec{k}) + \widetilde{\phi}^*(t,\vec{k}) \left( \widetilde{\rho}(t,\vec{k}) + i\vec{k} \cdot \widetilde{\vec{P}}_A(t,\vec{k}) \right) \right) d^3\vec{k}. \tag{1.10.11}$$

Let $\vec{\xi}_m(\vec{k})$ be the polarization vectors defined in Orig. problem 4.42. A Fourier transformed vector field $\widetilde{\vec{F}}$ can be decomposed into the helicity basis according to $\widetilde{F}_m = \vec{\xi}_m^\dagger \widetilde{\vec{F}}$. In the helicity basis the electromagnetic Hamiltonian becomes:

$$\mathcal{H} = \int_{\mathbb{K}} \left( \sum_{h \in \{-1,1\}} \left( \frac{|\widetilde{P}_{Ah}|^2 + |\vec{k}|^2 |\widetilde{A}_h|^2}{2} - \widetilde{J}_h^* \widetilde{A}_h \right) + \frac{|\widetilde{P}_{A0}|^2}{2} - \widetilde{J}_0^* \widetilde{A}_0 + \widetilde{\phi}^* \left( \widetilde{\rho} + i|\vec{k}| \widetilde{P}_{A0} \right) \right) d^3\vec{k} \tag{1.10.12}$$

with the Poisson brackets and gauge properties:

$$\{\widetilde{A}_m(t,\vec{k}), \widetilde{A}_{m'}(t,\vec{k}')\} = 0, \qquad \{\widetilde{A}_m(t,\vec{k}), \widetilde{P}_{Am'}(t,\vec{k}')\} = \delta_{mm'}\delta(\vec{k}+\vec{k}'), \qquad \{\widetilde{P}_{Am}(t,\vec{k}), \widetilde{P}_{Am'}(t,\vec{k}')\} = 0 \tag{1.10.13}$$

$$\widetilde{\phi}' = \widetilde{\phi} - \frac{\partial \widetilde{f}}{\partial t}, \qquad \widetilde{A}'_{-1} = \widetilde{A}_{-1}, \qquad \widetilde{A}'_0 = \widetilde{A}_0 + i|\vec{k}|\widetilde{f}, \qquad \widetilde{A}'_1 = \widetilde{A}_1. \tag{1.10.14}$$

Only the longitudinal mode of the vector potential is not gauge invariant. For the transverse modes, (1.10.12) is precisely the Hamiltonian of a driven harmonic oscillator. However, for the longitudinal modes there are still constraints to be handled for gauge invariance. We can clean up (1.10.12) by introducing "*John's Quantum Map*" coordinates just as in (1.2.3) or analogously just like in our analysis of the quantum harmonic oscillator or as originally in Orig. problem 4.42. Define $\omega(\vec{k}) = |\vec{k}|$ and the mode vector $\vec{a}(t,\vec{k})$ for $h \in \{-1,1\}$:

$$\vec{a}(t,\vec{k}) = a_{-1}(t,\vec{k})\vec{\xi}_{-1}(\vec{k}) + a_1(t,\vec{k})\vec{\xi}_1(\vec{k}), \qquad a_h(t,\vec{k}) = \sqrt{\frac{\omega}{2\hbar}}\widetilde{A}_h(t,\vec{k}) + \frac{i}{\sqrt{2\omega\hbar}}\widetilde{P}_{Ah}(t,\vec{k}). \tag{1.10.15}$$

Inverting these and returning them to the Hamiltonian gives:

$$\widetilde{A}_h(t,\vec{k}) = \sqrt{\frac{\hbar}{2\omega}} \left( a_h(t,\vec{k}) + a_h^*(t,-\vec{k}) \right), \qquad \widetilde{P}_{Ah}(t,\vec{k}) = \frac{1}{i}\sqrt{\frac{\hbar\omega}{2}} \left( a_h(t,\vec{k}) - a_h^*(t,-\vec{k}) \right) \tag{1.10.16}$$

$$\mathcal{H} = \int_{\mathbb{K}} \left( \sum_h \left( \hbar\omega(\vec{k}) a_h^* a_h - \sqrt{\frac{2\hbar}{\omega}} \mathrm{Re}(\widetilde{J}_h^* a_h) \right) + \frac{|\widetilde{P}_{A0}|^2}{2} - \widetilde{J}_0^* \widetilde{A}_0 + \widetilde{\phi}^* (\widetilde{\rho} + i|\vec{k}|\widetilde{P}_{A0}) \right) d^3\vec{k} \tag{1.10.17}$$

$$\{a_h(t,\vec{k}), a_{h'}(t,\vec{k}')\} = 0, \quad \{a_h(t,\vec{k}), a_{h'}^*(t,\vec{k}')\} = -\frac{i}{\hbar}\delta_{hh'}\delta(\vec{k}-\vec{k}'), \quad \{a_h, \widetilde{A}_0\} = \{a_h, \widetilde{P}_{A0}\} = 0. \tag{1.10.18}$$

The Gauss' law constraint on the longitudinal modes requires:

$$\widetilde{P}_{A0} = \frac{\partial \widetilde{A}_0}{\partial t} + |\vec{k}|\widetilde{\phi} = \frac{i}{|\vec{k}|}\widetilde{\rho}. \tag{1.10.19}$$

Let $\Lambda(t,\vec{x})$ be a function whose Laplacian matches the divergence of the vector potential:

$$\nabla^2 \Lambda = \nabla \cdot \vec{A}. \tag{1.10.20}$$

Such a function is guaranteed to exist for vector potentials which decay at infinity due to the existence of solutions to the Poisson equation. Then, we have $\widetilde{A}_0 = |\vec{k}|\widetilde{\Lambda}$ and that $\Lambda$ is completely arbitrary under gauge transformation as $\Lambda' = \Lambda + f(t,\vec{x})$. Thus, $\Lambda$ is fully unphysical and undetermined by the dynamics (it is pure gauge), but can be used to eliminate $\widetilde{A}_0$ from the theory. Using these constraints, the electromagnetic Hamiltonian becomes:

$$\mathcal{H} = \int_{\mathbb{K}} \sum_h \left( \hbar\omega(\vec{k}) a_h^* a_h - \sqrt{\frac{2\hbar}{\omega}} \mathrm{Re}(\widetilde{J}_h^* a_h) \right) d^3\vec{k} + \int_{\mathbb{S},\mathbb{S}} \frac{\rho(t,\vec{x})\rho(t,\vec{x}')}{8\pi|\vec{x}-\vec{x}'|} d^3\vec{x} d^3\vec{x}' - \int_{\mathbb{S}} \vec{J} \cdot \nabla\Lambda d^3\vec{x}. \tag{1.10.21}$$

These last two terms do not at all depend on the dynamical degrees of freedom. They are necessary for managing Lorentz and gauge invariance together, but are not dynamically independent. In terms of $\Lambda$ and the $a_h$, the potentials (and $\vec{P}_A$) can be written:

$$\phi(t,\vec{x}) = -\frac{\partial \Lambda}{\partial t} + \int_{\mathbb{S}} \frac{\rho(t,\vec{x}')}{4\pi|\vec{x}-\vec{x}'|} d^3\vec{x}', \qquad \vec{A}(t,\vec{x}) = \nabla\Lambda + \mathrm{Re}\int_{\mathbb{K}} e^{i\vec{k}\cdot\vec{x}}\vec{a}(t,\vec{k})\sqrt{\frac{2\hbar}{\omega}}\frac{d^3\vec{k}}{(2\pi)^{\frac{3}{2}}}, \tag{1.10.22}$$

$$\vec{P}_A(t,\vec{x}) = \nabla\int \frac{\rho(t,\vec{x}')}{4\pi|\vec{x}-\vec{x}'|} d^3\vec{x}' + \mathrm{Im}\int_{\mathbb{K}} e^{i\vec{k}\cdot\vec{x}}\vec{a}(t,\vec{k})\sqrt{2\hbar\omega}\frac{d^3\vec{k}}{(2\pi)^{\frac{3}{2}}}. \tag{1.10.23}$$

Thus, all of classical electromagnetism can be expressed in terms of a fully arbitrary function $\Lambda(t,\vec{x})$ and the gauge invariant functions $a_{\pm1}(t,\vec{k})$ which represent the physical degrees of freedom of the system.

Now that we have reduced to something in plain harmonic oscillator form, we make a quantum theory of electromagnetism by doing to it exactly what you would do to any other classical harmonic oscillator to make it a quantum one. The classical physical degrees of freedom $a_{\pm1}(t,\vec{k})$ will be given by the operator summary of appropriate operators $\hat{a}_{\pm1}(\vec{k})$ on an appropriate Hilbert space $\mathbb{H}$ with respect to the system's state vector $|\Psi(t)\rangle$:

$$a_{\pm1}(t,\vec{k}) = \langle\Psi(t)|\hat{a}_{\pm1}(\vec{k})|\Psi(t)\rangle. \tag{1.10.24}$$

$\Lambda$ does not receive an operator because it is just an arbitrary function, not a degree of freedom of the theory. By theorem 1.5, if the classical quantities $a_{\pm1}(t,\vec{k})$ are to satisfy the Poisson brackets in (1.10.18), the operators $\hat{a}_{\pm1}$ must satisfy the commutators:

$$[\hat{a}_h(\vec{k}),\hat{a}_{h'}(\vec{k}')] = 0, \qquad [\hat{a}_h(\vec{k}),\hat{a}^\dagger_{h'}(\vec{k}')] = \delta_{hh'}\delta(\vec{k}-\vec{k}'). \tag{1.10.25}$$

If the classical quantities $a_{\pm1}(t,\vec{k})$ are to have the correct classical equations of motion (those given by the Hamiltonian (1.10.21)) then because those are simply the driven harmonic oscillator equations of motion and we know that the quantum harmonic oscillator gives the classical harmonic oscillator equations of motion for the operator summaries, we simply take the Hamiltonian operator of quantum electromagnetism to be:

$$\hat{H} = \int_{\mathbb{K}}\sum_h \left(\hbar\omega(\vec{k})\hat{a}^\dagger_h\hat{a}_h - \sqrt{\frac{2\hbar}{\omega}}\mathrm{Sa}(\tilde{J}^*_h\hat{a}_h)\right)d^3\vec{k} + \int_{\mathbb{S},\mathbb{S}}\frac{\rho(t,\vec{x})\rho(t,\vec{x}')}{8\pi|\vec{x}-\vec{x}'|}d^3\vec{x}d^3\vec{x}'\mathbb{1} - \int_{\mathbb{S}}\vec{J}\cdot\nabla\Lambda d^3\vec{x}\mathbb{1} \tag{1.10.26}$$

where Sa indicates the self-adjoint part of an operator (we also use As as the anti-self-adjoint part):

$$\mathrm{Sa}\hat{\mathcal{O}} = \frac{\hat{\mathcal{O}}+\hat{\mathcal{O}}^\dagger}{2}, \quad \mathrm{As}\hat{\mathcal{O}} = \frac{\hat{\mathcal{O}}-\hat{\mathcal{O}}^\dagger}{2i} \implies \hat{\mathcal{O}} = \mathrm{Sa}(\hat{\mathcal{O}}) + i\mathrm{As}(\hat{\mathcal{O}}), \quad (\mathrm{Sa}(\hat{\mathcal{O}}))^\dagger = \mathrm{Sa}(\hat{\mathcal{O}}), \quad (\mathrm{As}(\hat{\mathcal{O}}))^\dagger = \mathrm{As}(\hat{\mathcal{O}}). \tag{1.10.27}$$

It is useful for calculations to repackage the $\hat{a}_{\pm1}(\vec{k})$ commutators into commutators explicitly in terms of $\hat{\vec{a}}(\vec{k})$. With $\hat{a}_b(\vec{k})$ meaning component $b$ of $\hat{\vec{a}}(\vec{k})$, (1.10.25) becomes:

$$[\hat{a}_a(\vec{k}),\hat{a}_b(\vec{k}')] = 0, \qquad [\hat{a}_a(\vec{k}),\hat{a}^\dagger_b(\vec{k}')] = \perp_{ab}(\vec{k})\delta(\vec{k}-\vec{k}')\mathbb{1} \tag{1.10.28}$$

where $\perp_{ab}(\vec{k}) = \delta_{ab} - \frac{k_a k_b}{|\vec{k}|^2}$ is the projector orthogonal to $\vec{k}$. Because the classical quantities $\phi$, $\vec{A}$, $\vec{P}_A$ depend at most linearly on the $a_{\pm1}(t,\vec{k})$ and the $a_{\pm1}(t,\vec{k})$ are given by operator summaries, it is natural to define the operators:

$$\hat{\phi}(t,\vec{x}) = -\frac{\partial\Lambda}{\partial t}\mathbb{1} + \int_{\mathbb{S}}\frac{\rho(t,\vec{x}')}{4\pi|\vec{x}-\vec{x}'|}d^3\vec{x}'\mathbb{1}, \qquad \hat{\vec{A}}(t,\vec{x}) = \nabla\Lambda\mathbb{1} + \mathrm{Sa}\int_{\mathbb{K}} e^{i\vec{k}\cdot\vec{x}}\hat{\vec{a}}(\vec{k})\sqrt{\frac{2\hbar}{\omega}}\frac{d^3\vec{k}}{(2\pi)^{\frac{3}{2}}}, \tag{1.10.29}$$

$$\hat{\vec{P}}_A(t,\vec{x}) = \nabla\int\frac{\rho(t,\vec{x}')}{4\pi|\vec{x}-\vec{x}'|}d^3\vec{x}'\mathbb{1} + \mathrm{As}\int_{\mathbb{K}} e^{i\vec{k}\cdot\vec{x}}\hat{\vec{a}}(\vec{k})\sqrt{2\hbar\omega}\frac{d^3\vec{k}}{(2\pi)^{\frac{3}{2}}} \tag{1.10.30}$$

where $\hat{\vec{a}}$ is given by $\hat{\vec{a}}(\vec{k}) = \hat{a}_1(\vec{k})\vec{\xi}_1(\vec{k}) + \hat{a}_{-1}(\vec{k})\vec{\xi}_{-1}(\vec{k})$ (i.e. just (1.10.15) with everyone wearing a hat). Then, the classical potentials and conjugate momenta can be expressed as operator summaries of these operators in the natural way:

$$\phi(t,\vec{x}) = \langle\Psi(t)|\hat{\phi}(t,\vec{x})|\Psi(t)\rangle, \quad \vec{A}(t,\vec{x}) = \langle\Psi(t)|\hat{\vec{A}}(t,\vec{x})|\Psi(t)\rangle, \quad \vec{P}_A(t,\vec{x}) = \langle\Psi(t)|\hat{\vec{P}}_A(t,\vec{x})|\Psi(t)\rangle. \tag{1.10.31}$$

Similarly, it is natural to define the electric and magnetic field operators:

$$\hat{\vec{E}} = -\hat{\vec{P}}_A, \qquad \hat{\vec{B}} = \nabla\times\hat{\vec{A}} \tag{1.10.32}$$

and again due to the linearity of these expressions in the relevant operators, the classical $\vec{E}$ and $\vec{B}$ fields are:

$$\vec{E}(t,\vec{x}) = \langle\Psi(t)|\hat{\vec{E}}(t,\vec{x})|\Psi(t)\rangle, \qquad \vec{B}(t,\vec{x}) = \langle\Psi(t)|\hat{\vec{B}}(t,\vec{x})|\Psi(t)\rangle. \tag{1.10.33}$$

By using these definitions, together with the Schrödinger equation for $|\Psi\rangle$ under the Hamiltonian operator in (1.10.26) and the commutators in (1.10.25), it turns out that **Maxwell's equations are exactly maintained in quantum electromagnetism**. The difference between classical and quantum electromagnetism is that in classical electromagnetism the potentials $\phi, \vec{A}$ contain all of the data in the theory but in quantum electromagnetism it is possible to consider more complicated objects such as $\langle\hat{A}_x(t_1,\vec{x}_1)\hat{A}_z(t_2,\vec{x}_2)\rangle$ and other operator summaries of operator products which are quantities not found in the classical theory. Sometimes those more complicated objects will couple to physical quantities of interest, leading to observable differences between the theories. Quantum electromagnetism contains all of classical electromagnetism but it also includes many many other degrees of freedom, due to fundamentally being described by a wavefunctional $|\Psi(t)\rangle$ on an enormous Hilbert space.

The free Hamiltonian operator of quantum electromagnetism (meaning without any applied charges or currents) is (by simply setting $\rho = 0$, $\vec{J} = 0$ in (1.10.26)):

$$\hat{H} = \int_{\mathbb{K}} \sum_h \hbar\omega \hat{a}_h^\dagger \hat{a}_h d^3\vec{k}. \tag{1.10.34}$$

From (1.10.25), $\hat{a}_h(\vec{k})$ has the canonical commutation relations of the annihilation operator (with $\hat{a}_h^\dagger$ being the creation operator) for a harmonic oscillator. Let $|0\rangle$ be the normalized ground state of $\hat{H}$, which because $\hat{H}$ is a sum of harmonic oscillators corresponds to the state which is annihilated by by $\hat{a}_h(\vec{k})$ for all $h$ and $\vec{k}$:

$$\hat{a}_h(\vec{k})|0\rangle = 0. \tag{1.10.35}$$

This $|0\rangle$ state is often called the "vacuum" state. However, it is dangerous (read: incorrect) to think of the vacuum state as a state in which there is nothing going on. To see why it is useful to do two simple (though slightly tedious) calculations. For the first, let us compute the electric field in vacuum:

$$\langle 0|\hat{\vec{E}}(\vec{x})|0\rangle = \frac{1}{2i}\int_{\mathbb{K}} \left(e^{i\vec{k}\cdot\vec{x}}\langle 0|\hat{\vec{a}}|0\rangle - e^{-i\vec{k}\cdot\vec{x}}\langle 0|\hat{\vec{a}}^\dagger|0\rangle\right)\sqrt{2\hbar\omega}\frac{d^3\vec{k}}{(2\pi)^{\frac{3}{2}}} = 0. \tag{1.10.36}$$

Both $\langle 0|\hat{\vec{a}}|0\rangle$ and $\langle 0|\hat{\vec{a}}^\dagger|0\rangle$ are 0 because $\hat{\vec{a}}$ annihilates $|0\rangle$ in the first and $\langle 0|\hat{\vec{a}}^\dagger = (\hat{\vec{a}}|0\rangle)^\dagger = 0$ in the second. For the second, let us compute the summary of a product of electric field components in vacuum:

$$\langle 0|\hat{E}_a(\vec{x})\hat{E}_b(\vec{x}')|0\rangle = -\int_{\mathbb{K},\mathbb{K}'} \frac{\hbar\sqrt{\omega\omega'}}{16\pi^3}\langle 0|\left(e^{i\vec{k}\cdot\vec{x}}\hat{a}_a(\vec{k}) - e^{-i\vec{k}\cdot\vec{x}}\hat{a}_a^\dagger(\vec{k})\right)\left(e^{i\vec{k}'\cdot\vec{x}'}\hat{a}_b(\vec{k}') - e^{-i\vec{k}'\cdot\vec{x}'}\hat{a}_b^\dagger(\vec{k}')\right)|0\rangle d^3\vec{k}d^3\vec{k}'. \tag{1.10.37}$$

Because $\hat{\vec{a}}$ annihilates $|0\rangle$ and $\hat{\vec{a}}^\dagger$ annihilates $\langle 0|$, only one of the four terms in the product survives, giving:

$$\langle 0|\hat{E}_a(\vec{x})\hat{E}_b(\vec{x}')|0\rangle = \int_{\mathbb{K},\mathbb{K}'} \frac{\hbar\sqrt{\omega\omega'}}{16\pi^3}e^{i\vec{k}\cdot\vec{x}-\vec{k}'\cdot\vec{x}'}\langle 0|\hat{a}_a(\vec{k})\hat{a}_b^\dagger(\vec{k}')|0\rangle d^3\vec{k}d^3\vec{k}'. \tag{1.10.38}$$

The final surviving summary inside the integral can be evaluated using the commutation relations:

$$\langle 0|\hat{a}_a(\vec{k})\hat{a}_b^\dagger(\vec{k}')|0\rangle = \langle 0|(\hat{a}_a(\vec{k})\hat{a}_b^\dagger(\vec{k}') - \hat{a}_b^\dagger(\vec{k}')\hat{a}_a(\vec{k}) + \hat{a}_b^\dagger(\vec{k}')\hat{a}_a(\vec{k}))|0\rangle \tag{1.10.39}$$

$$= \langle 0|[\hat{a}_a(\vec{k}), \hat{a}_b(\vec{k}')]|0\rangle + \langle 0|\hat{a}_b^\dagger(\vec{k}')\hat{a}_a(\vec{k})|0\rangle = \perp_{ab}(\vec{k})\delta(\vec{k}-\vec{k}'). \tag{1.10.40}$$

Returning this to (1.10.38) gives:

$$\langle 0|\hat{E}_a(\vec{x})\hat{E}_b(\vec{x}')|0\rangle = \int_{\mathbb{K}} \hbar|\vec{k}| \perp_{ab}(\vec{k})e^{i\vec{k}\cdot(\vec{x}-\vec{x}')}\frac{d^3\vec{k}}{16\pi^3}. \tag{1.10.41}$$

The remaining integration can be completed to give the right-hand side an explicit distributional form, though it isn't very insightful. What is insightful is that this is a nonzero quite complicated object. Even in the vacuum state, there are nontrivial correlations in quantum field theory.

Because the eigenstates of a single harmonic oscillator can be written all by acting the creation operator some number of times on the ground state, the eigenstates of $\hat{H}$ here can be written by acting the creation operator (for possibly many pairs of helicity and wavenumber) some number of times on $|0\rangle$. That is, we can take as an eigenbasis of $\hat{H}_0$ the states constructed by the following. Let $D$ be a nonnegative integer and let $(n_1, h_1, \vec{k}_1), ..., (n_D, h_D, \vec{k}_D)$ be a list of $D$ triplets of a nonnegative integer $n$, a helicity $h$, and a wavenumber

$\vec{k}$ with no triplets sharing simultaneously their values of $h$ and $\vec{k}$. Define the state $|n_1 h_1 \vec{k}_1, ...n_D h_D \vec{k}_D\rangle$ by:

$$|n_1 h_1 \vec{k}_1, ...n_D h_D \vec{k}_D\rangle = \frac{\left(\hat{a}^\dagger_{h_1}(\vec{k}_1)\right)^{n_1}}{\sqrt{n_1!}} ... \frac{\left(\hat{a}^\dagger_{h_D}(\vec{k}_D)\right)^{n_D}}{\sqrt{n_D!}}|0\rangle. \tag{1.10.42}$$

Ranging over all possible lists $\alpha$ for a given $D$ and then all values of $D$ produces an eigenbasis of $\hat{H}$. The eigenvalues of $\hat{H}$ are found by directly applying the canonical commutation relations and that $|0\rangle$ is the ground state:

$$\hat{H}|n_1 h_1 \vec{k}_1, ...n_D h_D \vec{k}_D\rangle = \sum_{d=1}^{D} \hbar\omega(\vec{k}_d) n_d |n_1 h_1 \vec{k}_1, ...n_D h_D \vec{k}_D\rangle. \tag{1.10.43}$$

This eigenvalue spectrum precisely is exactly the distribution corresponding to the density of states necessary to derive the black-body formula, as we hoped. We can now understand how unwieldy the Hilbert space of quantum electromagnetism is by expressing a general state $|\Psi\rangle$ in this basis. Defining the components:

$$\Psi_{(n_1 h_1, ..., n_D h_D)}(\vec{k}_1, ..., \vec{k}_D) = \langle n_1 h_1 \vec{k}_1, ...n_D h_D \vec{k}_D|\Psi\rangle \tag{1.10.44}$$

and using that these form a complete basis, we have that an arbitrary $|\Psi\rangle$ can be written as:

$$|\Psi\rangle = \sum_{D=0}^{\infty} \sum_{n_1=0}^{\infty} \sum_{h_1 \in \pm 1} \int_{\mathbb{K}_1} ... \sum_{n_D=0}^{\infty} \sum_{h_D \in \pm 1} \int_{\mathbb{K}_D} \Psi_{(n_1 h_1, ..., n_D h_D)}(\vec{k}_1, .., \vec{k}_D)|n_1 h_1 \vec{k}_1, ...n_D h_D \vec{k}_D\rangle d^3\vec{k}_1 ... d^3\vec{k}_D. \tag{1.10.45}$$

Oof.

In the free theory (under evolution by $\hat{H}$), of course each eigenstate simply evolves by a phase determined by its eigenvalue. Consequently, the occupation numbers $n_1$, ..., $n_D$ are constant in the free theory. We can get a simpler intuition of what is going on with $|\Psi\rangle$ in the free theory by looking at the subspaces of the Hilbert space of different occupation numbers. The total occupation number is of a state of definite number simply given by eigenvalue of the total number operator $\hat{N}$, which as always for each species of a harmonic oscillator is given by $\hat{a}^\dagger \hat{a}$ and so in total is giving by summing over such species. Therefore:

$$\hat{N} = \int_{\mathbb{K}} \sum_{h \in \pm 1} \hat{a}^\dagger_h(\vec{k}) \hat{a}_h(\vec{k}) d^3\vec{k}. \tag{1.10.46}$$

The subspace with all occupation numbers as 0 is simply the vacuum $|0\rangle$. A generic state $|\Psi\rangle_1$ with total occupation number 1 can be written as a linear combination of occupation 1 eigenstates of $\hat{H}$. Under the free evolution given by $\hat{H}_0$, an occupation 1 state remains occupation 1 (because $\hat{N}$ commutes with $\hat{H}$) so takes the form:

$$|\Psi(t)\rangle_1 = \int_{\mathbb{K}} (\psi_1(t, \vec{k}) \hat{a}^\dagger_1(\vec{k}) + \psi_{-1}(t, \vec{k}) \hat{a}^\dagger_{-1}(\vec{k}))|0\rangle d^3\vec{k} = \int_{\mathbb{K}} \vec{\psi}(t, \vec{k}) \cdot \hat{\vec{a}}^\dagger(\vec{k})|0\rangle d^3\vec{k} \tag{1.10.47}$$

for some functions $\psi_{\pm 1}(t, \vec{k})$ or equivalently $\vec{\psi}(t, \vec{k}) = \psi_1(t, \vec{k})\vec{\xi}_1(\vec{k}) + \psi_{-1}(t, \vec{k})\vec{\xi}_{-1}(\vec{k})$. Such a $\vec{\psi}(t, \vec{k})$ can also be thought of as an arbitrary vector function satisfying:

$$\vec{k} \cdot \vec{\psi}(t, \vec{k}) = 0 \quad \text{or equivalently} \quad \perp^a{}_b(\vec{k}) \psi^b(t, \vec{k}) = \psi^a(t, \vec{k}). \tag{1.10.48}$$

The normalization of $|\Psi\rangle_1$ requires:

$$1 = \langle\Psi|\Psi\rangle_1 = \int_{\mathbb{K}\mathbb{K}'} \psi^{*a}(t, \vec{k}) \psi^b(t, \vec{k}') \langle 0|\hat{a}_a(\vec{k}) \hat{a}^\dagger_b(\vec{k}')|0\rangle d^3\vec{k} d^3\vec{k}' = \int_{\mathbb{K}} |\psi(t, \vec{k})|^2 d^3\vec{k}. \tag{1.10.49}$$

Therefore, $\vec{\psi}(t, \vec{k})$ must be normalized as a wavefunction on the single particle wavenumber space. We can find the explicit time evolution of $\vec{\psi}(t, \vec{k})$ using the Schrödinger equation for the full theory:

$$i\hbar \frac{d|\Psi(t)\rangle_1}{dt} = \hat{H}|\Psi\rangle_1 \implies i\hbar \int_{\mathbb{K}} \frac{\partial \vec{\psi}(t, \vec{k})}{\partial t} a^\dagger(\vec{k})|0\rangle d^3\vec{k} = \int_{\mathbb{K}} \hbar\omega(\vec{k}) \vec{\psi}(t, \vec{k}) \cdot \hat{\vec{a}}^\dagger|0\rangle d^3\vec{k} \tag{1.10.50}$$

$$\implies i\hbar \frac{\partial \vec{\psi}(t, \vec{k})}{\partial t} = \hbar\omega(\vec{k}) \vec{\psi}(t, \vec{k}). \tag{1.10.51}$$

This is precisely the Schrödinger equation for a single particle wavefunction written in wavenumber space with Hamiltonian operator $\hbar\omega(\hat{\vec{k}})$ (where $\hat{\vec{k}}$ would be the wavenumber operator for that single particle quantum system)! Define then the Hilbert space $\mathbb{H}_1$ of states of $\vec{\psi}(t, \vec{k})$ so that we now write $|\psi(t)\rangle$, with Hamiltonian

operator $\hat{H}_1 = \hbar\omega(\hat{\vec{k}})$ and $|\vec{k}_\pm\rangle$ the wavenumber eigenstates of helicity $\pm 1$. Then, we can write:

$$i\hbar\frac{d|\psi\rangle}{dt} = \hat{H}_1|\psi\rangle, \qquad \psi_{\pm 1}(t,\vec{k}) = \langle\vec{k}_\pm|\psi(t)\rangle. \tag{1.10.52}$$

As well, for any operator $\hat{\mathcal{O}}$, the operator summary in the state $|\Psi\rangle_1$ can be reduced to an operator summary on the Hilbert space of configurations of $\vec{\psi}(t,\vec{k})$:

$$\langle\Psi|\hat{\mathcal{O}}|\Psi\rangle_1 = \langle\psi|\hat{\mathcal{O}}_1|\psi\rangle \qquad \text{where:} \qquad \hat{\mathcal{O}}_1 = \int_{\mathbb{K}\mathbb{K}'}\sum_{h\in\pm1}\sum_{h'\in\pm1}|\vec{k}_h\rangle\langle 0|\hat{a}_h(\vec{k})\hat{\mathcal{O}}\hat{a}^\dagger_{h'}(\vec{k}')|0\rangle\langle\vec{k}'_{h'}|d^3\vec{k}d^3\vec{k}'. \tag{1.10.53}$$

Therefore, **the subspace of occupation number one reduces to just single particle quantum mechanics**. The particle (meaning wave-packet, not point-particle) created by a single application of the electromagnetic creation operator is called the **photon**. Because of this, we call the occupation number the particle number (and in the case of electromagnetism, the **photon number**) and say that $|\Psi\rangle_1$ is a single photon state. We could even define a wavefunction $\vec{\psi}(t,\vec{x})$ on position space by Fourier transforming $\vec{\psi}(t,\vec{k})$ in the usual way to give an even more direct look at a single photon wavefunction if we like. **Quantum field theory contains single particle quantum mechanics.** Plainly, using the results of Orig. problem 4.42, the total energy and momentum of $|\Psi_1\rangle$ satisfy $p_\mu = \hbar k_\mu$ as seen in 1.5.130, justifying that $\hbar$ as it appears in Orig. problem 4.48 is actually the quantum mechanical $\hbar$ for such states.

Now for an upgrade. What happens when we look at the subspace of photon number two? A state of photon number two will involve acting two creation operators on the vacuum state in a generic linear combination. Again the photon number is conserved by the free Hamiltonian, so under such evolution a generic state will be characterized by some coefficients $\psi_{\pm1\pm1}(t,\vec{k},\vec{k}')$ in the form:

$$|\Psi(t)\rangle_2 = \frac{1}{\sqrt{2}}\int_{\mathbb{K}\mathbb{K}'}\sum_{hh'}\psi_{hh'}(t,\vec{k},\vec{k}')\hat{a}^\dagger_h(\vec{k})\hat{a}^\dagger_{h'}(\vec{k}')|0\rangle d^3\vec{k}d^3\vec{k}'. \tag{1.10.54}$$

We can just as well define the bivector with components $\psi^{ab}(t,\vec{k},\vec{k}')$ by:

$$\psi^{ab}(t,\vec{k},\vec{k}') = \sum_{hh'}\psi_{hh'}(t,\vec{k},\vec{k}')\xi^a_h(\vec{k})\xi^b_{h'}(\vec{k}') \implies \psi^{ab}k_b = 0. \tag{1.10.55}$$

Because $\hat{a}^\dagger_a(\vec{k})$ and $\hat{a}^\dagger_b(\vec{k}')$ commute, $\psi^{ab}(t,\vec{k},\vec{k}')$ is automatically a symmetric under interchanging the order of photon creation (because it is defined simply as whatever coefficients are in (1.10.54)):

$$\psi^{ba}(t,\vec{k}',\vec{k}) = \psi^{ab}(t,\vec{k},\vec{k}'). \tag{1.10.56}$$

The normalization of $|\Psi\rangle_2$ requires:

$$1 = \langle\Psi|\Psi\rangle_2 = \frac{1}{2}\int_{\mathbb{K}\mathbb{K}'\mathbb{K}''\mathbb{K}'''}\psi^{*ab}(t,\vec{k},\vec{k}')\psi^{cd}(t,\vec{k}'',\vec{k}''')\langle 0|\hat{a}_a(\vec{k})\hat{a}_b(\vec{k}')\hat{a}^\dagger_c(\vec{k}'')\hat{a}^\dagger_d(\vec{k}''')|0\rangle d^3\vec{k}d^3\vec{k}'d^3\vec{k}''d^3\vec{k}'''. \tag{1.10.57}$$

We can simplify by sufficient use of the commutation relations:

$$\begin{aligned}\langle 0|\hat{a}_a(\vec{k})\hat{a}_b(\vec{k}')\hat{a}^\dagger_c(\vec{k}'')\hat{a}^\dagger_d(\vec{k}''')|0\rangle &= \langle 0|\hat{a}_a(\vec{k})(\hat{a}^\dagger_c(\vec{k}'')\hat{a}_b(\vec{k}')+\perp_{bc}(\vec{k}')\delta(\vec{k}'-\vec{k}''))\hat{a}^\dagger_d(\vec{k}''')|0\rangle\\ &= \langle 0|\hat{a}_a(\vec{k})\hat{a}^\dagger_c(\vec{k}'')\hat{a}_b(\vec{k}')\hat{a}^\dagger_d(\vec{k}''')|0\rangle+\perp_{bc}(\vec{k}')\delta(\vec{k}'-\vec{k}'')\langle 0|\hat{a}_a(\vec{k})\hat{a}^\dagger_d(\vec{k}''')|0\rangle\\ &= \perp_{ac}(\vec{k})\perp_{bd}(\vec{k})\delta(\vec{k}-\vec{k}'')\delta(\vec{k}'-\vec{k}''')+\perp_{ad}(\vec{k})\perp_{bc}(\vec{k}')\delta(\vec{k}-\vec{k}''')\delta(\vec{k}'-\vec{k}'').\end{aligned}$$

Now returning this to the normalization and using the symmetry of $\psi^{ab}$:

$$\int_{\mathbb{K}\mathbb{K}'}\psi^*_{ab}(t,\vec{k},\vec{k}')\psi^{ab}(t,\vec{k},\vec{k}')d^3\vec{k}d^3\vec{k}' = 1. \tag{1.10.58}$$

Therefore, $\psi$ is like a normalized wavefunction on the Hilbert space $\mathbb{H}_2$ which is formed by the symmetric elements of the tensor product of the Hilbert space $\mathbb{H}_1$ with another copy of $\mathbb{H}_1$. We express this by:

$$\mathbb{H}_2 = \text{Sym}(\mathbb{H}_1\otimes\mathbb{H}_1). \tag{1.10.59}$$

Very interestingly and very importantly, $\mathbb{H}_2$ is not simply two copies of $\mathbb{H}_1$. That is, a state of photon number two is not the same as a pair of states of photon number one. If we had two states $|\Psi\rangle_1$ and $|\Psi'\rangle_1$ of photon number one (with corresponding single particle wavefunctions $\vec{\psi}(t,\vec{k})$ and $\vec{\psi}'(t,\vec{k})$), then we could form a state of photon number two by considering the creation of a photon with wavefunction $\vec{\psi}(t,\vec{k})$ and

then the creation of another photon with wavefunction $\vec{\psi}'(t,\vec{k})$. That is, we could define:

$$|\Psi\rangle_{1,1} = \mathcal{N}\int_{\mathbb{K}\mathbb{K}'} \vec{\psi}(t,\vec{k})\cdot\hat{\vec{a}}^\dagger(\vec{k})\vec{\psi}'(t,\vec{k}')\cdot\hat{\vec{a}}^\dagger(\vec{k}')|0\rangle d^3\vec{k}d^3\vec{k}'. \tag{1.10.60}$$

where $\mathcal{N}$ is some normalization factor. This is a two photon state with two photon wavefunction:

$$\psi^{ab}_{1,1}(t,\vec{k},\vec{k}') = \frac{\mathcal{N}}{\sqrt{2}}(\psi^a(t,\vec{k})\psi'^b(t,\vec{k}') + \psi'^a(t,\vec{k})\psi^b(t,\vec{k}')). \tag{1.10.61}$$

This does not represent a generic state of $\mathbb{H}_2$. In fact, the linear combination of just two states of this form with linearly independent single photon wavefunctions produces a state which is not in this form (see the homework). Therefore, **the state space of a two photon system is significantly larger than two independent copies of the state space of a single photon system.** If the two photon state can be written in the form of (1.10.61) it is said to be **separable.** Otherwise, it is said to be **entangled**. It should be stressed that entanglement is the generic much more common case and separability is highly restrictive. It is entanglement that is behind the spin-1/2 puzzle in (1.10.5). The state space of two spin-1/2 particles is not just two copies of the single particle state space, and while, with $\hat{\vec{S}}_1$ the spin operators for particle one, on a single particle Hilbert space $|\langle\hat{\vec{S}}_1\rangle| = \frac{\hbar}{2}$ for all states, on a multiparticle Hilbert space $|\langle\hat{\vec{S}}_1\rangle|$ can take any value from 0 to $\frac{\hbar}{2}$, exactly consistent with experiment. How exactly this works is shown in the homework.

Just like in the single photon case, under the evolution given by $\hat{H}$ it is possible to define a reduced effective Hamiltonian operator for $\psi^{ab}(t,\vec{k},\vec{k}')$. Inserting (1.10.54) into the Schrödinger equation with $\hat{H}$ as the Hamiltonian operator we find:

$$i\hbar\frac{d|\Psi\rangle_2}{dt} = \hat{H}|\Psi\rangle_2 \implies i\hbar\frac{\partial\psi^{ab}(t,\vec{k},\vec{k}')}{\partial t} = (\hbar\omega(\vec{k}) + \hbar\omega(\vec{k}'))\psi^{ab}(t,\vec{k},\vec{k}'). \tag{1.10.62}$$

Therefore the effective two photon Hamiltonian operator $\hat{H}_2$ on $\mathbb{H}_2$ is simply $\hat{H}_2 = \hbar\omega(\hat{\vec{k}}) + \hbar\omega(\hat{\vec{k}}')$. It is important to notice that if the initial two photon wavefunction is separable then it remains separable under the time evolution given by $\hat{H}_2$. However, under a more general time evolution (suppose there was some sort of interaction present so that we did not reduce from $\hat{H}$ but instead from $\hat{H}+$ a term that causes the two photons to interact) even an initially separable wavefunction will become entangled dynamically. This is analogous to (and related to) how initially statistically independent quantities in statistical mechanics lose their statistical independence when they interact with each other. So, quantum field theory also *defines* two particle quantum mechanics as living on the Hilbert space of the tensor product of two Hilbert spaces of single particle quantum mechanics (the symmetrization is because in this example both particles are photons, the same species).

Also just like the single photon case, it is possible to reduce operator summaries involving $|\Psi\rangle_2$ to operator summaries involving only $\psi^{ab}(t,\vec{k},\vec{k}')$:

$$\langle\Psi|\hat{\mathcal{O}}|\Psi\rangle_2 \stackrel{!}{=} \langle\psi|\hat{\mathcal{O}}_2|\psi\rangle \tag{1.10.63}$$

$$\implies \hat{\mathcal{O}}_2 = \int_{\mathbb{K}^4}\sum_{hh'h''h'''} |\vec{k}_h\rangle\otimes|\vec{k}'_{h'}\rangle\langle 0|\hat{a}_h(\vec{k})\hat{a}_{h'}(\vec{k}')\hat{\mathcal{O}}\hat{a}^\dagger_{h''}(\vec{k}'')\hat{a}^\dagger_{h'''}(\vec{k}''')|0\rangle\langle\vec{k}''_{h''}|\otimes\langle\vec{k}'''_{h'''}|d^3\vec{k}d^3\vec{k}'d^3\vec{k}''d^3\vec{k}'''. \tag{1.10.64}$$

It should be no surprise that at each photon number $N$ it is possible to define an effective $N$ photon wavefunction of $N$ different wavenumbers and helicities which is fully symmetric, evolves under a Schrödinger type equation in the free theory, and for which it is possible to define reduced operators for operator summaries in an $N$ photon state. Because $\mathbb{H}_N$ is formed from a (symmetrized) tensor product of $\mathbb{H}_1$'s rather than simply a direct sum (which would be a list of $N$ copies of $\mathbb{H}_1$), there are many independent quantities in an $N$ photon state which have no single photon description. $\mathbb{H}_N$ is properly written as:

$$\mathbb{H}_N = \text{Sym}(\mathbb{H}_1\otimes ...(N \text{ total factors})... \otimes\mathbb{H}_1). \tag{1.10.65}$$

The full Hilbert space $\mathbb{H}$ of the quantum field theory of the electromagnetic field consists of the direct sum of the Hilbert spaces of each particle number:

$$\mathbb{H} = \mathbb{H}_1\oplus\mathbb{H}_2\oplus\mathbb{H}_3\oplus ... \tag{1.10.66}$$

as a general state in $\mathbb{H}$ can be formed by linear combinations of states of various $D,n,h,\vec{k}$ as defined in (1.10.42). Such a Hilbert space constructed out of the direct sum of individual Hilbert spaces of various particle numbers is called a Fock space.

An $N$ photon state remains an $N$ photon state under the free Hamiltonian operator where $\rho = 0$, $\vec{J} = 0$. What happens if $\vec{J} \neq 0$ as in (1.10.26)? The number operator $\hat{N}$ no longer commutes with $\hat{H}$ so an $N$ photon state will not remain an $N$ photon state. The commutator of $\hat{N}$ with $\hat{H}$ gives:

$$\frac{1}{i\hbar}[\hat{N}, \hat{H}] = \int_{\mathbb{K}} \sum_h \sqrt{\frac{2}{\hbar\omega}} \text{As}(\widetilde{J}_h^* \hat{a}_h) d^3\vec{k}. \tag{1.10.67}$$

The easiest way to see that the presence of a current changes the number of photons is to simply look at what $\hat{H}$ does to the vacuum state. Suppose for simplicity we consider a nonzero current but a charge density of 0, and that we choose $\Lambda = 0$ by definition. Then, suppose we begin initially in the vacuum state $|0\rangle$ at $t = 0$. The resulting initial time derivative of $|\Psi\rangle$ at $t = 0$ is given by:

$$i\hbar \left.\frac{d|\Psi\rangle}{dt}\right|_{t=0} = \hat{H}|0\rangle = \int_{\mathbb{K}} \sum_h \sqrt{\frac{\hbar}{2\omega}} \widetilde{J}_h(\vec{k}) \hat{a}_h^\dagger(\vec{k})|0\rangle d^3\vec{k} \tag{1.10.68}$$

which is a linear combination of one photon states. Therefore, an arbitrarily short application of the current will cause a transition from the vacuum state to a linear combination of the vacuum state and one photon states (and as time passes the full $|\Psi\rangle$ generically becomes a linear combination of photon states of arbitrarily high and different numbers). **Interactions cause particle production**. As well, if we began in a state with a nonzero number of photons, the current can drive transitions to state of a lower number of photons (because $\widetilde{J}_h^*(\vec{k})$ multiplies the annihilation operator $\hat{a}_h(\vec{k})$ in the Hamiltonian operator).

The properties of quantum field theory we have seen in this section are not unique to electromagnetism. By theorem 1.2, the lowest order Hamiltonian of any stable classical field theory can be put into the form of creation and annihilation fields (as $\psi$ in the Schrödinger equation satisfies the canonical Poisson brackets with $\psi^*$ appropriate for creation and annihilation fields). So, when "quantized", meaning that classical field theory Hamiltonian is replaced by a quantum field theory with an appropriate Hamiltonian operator, its free Hamiltonian operator (the lowest order quadratic part) will be that of an infinite number of quantum harmonic oscillators and it will produce the subspaces of different particle numbers seen here for electromagnetism. The inclusion of interaction terms (anything beyond that minimal harmonic oscillator Hamiltonian operator) will generically cause transitions between the subspaces of different particle numbers. Therefore, if we describe matter by quantum field theory then we can explain the experimental fact that particles can be created and destroyed through interactions. Similarly, we can explain the existence of indistinguishable particles (as they would simply be particles formed from the same set of creation and annihilation operators). We can even explain the spin puzzle from (1.10.5). If we take electrons to be particles created by an appropriate electron creation operator, then a single electron system will have a two component (due to being spin-1/2) wavefunction $\psi^a(t, \vec{x})$ but a two electron system would have a four component (due to the tensor product of two spin-1/2 spaces) wavefunction $\psi^{ab}(t, \vec{x}, \vec{x}')$. The observation that no single electron wavefunction can explain the single electron spin statistics observed in a two electron system is completely explained by the fact that $\psi^{ab}(t, \vec{x}, \vec{x}')$ will generically be entangled and so there will be no single electron wavefunction appropriate for either individual electron.

The structure of quantum field theory explains the fact noted just after (1.2.22) that $\hbar$, although it that context seemingly derived from the specific wavefunction of interest, is actually universal. In QFT, $\hbar$ appears in (1.10.9) and $\langle\Psi|\Psi\rangle = 1$. All of the $N$ particle subspaces of QFT are derived from this single Lagrangian in the manner of this section, and each of those subspaces inherits the same value of $\hbar$. Even if you have multiple different species of particles, because all of their subspaces are derived from the same wavefunctional of everything, they all inherit that same $\hbar$. As explained around (1.5.130), this is what makes it possible to universally use units in which $\hbar = 1$, eliminating it from the theory. We will start to do that consistently in the next chapter.

QFT also tells us how to put two quantum systems together. Suppose we have two different fields for which we desire a quantum field theory (you could even take the plus helicity vector potential and minus helicity vector potential as examples, or the vector potential and a matter field). Then, following the same steps as done so far here, we would seek a Hamiltonian operator in terms of those field operators which would act on the Hilbert space of wavefunctionals of field configurations of those two fields. The Hilbert space of a function of two variables is the tensor product of the Hilbert spaces of functions of each variable, and so when we have multiple field the full theory has Hilbert space given by the tensor product of Hilbert spaces for each field. If we then look at the subspace of this Hilbert space in which there are $n_A$ particles of species $A$ and $n_B$ particles of species $B$, again that Hilbert space is automatically the tensor product of the Hilbert spaces of the individual two species types. So, in general in quantum mechanics, when you have one quantum

system described by a wavefunction in Hilbert space $\mathbb{H}_A$ and you bring it together with another quantum system described by a wavefunction in Hilbert space $\mathbb{H}_B$, then the combined system will actually have a wavefunction in the Hilbert space $\mathbb{H}_A \otimes \mathbb{H}_B$.

Quantum field theory fully contains the single particle quantum mechanics of this chapter, by simply restricting to the one particle Hilbert space for matter fields, but it also defines and explains what to do when we have more than a single matter particle. The Schrödinger equation for a single electron just describes the one electron subspace of a quantum field theory in which electrons are created and annihilated by appropriate operators. Quantum field theory is not just a formalism for electromagnetism, it is a formalism for everything. Ultimately, all species of matter of all interactions are described by interacting quantum field theory. It is, to this day, our basic understanding of how the universe actually works.

Perhaps surprisingly at first, many particle quantum mechanics of a fixed number and species of particles is inherently a nonrelativistic theory. This is because approximately the same energy scale determines whether a particle is relativistic as well as whether particles of that species can be created in an interaction. For a localized relativistic wave-packet, we have approximately $\mathrm{E} \approx \sqrt{|\vec{p}|^2 + m^2}$ in Lorentz natural units. If $|\vec{p}| \ll m$ then the nonrelativistic approximation is appropriate. However, if $|\vec{p}|$ is of the some order as $m$ or larger, a special relativistic treatment is necessary. Consequently, at energies of order $\mathrm{E} = 2m$, relativity is important. But, an energy of $2m$ is enough energy to transition from a state with a single particle of mass $m$ to two particles of mass $m$. Therefore, particle creation has the same characteristic energy scale as relativistic effects. Generically, the Hamiltonian of an interacting system will allow particle creation and annihilation and so it is not possible to maintain a fixed number of particles of fixed species when the particles have relativistic energies.

## Exercises

**Problem 1.81: Compton Scattering.** Consider Compton scattering, which is the scattering of a single photon off of a nonrelativistic electron without any particles being created or destroyed. Assume that in both the initial and final states the electron and photon are each sufficiently localized wave-packets for the purposes of relevant approximations.

(a) Explain that when they are far from each other the electron and photon energy can be written as:

$$\mathrm{E}_e = \sqrt{c^2|\vec{p}_e|^2 + m_e^2c^4}, \qquad \mathrm{E}_\gamma = \hbar c|\vec{k}_\gamma|$$

where $\vec{p}_e$ is the momentum of the electron and $\vec{k}_\gamma$ is the wavenumber of the photon.

(b) Let us work in the frame in which the electron is initially at rest. Let $\lambda$ be the initial wavelength of the photon and $\lambda'$ be its final wavelength. Let $\theta$, the scattering angle, be the angle between the photon's initial and final momenta. Show, using the conservation of energy and linear momentum, that the color of light scattered off of an electron is shifted according to:

$$\lambda' - \lambda = \frac{2\pi\hbar}{m_e c}(1 - \cos\theta).$$

**Problem 1.82: Statistics of a two particle system.** Consider two distinguishable spin-1/2 particles that are each approximately localized well enough that we can treat each of them as at some position with some spinor describing their polarization. Let $\mathbb{H}_1$ ($\mathbb{H}_2$) be the Hilbert space of the spinor of particle 1 (particle 2) with basis vectors $|+_1\rangle$ and $|-_1\rangle$ ($|+_2\rangle$ and $|-_2\rangle$) indicating the spin eigenvectors aligned and anti-aligned with $z$-axis.

(a) Explain that the interacting system's state vector is of the form (where $|\pm_1 \pm_2\rangle = |\pm_1\rangle \otimes |\pm_2\rangle$):

$$|\psi\rangle = \psi_{++}|+_1 +_2\rangle + \psi_{+-}|+_1 -_2\rangle + \psi_{-+}|-_1 +_2\rangle + \psi_{--}|-_1 -_2\rangle.$$

(b) Show that $|\psi\rangle$ is separable (meaning each particle has a well-defined wavefunction) if and only if:

$$\psi_{++}\psi_{--} = \psi_{+-}\psi_{-+}.$$

Show that in this case and only in this case, with $|\psi_1\rangle$ ($|\psi_2\rangle$) the state vector of particle 1 (particle 2):

$$\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle.$$

(c) Consider the state $|\psi\rangle$ below. Show that in this state the total spin of each particle is 0:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+_1 +_2\rangle + |-_1 -_2\rangle) \implies \vec{\mathrm{S}}_1 = \langle \hat{\vec{S}}_1 \otimes \mathbb{1}_2\rangle = \vec{\mathrm{S}}_2 = \langle \hat{\mathbb{1}}_1 \otimes \hat{\vec{S}}_2\rangle = 0.$$

Thus this state is consistent with the results of the experiment described in the section.

(d) Consider a generic state vector for which $\vec{\mathrm{S}}_1 = \vec{\mathrm{S}}_2 = 0$. Show that (after absorbing the meaningless overall phase of the wavefunction) for real parameters $\theta$, $\alpha$, $\beta$ the wavefunction can be written as:

$$|\psi\rangle = \frac{\cos\theta}{\sqrt{2}}\left(e^{i\alpha}|+_1+_2\rangle + e^{-i\alpha}|-_1-_2\rangle\right) + \frac{\sin\theta}{\sqrt{2}}\left(e^{i\beta}|+_1-_2\rangle - e^{-i\beta}|-_1+_2\rangle\right).$$

Show that $\theta, \alpha, \beta$ can be determined from the values of the correlation matrix $\langle \hat{S}_1^a \otimes \hat{S}_2^b \rangle$. Show further that if this matrix is required to be rotationally invariant then up to an overall phase:

$$|\psi\rangle = \frac{|+_1-_2\rangle - |-_1+_2\rangle}{\sqrt{2}}.$$

Thus this is the only spherically symmetric state vector for the spins of two spin-1/2 particles.

**Problem 1.83: Addition of Angular Momentum for Spin-1/2 Particles.** Consider two interacting well-localized spin-1/2 particles. Let $\mathbb{H}_1$ ($\mathbb{H}_2$) be the Hilbert space of the spinor of particle 1 (particle 2).

(a) Let $\vec{\mathrm{S}}$ be the total (free) spin angular momentum of the two particles. Show that $\vec{\mathrm{S}} = \langle \hat{\vec{S}} \rangle$ where the operator $\hat{\vec{S}}$ is given in terms of the individual particle spin operators $\hat{\vec{S}}_1$, $\hat{\vec{S}}_2$ as $\hat{\vec{S}} = \hat{\vec{S}}_1 \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \hat{\vec{S}}_2$.

(b) Show that the components of $\hat{\vec{S}}$ satisfy the usual su(2) commutation relations (1.7.13). Explain based on theorem 1.11 that this means the Hilbert space $\mathbb{H} = \mathbb{H}_1 \otimes \mathbb{H}_2$ can be decomposed into a direct sum of Hilbert spaces so that on each component of the direct sum $\hat{\vec{S}}$ acts in an irreducible representation of su(2).

(c) Compute how $\hat{S}_a$ acts on each of the basis states $|+_1+_2\rangle$, $|+_1-_2\rangle$, $|-_1+_2\rangle$, $|-_1-_2\rangle$.

(d) From the results of the previous part, diagonalize $\hat{S}_z$ and $\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$. Show that $\mathbb{H}$ decomposes specifically as $\mathbb{H} = \mathbb{H}_0 \oplus \mathbb{H}_1$ where for states in $\mathbb{H}_0$ $\hat{\vec{S}}$ acts in an irreducible spin-0 representation and for states in $\mathbb{H}_1$ $\hat{\vec{S}}$ acts in an irreducible spin-1 representation. Let $|s,m\rangle$ be the eigenstates of $\hat{S}^2, \hat{S}_z$ with spin-$s$ and $z$-component eigenvalue $\hbar m$. Show that the eigenstates of total spin are related to the individual particle bases by:

$$|0,0\rangle = \frac{|+_1-_2\rangle - |-_1+_2\rangle}{\sqrt{2}}, \quad |1,1\rangle = |+_1+_2\rangle, \quad |1,0\rangle = \frac{|+_1-_2\rangle + |-_1+_2\rangle}{\sqrt{2}}, \quad |1,-1\rangle = |-_1-_2\rangle.$$

**Problem 1.84: Interaction of Spin-1/2 Particles.** Consider, for some constant $\omega$, the interaction of two localized spin-1/2 particles according to the Hamiltonian operator:

$$\hat{H} = \frac{\omega}{\hbar}\hat{\vec{S}}_1 \cdot \hat{\vec{S}}_2.$$

Let $\hat{\vec{S}}$ be the generating operators of the total spin of the two particles and $|s,m\rangle$ be the basis constructed in the previous question. Show that:

$$\hat{H} = \frac{\omega}{2\hbar}(\hat{S}^2 - \hat{S}_1^2 - \hat{S}_2^2), \qquad \hat{H}|s,m\rangle = \frac{\hbar\omega}{4}(2s^2+2s-3)|s,m\rangle.$$

**Problem 1.85: ⋆ Addition of Angular Momentum in general.** Consider two quantum systems, the first in an irreducible representation of angular momentum $j_1$ and the second in an irreducible representation of angular momentum $j_2$. Label a basis of states of system 1 by $|m_1\rangle$, where $\hbar m_1$ is the eigenvalue of $\hat{J}_{1z}$ and of system 2 by $|m_2\rangle$, where $\hbar m_2$ is the eigenvalue of $\hat{J}_{2z}$.

(a) Let $\vec{\mathrm{J}}$ be the total angular momentum of the two systems. Show that $\vec{\mathrm{J}} = \langle \hat{\vec{J}} \rangle$ where $\hat{\vec{J}} = \hat{\vec{J}}_1 \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \hat{\vec{J}}_2$ and $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ are the angular momentum operators of the two systems.

(b) Show that the components of $\hat{\vec{J}}$ satisfy the usual su(2) commutation relations (1.7.13). Explain based on theorem 1.11 that this means the Hilbert space $\mathbb{H} = \mathbb{H}_1 \otimes \mathbb{H}_2$ can be decomposed into a direct sum of Hilbert spaces so that on each component of thes sum $\hat{\vec{J}}$ acts in an irreducible representation of su(2).

(c) Let $|j,m\rangle$ be a state in total angular momentum representation $j$ with $\hat{J}_z$ eigenvalue $\hbar m$. Explain that there exist some so-called Clebsch-Gordan coefficients $C^{jm}_{m_1 m_2}$ so that:

$$|j,m\rangle = \sum_{m_1 m_2} C^{jm}_{m_1 m_2} |m_1\rangle \otimes |m_2\rangle$$

which do also depend on $j_1$ and $j_2$, but because those are fixed in this question we suppress the visibility of that dependence to keep notation condensed.

(d) Define the operators $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$ and compute $\hat{J}_\pm|j,m\rangle$ following the calculations of section 1.7. Show that the Clebsch-Gordan coefficients satisfy the recurrence relation:

$$\sqrt{j(j+1)-m(m+1)}C^{j(m\pm1)}_{m_1m_2} = \sqrt{j_1(j_1+1)-m_1(m_1\mp1)}C^{jm}_{(m_1\mp1)m_2} + \sqrt{j_2(j_2+1-m_2(m_2\mp1)}C^{jm}_{m_1(m_2\mp1)}.$$

(e) By using the definition of $\hat{S}_z$, show that $C^{jm}_{m_1m_2} = 0$ unless $m = m_1 + m_2$. Consequently, show that the maximum value of $j$ is $j_1 + j_2$. Then, show that the dimension of the direct sum of Hilbert spaces of irreducible representations of angular momenta $|j_1 - j_2|$, $|j_1 - j_2| + 1$, ..., $j_1 + j_2$ equals the dimension of the tensor product Hilbert space of $\mathbb{H}_1 \otimes \mathbb{H}_2$.

(f) For simplicity, assume we have labeled the systems so that $j_1 \geq j_2$. Explain that using $\hat{J}_-$, if we know all of the coefficients for $C^{jj}_{m_1(j-m_1)}$ we can determine all of $C^{jm}_{m_1m_2}$. Then show that:

$$C^{jj}_{(m_1-1)(j-m_1+1)} = -\sqrt{\frac{j_2(j_2+1)-(j-m_1)(j-m_1+1)}{j_1(j_1+1)-m_1(m_1-1)}}C^{jj}_{m_1(j-m_1)}.$$

Solve this recurrence relation and normalize $|j,j\rangle$ with the convention that $(\langle j_1| \otimes \langle j-j_1|)|jj\rangle$ be real and greater than 0. Use this to show that the Clebsch-Gordan coefficients for $m \geq 0$ are given by:

$$\begin{aligned} C^{jm}_{m_1m_2} =& \delta^m_{m_1+m_2}\sqrt{\frac{(2j+1)(j+j_1-j_2)!(j-j_1+j_2)!(j_1+j_2-j)!}{(j_1+j_2+j+1)!}} \\ &\times \sqrt{(j+m)!(j-m)!(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!} \\ &\times \sum_k \frac{(-1)^k}{k!(j_1+j_2-j-k)!(j_1-m_1-k)!(j_2+m_2-k)!(j-j_2+m_1+k)!(j-j_1-m_2+k)!} \end{aligned}$$

where the $k$ sum is over all integers for which the arguments of all the factorials are nonnegative.

(g) Show that the coefficients for $m < 0$ can be determined by:

$$C^{jm}_{m_1m_2} = (-1)^{j-j_1-j_2}C^{j(-m)}_{(-m_1)(-m_2)}.$$

In a general context, we will write $C^{jm}_{j_1m_1j_2m_2}$ for the Clebsch-Gordan coefficients. As well, for symmetry convenience, it is typical to define the coefficients for $j_1 < j_2$ by:

$$C^{jm}_{j_1m_1j_2m_2} = (-1)^{j-j_1-j_2}C^{jm}_{j_2m_2j_1m_1}.$$

(h) Let $\mathbb{H}_j$ indicate the Hilbert space of states in the irreducible representation of angular momentum $j$. Show that all of the results of this question together imply:

$$\mathbb{H}_{j_1} \otimes \mathbb{H}_{j_2} = \mathbb{H}_{|j_1-j_2|} \oplus \mathbb{H}_{|j_1-j_2|+1} \oplus ... \oplus \mathbb{H}_{j_1+j_2-1} \oplus \mathbb{H}_{j_1+j_2}.$$

The collective results of this question are known as the Clebsch-Gordan theorem.

**Problem 1.86: Wigner-Eckart Theorem.** Let $\hat{R}(\vec{\theta})$ be a generic $3 \times 3$ rotation matrix, with components $R^a{}_b$. As we have previously defined, a rank $N$ rotational tensor $T$ is an object with components $T^{a_1\ldots a_N}$ which under a rotation transform according to $T'^{a_1\ldots a_N} = R^{a_1}{}_{b_1}...R^{a_N}{}_{b_N}T^{b_1\ldots b_N}$.

(a) Let $\mathbb{V}$ be the vector space of 3 dimensional vectors. Explain why the vector space of rank $N$ rotational tensors is simply the $N$ fold tensor product of $\mathbb{V}$ with itself, i.e. $\mathbb{V} \otimes \mathbb{V} \otimes ... \otimes \mathbb{V}$ with $N$ factors of $\mathbb{V}$.

(b) Explain that $\mathbb{V}$ is the Hilbert space of the irreducible $j = 1$ representation of su(2).

(c) Consider the irreducible angular momentum $j$ representation of su(2). Define Wigner's $D$-matrix:

$$D^j_{m'm}(\vec{\theta}) = \langle j,m'|e^{-\frac{i}{\hbar}\vec{\theta}.\hat{\vec{J}}}|j,m\rangle.$$

Explain how $D^j_{m'm}$ completely characterizes the $j$ angular momentum representation of rotations. Explicit formulas exist for $D^j_{m'm}$, but they are extremely too complicated to usually be useful. Let $|v\rangle$ be a state vector in the Hilbert space of irreducible angular momentum $j$ representation of su(2) with components $v_m$ as below. Show that under a rotation the components of $|v\rangle$ transform as:

$$|v\rangle = \sum_{m=-j}^{j} v_m|j,m\rangle \implies v'_m = \sum_{m'=-j}^{j} D^j_{mm'}v_{m'}.$$

(d) Using the Clebsch-Gordan theorem, explain why the vector space of rank $N$ rotational tensors is not an

irreducible representation of angular momentum and in fact actually contains a tower of representations of different angular momenta. Show that there exists a single subspace of the vector space of rank $N$ rotational tensors which is an irreducible representation of su(2) with angular momentum $N$. A rank $N$ tensor which is in this subspace is called a rank $N$ spherical tensor.

(e) Let $T^{(N)}$ be a rank $N$ spherical tensor. Explain that there exist $2N+1$ rank $N$ spherical tensors $M_k^{(N)}$ with the label $k$ running from $-N$ to $N$ which form a basis for such tensors and components $T_k^{(N)}$ so that $T^{(N)}$ can be written as below and so that under a rotation the $T_k^{(N)}$ transform like:

$$T^{(N)} = \sum_{k=-N}^{N} T_k^{(N)} M_k^{(N)} \implies T_k'^{(N)} = \sum_{m=-N}^{N} D_{km}^N T_m^{(N)}.$$

(f) In quantum mechanics, a spherical tensor operator of rank $N$ is an operator $\hat{T}^{(N)}$ so that $\langle \hat{T}^{(N)} \rangle$ is a spherical tensor of rank $N$. Show that this requirement implies:

$$[\hat{T}_k^{(N)}, \hat{J}_a] = \sum_{m=-j}^{j} \langle N, k | \hat{J}_a | N, m \rangle \hat{T}_m^{(N)}.$$

Equivalently, show that this requires:

$$[\hat{T}_k^{(N)}, \hat{J}_\pm] = \hbar \sqrt{N(N+1) - k(k \mp 1)} \hat{T}_{k\mp 1}^{(N)}, \qquad [\hat{T}_k^{(N)}, \hat{J}_z] = \hbar k \hat{T}_k^{(N)}.$$

(g) From the above commutators, show that the matrix elements of components of a spherical tensor operator satisfy $\langle j, m | \hat{T}_k^{(N)} | j', m' \rangle = 0$ unless $m' = m + k$ and also satisfy the recurrence relation:

$$\begin{aligned}\sqrt{j'(j'+1) - m'(m' \pm 1)} \langle j, m | \hat{T}_k^{(N)} | j', m' \pm 1 \rangle =& \sqrt{j(j+1) - m(m \mp 1)} \langle j, m \mp 1 | \hat{T}_k^{(N)} | j', m' \rangle \\ &+ \sqrt{N(N+1) - k(k \mp 1)} \langle j, m | \hat{T}_{k\mp 1}^{(N)} | j', m' \rangle.\end{aligned}$$

(h) Show that the matrix elements of a spherical tensor operator satisfy all of the same conditions as the Clebsch-Gordan coefficients except normalization, and thus based on the Clebsch-Gordan theorem show that the matrix elements of the components of a spherical tensor operator satisfy:

$$\langle j, m | \hat{T}_k^{(N)} | j', m' \rangle = T_{jj'}^{(N)} C_{jmNk}^{j'm'}$$

where $C_{j_1 m_1 j_2 m_2}^{jm}$ are the Clebsch-Gordan coefficients and $T_{jj'}^{(N)}$ are called the Wigner-Eckart reduced matrix elements of $\hat{T}^{(N)}$. This is known as the Wigner-Eckart theorem.

**Problem 1.87: Quantum Maxwell's Equations.** Show by explicit calculation, using (1.10.9), (1.10.25), (1.10.26), (1.10.29), (1.10.30), and (1.10.33) that $\vec{E}$ and $\vec{B}$ exactly satisfy Maxwell's equations even in the quantum theory. Explain why this is unsurprising given the results of problem 1.34.

**Problem 1.88: Electromagnetic Wavefunctionals.** Consider free quantum electromagnetism, defined by the Hamiltonian operator in (1.10.26) with $\rho = 0$ and $\vec{J} = 0$.

(a) Let $|\vec{A}\rangle$ be a functional Dirac delta in the rigged Hilbert space of functionals of all configurations of the vector potential, i.e. an eigenstate of $\hat{\vec{A}}(\vec{x})$ with eigenvalue some specific field configuration at a point $\vec{A}(\vec{x})$. We define the wavefunctional $\Psi[\vec{A}] = \langle \vec{A} | \Psi \rangle$. Show:

$$\langle \vec{A} | \hat{\vec{A}}(\vec{x}) | \Psi \rangle = \vec{A}(\vec{x}) \Psi[\vec{A}], \qquad \langle \vec{A} | \hat{\vec{P}}_A(\vec{x}) | \Psi \rangle = -i\hbar \frac{\delta \Psi[\vec{A}]}{\delta \vec{A}(\vec{x})}.$$

(b) Let $\Psi_0[\vec{A}]$ be the free vacuum state of $\hat{H}$. Show that (1.10.35) implies:

$$\Psi_0[\vec{A}] = \Psi_0[0] e^{-\int \frac{|\widetilde{\vec{B}}|^2}{2\hbar |\vec{k}|} d^3\vec{k}}$$

where $\widetilde{\vec{B}}$ is the Fourier transform of the magnetic field $\vec{B}$ and $\Psi_0[0]$ is an irrelevant normalization that would somehow be fixed by "integrating over all possible field configurations $\vec{A}$" and normalizing. Such functional integrals are difficult to define rigorously, but conveniently $\Psi_0[0]$ can always canceled from observables, so there is no need to worry about it. Plainly, the electromagnetic vacuum is a gauge invariant Gaussian in the space of field configurations. A true state of "nothing going on" by contrast would have a Dirac delta at a field value of 0.

(c) Let $\Psi_1[\vec{A}]$ be a generic 1 photon state. Show that for some $\vec{\psi}(\vec{k})$ satisfying (1.10.48) and (1.10.49):

$$\Psi_1[\vec{A}] = \sqrt{\frac{2}{\hbar}} \int \sqrt{|\vec{k}|}\, \widetilde{\vec{A}}^*(\vec{k}) \cdot \vec{\psi}(\vec{k}) d^3\vec{k}\, \Psi_0[\vec{A}].$$

Describe how this functional looks in field space. Show that $\langle \hat{\vec{E}}(\vec{x}) \rangle$ and $\langle \vec{B}(\vec{x}) \rangle$ are both 0 for any 1 photon state.

(d) Let $\Psi_2[\vec{A}]$ be a generic 2 photon state. Show that for some $\psi^{ab}(\vec{k}, \vec{k}')$ satisfying (1.10.55), (1.10.56), and (1.10.58):

$$\Psi_2[\vec{A}] = \frac{1}{\sqrt{2}} \int \psi^{ab}(\vec{k}, \vec{k}') \left( \frac{2}{\hbar} \sqrt{|\vec{k}||\vec{k}'|} \widetilde{A}^*_a(\vec{k}) \widetilde{A}^*_b(\vec{k}') - \delta_{ab}\delta(\vec{k} + \vec{k}') \right) d^3\vec{k} d^3\vec{k}' \Psi_0[\vec{A}].$$

Describe how this functional looks in field space. Show that $\langle \hat{\vec{E}}(\vec{x}) \rangle$ and $\langle \vec{B}(\vec{x}) \rangle$ are both 0 for any 2 photon state.

(e) Show that $\langle \hat{\vec{E}}(\vec{x}) \rangle$ and $\langle \vec{B}(\vec{x}) \rangle$ are both 0 for any state which has exactly $N$ photons but may be non-zero for a state that is a linear combination of an $N$ photon state with an $N+1$ photon state. Thus, classical electromagnetic fields actually only arise due to states that are linear combinations of states of different photon numbers.

## 1.11 Summary and Additional Reading

Important homework questions from this chapter (the QHO questions are **mandatory**):

- **Problem 1.3: Silver Bullet Solution to the Schrödinger equation** and
- **Problem 1.7: Time Evolution Operator** establish the basics of how to solve the time evolution of a general quantum mechanical system.
- **Problem 1.10: Ehrenfest's Theorem and the Quantum Noether Theorem** derives the special form that Noether's theorem takes in quantum mechanics.
- **Problem 1.11: Active vs. Passive Picture** introduces the active description of quantum mechanics, in which the state vector of the system is constant and all of the time evolution is shuffled into the operators and shows that it is equivalent to the more typical passive picture, in which the state vector is dynamical and satisfies the Schrödinger equation.
- **Problem 1.12: Commutators becoming derivatives** shows how commutators of an operator with a variable effectively differentiates that operator with respect to the conjugate to that variable.
- **Problem 1.13: Restoring factors of $\hbar$** explains how to remove and how to restore factors of $\hbar$ when going between natural units in which $\hbar = 1$ and unnatural unit systems in which it is not.
- **Problem 1.21: Baker–Campbell–Hausdorff formula** derives the basic extent to which the operators in the exponents of a product of operator exponentials fail to add.
- **Problem 1.24: First Look at Coherent States** introduces the "maximally classical" coherent states which are as simultaneously localized in both position and momentum space as possible.
- **Problem 1.26: Double-Slit Experiment Revisited** explains the experimental results of the high-rate double slit experiment in terms of the wavefunction, resolving the tension of Orig. section 4.6.4.
- **Problem 1.31: Aharanov-Bohm Effect** derives experimental consequences which involve the local interaction of the wavefunction with the vector potential, not simply its gauge invariant byproducts, thus demonstrating unambiguously that the vector potential is genuinely physical.
- **Problem 1.32: Generator of Gauge Transformations** derives the canonical structure of gauge transformations in the interacting theory of a single charged matter wave with electromagnetism.
- **Problem 1.33: Quantum Harmonic Oscillator (QHOI) - Creation and Annihilation Operators** and
- **Problem 1.34: QHOII - Moment Evolution Equations** and
- **Problem 1.35: QHOIII - Evolution of Coherent States and Propagator** and
- **Problem 1.36: QHOIV - Driven Oscillations** and
- **Problem 1.37: QHOV - Multivariable Oscillator** and
- **Problem 1.38: QHOVI - Fully general quadratic Hamiltonian** derive all of the basic properties of a quantum harmonic oscillator, which is the most important quantum system.
- **Problem 1.49: Solving the Hydrogen Atom** shows how the spectrum of the hydrogen atom is

directly determined by the Schrödinger equation, in excellent agreement with experiment.

- **Problem 1.50: Atomic Spectral Lines and Measuring $\hbar$** shows how the atomic spectra can be measured experimentally and how the placement of spectral lines can be used to infer the value of $\hbar$ in unnatural units or calibrate the units of the action (so that $\hbar = 1$) in natural units.
- **Problem 1.51: Photoelectric Effect** gives another way of measuring $\hbar$ and derives the photoelectric effect, which is an experimental phenomenon that is not explained by classical mechanics.
- **Problem 1.52: Fine-Structure Constant** introduces the fine-structure constant $\alpha$, which is a fundamental unitless physical constant that determines the strength of the electromagnetic interaciton. From $\alpha$, we defined the elementary electric charge `e`.
- **Problem 1.60: Stern-Gerlach Experiment** describes and explains how the spin of electrons can be measured and how the presence of an inhomogeneous magnetic field prevents the classical limit from being a good approximation for the dynamics of an electron wave-packet.
- **Problem 1.73: Rutherford Scattering Revisited** shows that the scattering of one charged particle off of another leads to exactly the same differential cross section in both classical and quantum mechanics.
- **Problem 1.85: Addition of Angular Momentum in general** defines the Clebsch-Gordan coefficients and explains how to construct the total angular momentum Hilbert space when two different systems in their own angular momentum representations are brought together into one system.

In this chapter we developed the theory of single particle quantum mechanics. This theory replaced the classical point-particle description of matter by a wave description, i.e. a classical field theory of a single particle in terms of the wavefunction $\psi$ which satisfies the Schrödinger equation. Because that classical field is not perfectly point-localized, it has finite interactions with electromagnetism, fixing the problem of point-particle divergences. In section 1.1, we introduced the Dirac equation and the concept of a Hilbert space to help us cast classical field theory in the language of linear algebra. In section 1.2, we showed that a generic stable lowest order classical field theory when put into the right coordinates simply becomes the Schrödinger equation with some self-adjoint Hamiltonian operator, introduced $\hbar$ as the constant relating the energy of a normalized waves to its angular frequency, and showed that projective intensities ("quantum probabilities") are the adiabatic invariants of quantum mechanics. In section 1.3 we proved Wigner's theorem that all time nonreversing quantum canonical transformations are implemented by a unitary transformation and all time reversing such transformations are implemented by a unitary transformation and a complex conjugation. In section 1.4 we showed that all continuous transformation generators in quantum mechanics take the form of an operator summary and showed how quantum tomography can be used to reconstruct the entire wavefunction up to an overall phase from those operator summaries. In section 1.5 we found the position and momentum operators as the natural quantities whose summaries track the wavefunction's center of mass and its total linear momentum. We studied the basics of the classical limit and the Born rule. In section 1.6 we finally found the specific Hamiltonian operator of a nonspinning nonrelativistic charged particle. We also studied the geometry of gauge transformations in which the vector potential provides the connection on the wavefunction's internal complex space with the electric and magnetic fields providing the internal curvature. In section 1.7 we derived the basic facts of the quantum theory of angular momentum in which we found that rotations of the wavefunction are captured by the SU(2) group. We classified all irreducible representations of SU(2), computed the angular momentum operators in each representation, and studied the spherical harmonics. In section 1.8 we found the specific Hamiltonian operator of a spin-1/2 nonrelativistic charged particle. In section 1.9 we derived the basics of the theory of quantum scattering, which is the most powerful way in which we can access quantum properties experimentally. In section 1.10 we discussed various experimental troubles faced by naive applications of single particle quantum mechanics and found that these troubles as well as our previous confusion around black-body radiation can all be explained by quantum field theory, the deep properties of which will be the subject of the rest of our studies.

This chapter is simply not a complete treatment of essential topics in single particle quantum mechanics. Refer to [2, 33, 48, 51, 57, 63] for general treatments. For more on single particle quantum mechanics as a classical field theory, refer to [25]. For more on the interpretative elements on quantum mechanics (which we will also discuss thoroughly in the next chapter), refer to [38]. There are many more specialized and technical topics in basic quantum mechanics not covered much in any of these. Because nonrelativistic quantum mechanics of one or a few particles remains actively important for scientific research and is so young compared to classical mechanics, there are far too many specialized applications for any one book to cover.

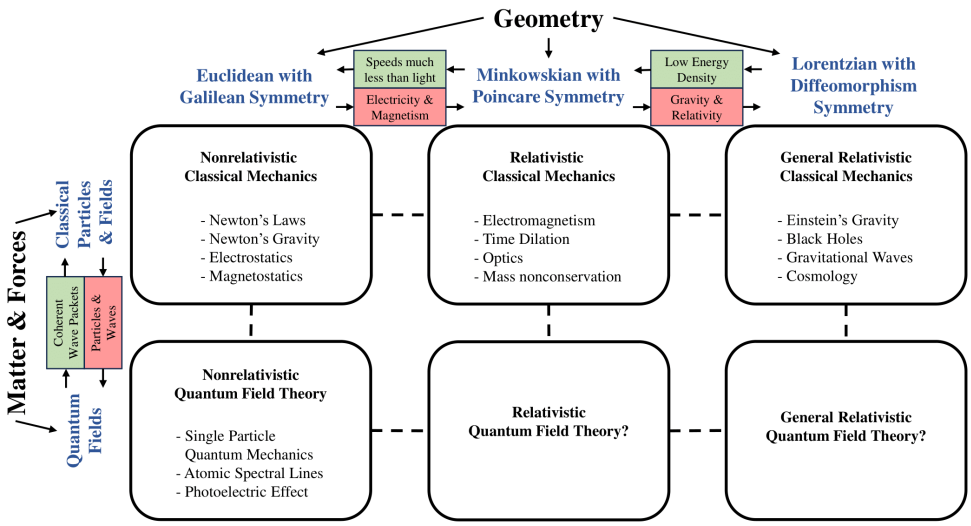



**Figure 1.1:** Our perspective on physics as of the end of this chapter. Replacing our point particle theory of matter with a wave theory has completely solved the problem of point particle divergences. However, so far we have only done so using Galilean symmetry. We know that to combine Lorentz symmetry, quantum mechanics, and the correct multiparticle statistics, we need to use quantum field theory.

# Chapter 2

# Quantum Mechanics in General

*Ever since the advent of modern quantum mechanics in the late 1920's, the idea has been prevalent that the classical laws of probability cease, in some sense, to be valid in the new theory. More or less explicit statements to this effect have been made in large number and by many of the most eminent workers in the new physics. Some authors have even gone farther and stated that the formal structure of logic must be altered to conform to the terms of reference of quantum physics. Such a thesis is surprising, to say the least, to anyone holding more or less conventional views regarding the positions of logic, probability, and experimental science: many of us have been apt – perhaps too naively – to assume that experiments can lead to conclusions only when worked up by means of logic and probability, whose laws seem to be on a different level from those of physical science. The primary object of this presentation is to show that the thesis in question is entirely without validity and is the product of a confused view of the laws of probability.*

– Bernard Koopman [28]

*Something [is] rotten in the eigenstate of Denmark.*

– Adam Becker [3]

*Niels Bohr brain-washed a whole generation of physicists into believing that the problem had been solved fifty years ago.*

– Murray Gell-Mann [27]

We discovered at the end of the previous chapter that quantum mechanics can be more general than simply replacing our point particle theory of matter by a wave theory. In this chapter we will understand that quantum mechanics is actually a second branch of possibilities that we sped past all the way back in Orig. chapter 1, though I did warn you. In order to get there, we need to remember our line of thinking in Orig. section 1.3. This second branch of possibilities combined with the principles of locality and relativity will lead naturally to quantum field theory, which is the formalism that is at the foundation of our status quo understanding of physics.

## 2.1 Why Quantum Mechanics?

### 2.1.1 Quantum Ex Nihilo

There is another way of arriving at exactly the same Lagrangian and Hamiltonian as in section 1.2 but by apparently very different thinking. In Orig. section 1.3 we found that whatever dynamical variables we have of interest, we can encode all of them as well as all transformation generators transforming those variables as linear operators (for example an operator $\hat{F}$) so that under the transformation generated by some

other generator (say $\hat{G}$, with transformation parameter $\beta$):

$$\frac{d\hat{F}}{d\beta} = [\hat{G}, \hat{F}]. \tag{2.1.1}$$

This operator formalism is totally abstract and in order to do physics we need some way to map it into numbers. In Orig. section 1.4 we found that if we wish to represent these operators with real valued variables and functions then result is the Poisson representation. However, this is not the only possible representation. In fact, it should be clear that the Poisson representation is not a faithful representation of the linear operators. In particular, if the linear operators $\hat{F}$ and $\hat{G}$ do not commute, then $\hat{F}\hat{G}$ and $\hat{G}\hat{F}$ are of course different objects while if $\hat{F}$ and $\hat{G}$ are represented by phase space functions $f$ and $g$ in the Poisson representation, then the products $fg$ and $gf$ are the same. We will prove more directly and explicitly in section 2.2 that there is no bijection between our linear operators and real valued functions that is compatible with the Lie bracket. In the rest of this section we will discuss a faithful representation of the objects of Orig. section 1.3.

The generators as written in (2.1.1) are normalized the way mathematicians typically normalize them[1]. However, for physics that normalization is inconvenient. Let us label the generators as mathematicians normalize them with a subscript $L$, short for Lie:

$$\frac{d\hat{F}_L}{d\beta} = [\hat{G}_L, \hat{F}_L]. \tag{2.1.2}$$

The coordinate operators $\hat{Q}^a$ of Orig. section 1.3 were defined so that when acting on a function of the coordinates at a particular coordinate value they simple multiply by the coordinate value. Suppose we have $d$ coordinate operators $\hat{Q}^1, ..., \hat{Q}^d$. Due to the commutativity of multiplication, these coordinate operators automatically commute with each other and due to the use of real coordinates, these coordinate operators are automatically self-adjoint. Then, consider a Dirac delta centered at coordinates $Q^a$, which as usual we write as a vector using the Dirac notation $|Q\rangle$. Then, at the coordinate center $\hat{Q}^a|Q\rangle = Q^a|Q\rangle$ because the coordinate operators multiply the coordinate value by definition and away from the coordinate center the Dirac delta vanishes and so both sides of this equation evaluate to 0. Thus:

$$\hat{Q}^a|Q\rangle = Q^a|Q\rangle, \qquad [\hat{Q}^a, \hat{Q}^b] = 0, \qquad \hat{Q}^{a\dagger} = \hat{Q}^a. \tag{2.1.3}$$

Hence these coordinate operators $\hat{Q}^a$ are exactly like the $\hat{x}^a$ operators found in the previous chapter. Automatically then we can recycle the identities:

$$\langle Q|Q'\rangle = \delta(Q - Q'), \qquad \int |Q\rangle\langle Q| d^dQ = \mathbb{1}, \qquad \hat{Q}^a = \int Q^a |Q\rangle\langle Q| d^dQ. \tag{2.1.4}$$

If $|\Psi\rangle$ is a function of $Q$ represented as a vector, then we can also recycle:

$$|\Psi\rangle = \int \Psi(Q)|Q\rangle d^dQ, \qquad \Psi(Q) = \langle Q|\Psi\rangle. \tag{2.1.5}$$

The momentum operators of Orig. section 1.3 are generators and so we write them as $\hat{P}_{La}$ and they were defined so that they differentiate functions of the coordinates with respect to the appropriate coordinate. Written in the Dirac notation, this is simply:

$$\langle Q|\hat{P}_{La}|\Psi\rangle = \frac{\partial\Psi(Q)}{\partial Q^a}. \tag{2.1.6}$$

Thus the $\hat{P}_{La}$ are not quite exactly like the $\hat{p}_a$ operators of the previous chapter. As defined, $\hat{P}_{La}$ is anti-self-adjoint as it is equal to the negative of its own adjoint (the derivative picks up a negative sign under integration by parts). If we include an $i$ in the definition of the momentum operators, then like the $\hat{p}_a$ of the previous chapter they will be self-adjoint. Thus, we introduce the physicist normalization (which we write without a subscript) by:

$$\hat{G}_a = -i\hbar\hat{G}_{La} \tag{2.1.7}$$

for a generic transformation generator $\hat{G}_a$. The factor of $\hbar$ is there to control units of the action. The momentum conjugate to a coordinate $Q$ has units of action divided by the units of $Q$ while the mathematician's momentum operator $\hat{P}_{La}$ has units of $\frac{1}{Q}$, thus effectively being a convention in which the action is unitless. The choice of the value of $\hbar$ as some specific real number implicitly defines the units of the action or conversely

[1]There are possible convention differences by a minus sign between sources.

so. So, we write the momentum operators $\hat{P}_a$ so that:

$$\langle Q|\hat{P}_a|\Psi\rangle = -i\hbar\frac{\partial\Psi(Q)}{\partial Q^a}. \tag{2.1.8}$$

From this we can automatically recycle:

$$\hat{P}_a = \int |Q\rangle\left(-i\hbar\frac{\partial}{\partial Q^a}\right)\langle Q|d^dQ, \qquad [\hat{Q}^a,\hat{P}_b] = i\hbar\delta^a_b\mathbb{1}, \qquad [\hat{P}_a,\hat{P}_b] = 0, \qquad \hat{P}_a^\dagger = \hat{P}_a. \tag{2.1.9}$$

As well, we already know we can diagonalize the $\hat{P}_a$ operators and we know how to do so:

$$|P\rangle = \int \frac{e^{\frac{i}{\hbar}Q^aP_a}}{(2\pi\hbar)^{\frac{d}{2}}}|Q\rangle d^dQ, \qquad \langle P|P'\rangle = \delta(P-P'), \qquad \int |P\rangle\langle P|d^dP = \mathbb{1}, \qquad \hat{P}_a = \int P_a|P\rangle\langle P|d^dP. \tag{2.1.10}$$

Using the physicist normalization for generators, (2.1.1) becomes:

$$\frac{d\hat{F}}{d\beta} = \frac{1}{i\hbar}[\hat{F},\hat{G}]. \tag{2.1.11}$$

We can reorganize the generating operators of Orig. section 1.3 into a form that looks suggestive and familiar knowing quantum mechanics, but that isn't exactly the same as arriving at quantum mechanics. The next step is where the real magic happens. So far we have been discussing these abstract mathematical objects as linear operators. Experimenters do not read out linear operators from their equipment; they measure numbers. Therefore, somehow these linear operators have to map somehow to a number that represents the physical property the experimenter desires to measure. That is, there must be some map from our operators into an appropriate space of numbers (we will take that space to be the complex numbers $\mathbb{C}$, as they contain the usual real numbers and sometimes provide convenient additional structure) with the output numerical value of this map meant to represent the experimentally relevant number assigned to the linear operator under consideration. Let the map from the operators into complex numbers be denoted $\langle\cdot\rangle$ so that $\langle\hat{F}\rangle$ is a complex number. We will call $\langle\hat{F}\rangle$ the summary of $\hat{F}$. This notation obviously evokes the memory of the operator summary notation used in the last chapter, but pretend you don't know what an operator summary means in that way for the time being.

Many of our quantities of interest are vectors in some vector space (such as literal spatial vectors or 4-vectors, higher rank tensors, et cetera) and so we express them as linear combinations of basis vectors in whatever appropriate space. The summary map maps operators into complex numbers. So, the natural definition for a summary of a vector of operators is to make it the vector whose components are the summary of each of those respective operators. That is:

$$\hat{\vec{F}} = \hat{F}^a\vec{\iota}_a \implies \langle\hat{\vec{F}}\rangle \stackrel{\text{def}}{=} \langle\hat{F}^a\rangle\vec{\iota}_a. \tag{2.1.12}$$

Under a change of basis for some real coefficients $M^a{}_b$:

$$\vec{\iota}_a = M^b{}_a\vec{\iota}'_b, \qquad \hat{\vec{F}} = \hat{F}^a\vec{\iota}_a = \hat{F}^bM^a{}_b\vec{\iota}'_a = \hat{F}'^a\vec{\iota}'_a \implies \hat{F}'^a = \hat{F}^bM^a{}_b. \tag{2.1.13}$$

A vector in a vector space is the same object whatever basis is used. Therefore, to respect the vector space structure, we need the summary of a vector to be independent of the basis chosen. Thus:

$$\langle\hat{\vec{F}}\rangle = \langle\hat{F}^a\rangle\vec{\iota}_a \stackrel{!}{=} \langle\hat{F}'^a\rangle\vec{\iota}'_a \implies \langle\hat{F}'^a\rangle\vec{\iota}'_a = \langle\hat{F}^a\rangle\vec{\iota}_a \implies \langle\hat{F}^bM^a{}_b\rangle\vec{\iota}'_a = \langle\hat{F}^b\rangle M^a{}_b\vec{\iota}'_a \implies \langle\hat{F}^bM^a{}_b\rangle = \langle\hat{F}^b\rangle M^a{}_b. \tag{2.1.14}$$

This must be true for all change of basis coefficients $M^a{}_b$. Component by component this is just the linearity property. That is, in order for the summary map to respect the basis independence of vectors, it must be linear over complex numbers:

$$\langle a\hat{F}+b\hat{G}\rangle = a\langle\hat{F}\rangle + b\langle\hat{G}\rangle \tag{2.1.15}$$

and so we will require linearity of the summary map going forward.

The linear operators which we are taking the summary of act on some complex vector space (for example, on some appropriate set of functions of the coordinates $Q$). Let us call the associated vector space $\mathbb{H}$. We can equip that vector space with a complex inner product (for functions on the coordinates, the usual integral expression for the inner product of complex valued functions). Then, with that inner product, we can develop an orthonormal basis of vectors $|n\rangle$ for $\mathbb{H}$ with usual adjoint vectors $\langle n|$. Due to our desire to be able to take limits in physics, we require $\mathbb{H}$ to be complete, and thus we have that $\mathbb{H}$ is a Hilbert space. Then, the

orthonormality and completeness relations read:

$$\langle n|m\rangle = \delta_{nm}, \qquad \sum_n |n\rangle\langle n| = \mathbb{1}. \tag{2.1.16}$$

As well, for any operator $\hat{F}$ we can define the matrix elements of that operator by:

$$F_{nm} \stackrel{\text{def}}{=} \langle n|\hat{F}|m\rangle. \tag{2.1.17}$$

Then, any operator may be expressed as:

$$\hat{F} = \sum_n |n\rangle\langle n|\hat{F}\sum_m |m\rangle\langle m| = \sum_{nm} |n\rangle\langle n|\hat{F}|m\rangle\langle m| = \sum_{nm} = F_{nm}|n\rangle\langle m|. \tag{2.1.18}$$

It is useful to define the trace of an operator by:

$$\mathrm{Tr}(\hat{F}) \stackrel{\text{def}}{=} \sum_n \langle n|\hat{F}|n\rangle = \sum_n F_{nn}. \tag{2.1.19}$$

Naively it appears that the trace is a basis dependent quantity. However, consider a new basis $|n'\rangle$ which is also orthonormal and complete. Then, this basis must have some expansion in terms of the old basis elements:

$$|n'\rangle = \sum_m C_{n'm}|m\rangle. \tag{2.1.20}$$

Define the change of basis operator $\hat{U}$ so that it has matrix elements $C_{n'm}$:

$$U_{mn} = \langle m|\hat{U}|n\rangle = C_{n'm} \implies |n'\rangle = \sum_m |m\rangle\langle m|\hat{U}|n\rangle = \mathbb{1}\hat{U}|n\rangle = \hat{U}|n\rangle. \tag{2.1.21}$$

Then the orthonormality condition becomes:

$$\langle m'|n'\rangle = \langle m|\hat{U}^\dagger\hat{U}|n\rangle = \delta_{mn} \stackrel{!}{=} \langle m|n\rangle \implies \langle m|\left(\hat{U}^\dagger\hat{U} - \mathbb{1}\right)|n\rangle = 0 \tag{2.1.22}$$

$$\implies \sum_{mn} |m\rangle\langle m|\left(\hat{U}^\dagger\hat{U} - \mathbb{1}\right)|n\rangle\langle n| = 0 \implies \hat{U}^\dagger\hat{U} = \mathbb{1}. \tag{2.1.23}$$

The completeness condition becomes:

$$\sum_n |n'\rangle\langle n'| \stackrel{!}{=} \mathbb{1} \implies \sum_n \hat{U}|n\rangle\langle n|\hat{U}^\dagger = \mathbb{1} \implies \hat{U}\hat{U}^\dagger = \mathbb{1}. \tag{2.1.24}$$

Hence, $\hat{U}$ must be a unitary operator. For such a change of basis, the transformed trace is then:

$$\mathrm{Tr}(\hat{F})' = \sum_n \langle n'|\hat{F}|n'\rangle = \sum_n \langle n|\hat{U}^\dagger\hat{F}\hat{U}|n\rangle = \sum_{nm} \langle n|\hat{U}^\dagger|m\rangle\langle m|\hat{F}\hat{U}|n\rangle \tag{2.1.25}$$

$$= \sum_{nm} \langle m|\hat{F}\hat{U}|n\rangle\langle n|\hat{U}^\dagger|m\rangle = \sum_m \langle m|\hat{F}\hat{U}\mathbb{1}\hat{U}^\dagger|m\rangle = \sum_n \langle n|\hat{F}|n\rangle = \mathrm{Tr}(\hat{F}). \tag{2.1.26}$$

Thus, the trace of an operator is in fact basis independent. There is a subtly here. In an infinite dimensional vector space, a change of basis and the trace both involve infinite series which may or may not converge absolutely. Because of this, there may be ways to infinitely rearrange the series so that the result of the trace operation appears basis dependent. The convergence of these series depend on the matrix elements of the operator whose trace is being taken, and so there are operators for which the trace is not well-defined. We are only allowed to properly apply the trace to so-called trace-class operators, which are operators whose trace series/integrals converge absolutely. Going forward if we take the trace of an operator, we will always be assuming it is trace-class. A very simple example of non-trace class operators is given in the homework. Another important property of the trace is cyclicity. For any three operators $\hat{A}, \hat{B}, \hat{C}$:

$$\mathrm{Tr}(\hat{A}\hat{B}\hat{C}) = \sum_n \langle n|\hat{A}\hat{B}\hat{C}|n\rangle = \sum_{nml} \langle n|\hat{A}|m\rangle\langle m|\hat{B}|l\rangle\langle l|\hat{C}|n\rangle = \sum_{nml} \langle l|\hat{C}|n\rangle\langle n|\hat{A}|m\rangle\langle m|\hat{B}|l\rangle$$

$$= \sum_{nml} \langle n|\hat{C}|m\rangle\langle m|\hat{A}|l\rangle\langle l|\hat{B}|n\rangle = \mathrm{Tr}(\hat{C}\hat{A}\hat{B}). \tag{2.1.27}$$

Due to (2.1.15), $\langle\hat{F}\rangle$ is a linear functional of $\hat{F}$. Because $\hat{F}$ depends linearly on its matrix elements and $\langle\hat{F}\rangle$ depends linearly on $\hat{F}$, $\langle\hat{F}\rangle$ also depends linearly on the matrix elements of $\hat{F}$. So, for some coefficients $\rho^{nm}$ which are independent of $\hat{F}$:

$$\langle\hat{F}\rangle = \sum_{nm} \rho_{nm}F_{mn}. \tag{2.1.28}$$

Define then the operator $\hat{\rho}$, called the **density matrix**, to have matrix elements $\rho_{nm}$:

$$\rho_{nm} = \langle n|\hat{\rho}|m\rangle \implies \langle \hat{F}\rangle = \sum_{nm}\langle n|\hat{\rho}|m\rangle\langle m|\hat{F}|n\rangle = \sum_{n}\langle n|\hat{\rho}\hat{F}|n\rangle = \mathrm{Tr}(\hat{\rho}\hat{F}). \tag{2.1.29}$$

Hence, there must be some distinguished operator $\hat{\rho}$ which characterizes the summary map so that the summary of any operator can always be written as:

$$\boxed{\langle \hat{F}\rangle = \mathrm{Tr}(\hat{\rho}\hat{F}).} \tag{2.1.30}$$

This is an essential discovery. The entire relationship between our abstract operator formalism and the numbers we wish to measure is characterized by this single operator $\hat{\rho}$. Let us now study its properties.

Consider the identity operator. Because the identity operator is always 1 when acting on any vector or multiplied on any operator, its effective numerical value is necessarily 1. Thus, its summary must be 1:

$$\langle \mathbb{1}\rangle = \mathrm{Tr}(\hat{\rho}\mathbb{1}) = \mathrm{Tr}(\hat{\rho}) \stackrel{!}{=} 1. \tag{2.1.31}$$

And so, the density matrix must have a trace of 1.

Consider our self-adjoint coordinate operators $\hat{Q}^a$. Because they are intended to represent real-valued coordinates, they must have real-valued operator summaries. However, physics knows nothing about what coordinates we choose to describe it with, so any other choice of self-adjoint coordinate operators must also have real-valued operator summaries. Consider some unspecified self-adjoint generator $\hat{G}$ and the transformation given by using $\hat{G}$ as a generator:

$$\frac{d\hat{Q}^a}{d\beta} = \frac{1}{i\hbar}[\hat{Q}^a, \hat{G}], \qquad \frac{d\hat{P}_a}{d\beta} = \frac{1}{i\hbar}[\hat{P}_a, \hat{G}]. \tag{2.1.32}$$

Because $\hat{Q}^a$ and $\hat{G}$ are self-adjoint, automatically so is $\frac{1}{i\hbar}[\hat{Q}^a, \hat{G}]$. Therefore, $\frac{d\hat{Q}^a}{d\beta}$ is self-adjoint and so the transformed coordinate is always self-adjoint. Thus, the operator summary of $\frac{d\hat{Q}^a}{d\beta}$ must be real-valued for any choice of $\hat{G}$, as that represents a valid coordinate transformation. For a given value of $a$ (say 1), the actual operator on the right hand side defining $\frac{d\hat{Q}^1}{d\beta}$ can be made to be any self-adjoint operator by appropriate choice of $\hat{G}$. Therefore, for $\frac{d\hat{Q}^1}{d\beta}$ to have real operator summary for any choice of self-adjoint $\hat{G}$, it must be the case that the operator summary of any self-adjoint operator is itself real.

Let $\hat{F}$ be any operator and split it into self-adjoint and anti-self-adjoint parts:

$$\mathrm{Sa}(\hat{F}) = \frac{\hat{F}+\hat{F}^\dagger}{2}, \qquad \mathrm{As}(\hat{F}) = \frac{\hat{F}-\hat{F}^\dagger}{2i} \implies \hat{F} = \mathrm{Sa}(\hat{F}) + i\mathrm{As}(\hat{F}). \tag{2.1.33}$$

Since both $\mathrm{Sa}(\hat{F})$ and $\mathrm{As}(\hat{F})$ are self-adjoint, they necessarily have real operator summaries:

$$\langle \mathrm{Sa}(\hat{F})\rangle^* = \langle \mathrm{Sa}(\hat{F})\rangle, \qquad \langle \mathrm{As}(\hat{F})\rangle^* = \langle \mathrm{As}(\hat{F})\rangle. \tag{2.1.34}$$

From this and linearity it immediately follows that:

$$\langle \hat{F}\rangle^* = \langle (\mathrm{Sa}(\hat{F}) + i\mathrm{As}(\hat{F}))\rangle^* = \langle \mathrm{Sa}(\hat{F})\rangle - i\langle \mathrm{As}(\hat{F})\rangle = \langle \hat{F}^\dagger\rangle \tag{2.1.35}$$

Looking at each side in components:

$$\langle \hat{F}\rangle^* = \sum_{mn}(\rho_{nm})^* F^*_{mn}, \qquad \langle \hat{F}^\dagger\rangle = \sum_{mn}\rho_{nm}F^*_{nm} \tag{2.1.36}$$

$$\implies \sum_{mn}\rho^*_{nm}F^*_{mn} = \sum_{mn}\rho_{mn}F^*_{mn} \implies \sum_{mn}(\rho_{mn}-\rho^*_{nm})F^*_{mn} = 0, \qquad \forall F_{mn} \tag{2.1.37}$$

$$\implies \rho_{mn} = \rho^*_{nm} \implies \hat{\rho} = \hat{\rho}^\dagger. \tag{2.1.38}$$

Hence, requiring $\langle \hat{Q}^a\rangle$ to be real for any choice of coordinates demands that the density matrix $\hat{\rho}$ is self-adjoint.

Finally let us consider a nonnegative function of our coordinates $f(\hat{Q})$. That is, $f(Q)$ is a function such that for any real valued $Q$, $f(Q) \geq 0$. Because $\langle f(\hat{Q})\rangle$ is intended to represent the physical value of this function and this function is nonnegative, we need $\langle f(\hat{Q})\rangle$ to be nonnegative. Because this $f$ may be any nonnegative function and we may perform arbitrary coordinate transformations of our coordinates which keep the coordinate operators self-adjoint, we really need $\langle \hat{F}\rangle \geq 0$ for every nonnegative operator $\hat{F}$. Because $\hat{\rho}$ is self-adjoint, the spectral theorem guarantees it is diagonalizable. Let $|\rho; n\rangle$ denote an orthonormal eigenbasis

of $\hat{\rho}$, with $n$ a label, and let $\rho_n$ indicate the eigenvalues of $\hat{\rho}$:

$$\hat{\rho}|\rho;n\rangle = \rho_n|\rho;n\rangle, \qquad \langle\rho;n|\rho;n'\rangle = \delta_{nn'}, \qquad \sum_n |\rho;n\rangle\langle\rho;n| = \mathbb{1} \implies \hat{\rho} = \sum_n \rho_n|\rho;n\rangle\langle\rho;n|. \tag{2.1.39}$$

The operator $|\rho;n\rangle\langle\rho;n|$ is automatically nonnegative for each $n$ because for any vector $|\psi\rangle$:

$$\langle\psi|\left(|\rho;n\rangle\langle\rho;n|\right)|\psi\rangle = |\langle\psi|\rho;n\rangle|^2 \geq 0. \tag{2.1.40}$$

Thus, we need $\langle|\rho;n\rangle\langle\rho;n|\rangle$ to be nonnegative:

$$0 \leq \langle|\rho;n\rangle\langle\rho;n|\rangle = \mathrm{Tr}(\hat{\rho}|\rho;n\rangle\langle\rho;n|) = \rho_n. \tag{2.1.41}$$

Therefore, the eigenvalues of $\hat{\rho}$ are nonnegative. Consequently, for any vector $|\psi\rangle$:

$$\langle\psi|\hat{\rho}|\psi\rangle = \sum_n \rho_n|\langle\psi|\rho;n\rangle|^2 \geq 0 \tag{2.1.42}$$

and so $\hat{\rho}$ is a necessarily a nonnegative operator.

The operator Lie bracket formalism is the most general way to write down continuous transformations. In that formalism every quantity is associated with a linear operator. The requirement that experiments yield numbers demands that there exists a map from those operators into the complex numbers which we call the operator summary. The requirement that the summary of a vector transforms according to the change of basis formula for a vector demands that the summary depends linearly on the operator one is taking the summary of. This linearity requirement implies that the summary can be written as a trace of the product of the operator in question with a distinguished operator called the density matrix $\hat{\rho}$. The requirement that the summary of a constant number be that constant implies that the density matrix has trace 1. The requirement that the summary of a self-adjoint operator be real implies that the density matrix is self-adjoint. The requirement that the summary of a nonnegative operator be nonnegative implies that the density matrix is nonnegative. So, we have:

$$\boxed{\begin{aligned} \langle\hat{F}\rangle &= \mathrm{Tr}(\hat{\rho}\hat{F}), \qquad \forall\hat{F} && (2.1.43)\\ \mathrm{Tr}(\hat{\rho}) &= 1 && (2.1.44)\\ \hat{\rho} &= \hat{\rho}^\dagger && (2.1.45)\\ \langle\psi|\hat{\rho}|\psi\rangle &\geq 0, \qquad \forall|\psi\rangle. && (2.1.46) \end{aligned}}$$

Time is itself a continuous parameter and so time evolution is a continuous transformation with some generator which we call $\hat{H}_A$ (glancing ahead, this will be the Hamiltonian operator in the active picture). Because time is a real valued parameter its generator is real valued and so $\hat{H}_A$ must be self-adjoint. Then, any operator corresponding to a generator (call it $\hat{F}_A$) within our system evolves in time according to (2.1.11):

$$\frac{d\hat{F}_A}{dt} = \frac{1}{i\hbar}[\hat{F}_A, \hat{H}_A]. \tag{2.1.47}$$

This is precisely the evolution equation for an operator in the active picture of quantum mechanics. Define the operator $\hat{U}(t,t_0)$, called the time evolution operator, by the conditions:

$$\hat{U}(t_0,t_0) = \mathbb{1}, \qquad i\hbar\frac{d\hat{U}(t,t_0)}{dt} = \hat{U}(t,t_0)\hat{H}_A(t). \tag{2.1.48}$$

Then, we can see that at $t = t_0$ the operator $\hat{U}$ is unitary:

$$\hat{U}^\dagger(t_0,t_0)\hat{U}(t_0,t_0) = \mathbb{1}, \qquad \hat{U}(t_0,t_0)\hat{U}^\dagger(t_0,t_0) = \mathbb{1}. \tag{2.1.49}$$

As well, we can see that unitarity is preserved by the equation of motion:

$$\frac{d}{dt}\left(\hat{U}\hat{U}^\dagger\right) = \frac{d\hat{U}}{dt}\hat{U}^\dagger + \hat{U}\frac{d\hat{U}^\dagger}{dt} = \left(\frac{1}{i\hbar}\hat{U}\hat{H}_A\right)\hat{U}^\dagger + \hat{U}^\dagger\left(\frac{1}{i\hbar}\hat{U}\hat{H}_A\right)^\dagger = -\frac{1}{i\hbar}\hat{U}\hat{H}_A\hat{U}^\dagger + \frac{1}{i\hbar}\hat{U}\hat{H}_A\hat{U}^\dagger = 0. \tag{2.1.50}$$

$$\frac{d}{dt}\left(\hat{U}^\dagger\hat{U}\right) = \frac{d\hat{U}^\dagger}{dt}\hat{U} + \hat{U}^\dagger\frac{d\hat{U}}{dt} = \left(\frac{1}{i\hbar}\hat{U}\hat{H}_A\right)^\dagger\hat{U} + \hat{U}^\dagger\frac{1}{i\hbar}\hat{U}\hat{H}_A = -\frac{1}{i\hbar}\hat{H}_A\hat{U}^\dagger\hat{U} + \frac{1}{i\hbar}\hat{U}^\dagger\hat{U}\hat{H}_A = 0 \tag{2.1.51}$$

The first one is 0 automatically using that $\hat{H}_A$ is self-adjoint. The second is 0 because we see that $\hat{U}^\dagger\hat{U} - \mathbb{1}$ satisfies a linear homogeneous equation of motion in terms of itself and that combination is initially 0 and thus remains 0 for all time under homogeneous linear evolution. So, $\hat{U}(t,t_0)$ is unitary for all $t, t_0$. From here we can immediately define the passive picture Hamiltonian operator $\hat{H}(t)$ written without a subscript:

$$\hat{H}(t) = \hat{U}(t,t_0)\hat{H}_A(t)\hat{U}^\dagger(t,t_0) \tag{2.1.52}$$

$$\implies \hat{H}(t)\hat{U}(t,t_0) = \hat{U}(t,t_0)\hat{H}_A(t). \tag{2.1.53}$$

So, (2.1.48) can equivalently be written as:

$$i\hbar\frac{d\hat{U}(t,t_0)}{dt} = \hat{H}(t)\hat{U}(t,t_0). \tag{2.1.54}$$

Recycling (1.9.24) we can in principle explicitly solve for $\hat{U}(t,t_0)$ using the time ordering transformation:

$$\hat{U}(t,t_0) = \mathscr{T} : e^{-\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t')dt'}. \tag{2.1.55}$$

For an operator $\hat{F}_A(t)$ representing some generator within our system evolving with time we see that:

$$\frac{d}{dt}\left(\hat{U}(t,t_0)\hat{F}_A(t)\hat{U}^\dagger(t,t_0)\right) = \frac{d\hat{U}}{dt}\hat{F}_A\hat{U}^\dagger + \hat{U}\frac{d\hat{F}_A}{dt}\hat{U}^\dagger + \hat{U}\hat{F}_A\frac{d\hat{U}^\dagger}{dt} \tag{2.1.56}$$

$$= \frac{1}{i\hbar}\hat{U}\hat{H}_A\hat{F}_A\hat{U}^\dagger + \hat{U}\frac{1}{i\hbar}(\hat{F}_A\hat{H}_A - \hat{H}_A\hat{F}_A)\hat{U}^\dagger - \frac{1}{i\hbar}\hat{F}_A\hat{H}_A\hat{U}^\dagger \tag{2.1.57}$$

$$= 0. \tag{2.1.58}$$

So, this combination is constant. At $t = t_0$ we see:

$$\hat{U}(t_0,t_0)\hat{F}_A(t_0)\hat{U}^\dagger(t_0,t_0) = \mathbb{1}\hat{F}_A(t_0)\mathbb{1} = \hat{F}_A(t_0) \tag{2.1.59}$$

$$\implies \hat{F}_A(t) = U^\dagger(t,t_0)\hat{F}_A(t_0)\hat{U}(t,t_0). \tag{2.1.60}$$

Hence, the time evolution operator explicitly solves the evolution of any generator within our system by this form. This is precisely the evolution of an operator in the active picture, and so we can honestly interpret $\hat{F}_A(t)$ with this time dependence as the active picture operator and define $\hat{F} = \hat{F}_A(t_0)$ as the passive picture operator.

As we have discussed the density matrix so far, it simply exists to define the summary map $\langle\cdot\rangle$ and so is not a dynamical quantity; all the dynamics have been carried in the active picture operators. Going forward, when viewing the density matrix this way, we will call that static density matrix $\hat{\rho}_A$. However, when we reorganize to the passive picture, the linear operators representing properties of our system are stripped of their time dependence, and we repackage that time dependence into the density matrix. We call the passive picture density matrix simply $\hat{\rho}(t)$. Looking at the operator summary and using the cyclicity of the trace:

$$\langle\hat{F}_A(t)\rangle = \mathrm{Tr}(\hat{F}_A(t)\hat{\rho}_A) = \mathrm{Tr}(\hat{U}^\dagger(t,t_0)\hat{F}\hat{U}(t,t_0)\hat{\rho}_A) = \mathrm{Tr}(\hat{F}\hat{U}(t,t_0)\hat{\rho}_A\hat{U}^\dagger(t,t_0)). \tag{2.1.61}$$

Thus, if we define the passive picture density matrix $\hat{\rho}(t)$ by:

$$\hat{\rho}(t) = \hat{U}(t,t_0)\hat{\rho}_A\hat{U}^\dagger(t,t_0) \implies \mathrm{Tr}(\hat{F}_A(t)\hat{\rho}_A) = \mathrm{Tr}(\hat{F}\hat{\rho}) \tag{2.1.62}$$

and so we can equivalently capture all of the dynamical information of the theory by time evolving the density matrix while keeping all of the generator operators $\hat{F}$ constant. Thus, in the passive picture, **the density matrix contains the entire state of the system**. By construction, the operator summaries in the active picture using $\hat{\rho}_A$ exactly match the operator summaries in the passive picture using $\hat{\rho}$ and so nothing is lost or gained in either approach.

We can find a nice equation of motion for the density matrix by differentiating (2.1.62) in time:

$$\frac{d\hat{\rho}}{dt} = \frac{d\hat{U}}{dt}\hat{\rho}_0\hat{U}^\dagger + \hat{U}\hat{\rho}_0\frac{d\hat{U}^\dagger}{dt} = \frac{1}{i\hbar}\hat{H}\hat{U}\hat{\rho}_0\hat{U}^\dagger - \frac{1}{i\hbar}\hat{U}\hat{\rho}_0\hat{U}^\dagger\hat{H} \implies \boxed{\frac{d\hat{\rho}}{dt} + \frac{1}{i\hbar}[\hat{\rho},\hat{H}] = 0.} \tag{2.1.63}$$

Equation (2.1.63) is known as the **von Neumann equation**. It is the fundamental equation of motion of the density matrix, and thus of the entire state of our system. The von Neumann equation perfectly reproduces Ehrenfest's theorem; to see this consider a generator $\hat{F}(t)$ which in the passive picture we allow for generality to possibly have some explicit time dependence:

$$\frac{d\langle\hat{F}\rangle}{dt} = \frac{d}{dt}\mathrm{Tr}(\hat{\rho}\hat{F}) = \mathrm{Tr}\left(\frac{d\hat{\rho}}{dt}\hat{F} + \hat{\rho}\frac{\partial\hat{F}}{\partial t}\right) = \mathrm{Tr}\left(\frac{1}{i\hbar}(\hat{H}\hat{\rho} - \hat{\rho}\hat{H})\hat{F}\right) + \langle\frac{\partial\hat{F}}{\partial t}\rangle = \langle\left(\frac{\partial\hat{F}}{\partial t} + \frac{1}{i\hbar}[\hat{F},\hat{H}]\right)\rangle \tag{2.1.64}$$

using the cyclicity of the trace.

### 2.1.2 Why Density Matrices?

All of this density matrix formalism is mathematically interesting, but how is it quantum mechanics? Previously when we studied quantum mechanics, the state of the system was represented by a state vector

$|\Psi\rangle$. The density matrix can be seen as a strict generalization of the state vector description. Consider the special case that there exists at some particular time $t_0$ a state vector $|\Psi_0\rangle$ so that $\hat{\rho}$ takes the form:

$$\hat{\rho}(t_0) = |\Psi_0\rangle\langle\Psi_0|. \tag{2.1.65}$$

Then, at a general time $t$ the density matrix can be written as:

$$\hat{\rho}(t) = \hat{U}(t,t_0)\hat{\rho}(t_0)\hat{U}^\dagger(t,t_0) = (\hat{U}(t,t_0)|\Psi_0\rangle)(\hat{U}(t,t_0)|\Psi_0\rangle)^\dagger = |\Psi(t)\rangle\langle\Psi(t)| \tag{2.1.66}$$

where we have defined the time dependent state vector $|\Psi(t)\rangle$ by:

$$|\Psi(t)\rangle = \hat{U}(t,t_0)|\Psi_0\rangle \tag{2.1.67}$$

which then automatically satisfies the Schrödinger equation by differentiation:

$$i\hbar\frac{d|\Psi\rangle}{dt} = \hat{H}|\Psi\rangle. \tag{2.1.68}$$

Thus the state vector description of quantum mechanics and the Schrödinger equation emerge as the special case that the density matrix ever takes the specific form of (2.1.65). Such a density matrix is called a **pure state**. Otherwise, it is called a **mixed state**. There is a simple basis independent way to tell if a density matrix is a mixed state or a pure state. For a pure state:

$$1 = \mathrm{Tr}(\hat{\rho}) = \sum_n \langle n|\Psi\rangle\langle\Psi|n\rangle = \langle\Psi|\Psi\rangle \tag{2.1.69}$$

and thus $|\Psi\rangle$ is normalized. Then:

$$\mathrm{Tr}(\hat{\rho}^2) = \sum_n \langle n|\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi|n\rangle = \langle\Psi|\Psi\rangle = 1. \tag{2.1.70}$$

Hence, for a pure state $\mathrm{Tr}(\hat{\rho}^2) = 1$. Alternatively, for a mixed state, we can still diagonalize $\hat{\rho}$:

$$\hat{\rho} = \sum_n \rho_n |\rho;n\rangle\langle\rho;n|. \tag{2.1.71}$$

Because $\mathrm{Tr}(\rho) = 1$, the $\rho_n$ sum to 1:

$$1 = \mathrm{Tr}(\rho) = \sum_n \rho_n \langle\rho;n|\rho;n\rangle = \sum_n \rho_n. \tag{2.1.72}$$

And, because $\hat{\rho}$ is nonnegative, each of its eigenvalues is nonnegative. Thus, each eigenvalue must be less than or equal to 1 and if any of the eigenvalues is 1, all of the rest must be 0. If one of the $\rho_n$ is 1 and the rest are 0 then $\hat{\rho}$ is a pure state. Consequently, for a mixed state:

$$0 < \rho_n < 1. \tag{2.1.73}$$

Multiplying by $\rho_n$:

$$0 < \rho_n^2 < \rho_n \tag{2.1.74}$$

and so for a mixed state:

$$\mathrm{Tr}(\hat{\rho}^2) = \sum_{nml} \langle n|\rho_m|\rho;m\rangle\langle\rho;m|\rho_l|\rho;l\rangle\langle\rho;l|n\rangle = \sum_{ml} \delta_{ml}\rho_m\rho_l\langle\rho;l|\rho;m\rangle = \sum_n \rho_n^2 < \sum_n \rho_n = \mathrm{Tr}(\hat{\rho}) = 1. \tag{2.1.75}$$

Therefore, for all states $\mathrm{Tr}(\hat{\rho}^2) \leq 1$ with equality if and only if $\hat{\rho}$ is a pure state.

The density matrix description is more general than the state vector description and so it doesn't hurt to take interest in it, but more fundamentally there are several reasons it is actually necessary for doing quantum mechanics in general. In a real experiment, many trials have to be done and statistics have to be collected. Consequently, the "same" state has to be attempted to be prepared many times (at least same within some experimental equivalence class). As well, when state preparation is attempted, we do not actually know everything about the state of the system we are preparing due to the enormity of variables in the world beyond our control. Thus, in any real experimental setup there will be some probability distribution on the state of the system at the time of setup. This is just a repeat of the logic used for classical statistical mechanics, and really is the gateway into quantum statistical mechanics. Suppose we believed in a description of quantum mechanics using only state vectors and suppose we had a probability distribution in our experiment so that we prepare state $|\Psi_n\rangle$ with probability $p_n$ (these $|\Psi_n\rangle$ are normalized but we assume nothing about how they relate to each other). Then, let $G = \langle\Psi|\hat{G}|\Psi\rangle$ be some generator of experimental interest. The expected value

of this quantity in our experiment is:

$$\mathbb{E}[G] = \sum_n p_n G_n = \sum_n p_n \langle \Psi_n | \hat{G} | \Psi_n \rangle. \tag{2.1.76}$$

It would be mathematically convenient if we could identify a state vector $|\widetilde{\Psi}\rangle$ so that $\mathbb{E}[G] = \langle \widetilde{\Psi} | \hat{G} | \widetilde{\Psi} \rangle$ with $|\widetilde{\Psi}\rangle$ independent of $\hat{G}$ as then this $|\widetilde{\Psi}\rangle$ would contain all of the information of statistical expectations in our experiment. However, this is almost never possible. Notice though that we can write:

$$\mathbb{E}[G] = \mathrm{Tr}\left( \hat{G} \sum_n p_n |\Psi_n\rangle\langle\Psi_n| \right) \tag{2.1.77}$$

and thus can define the effective density matrix $\hat{\rho}$:

$$\hat{\rho} = \sum_n p_n |\Psi_n\rangle\langle\Psi_n| \tag{2.1.78}$$

so that:

$$\mathbb{E}[G] = \mathrm{Tr}(\hat{G}\hat{\rho}). \tag{2.1.79}$$

Quite short calculations reveal that $\hat{\rho}$ as defined here is self-adjoint, has trace 1, is nonnegative, and satisfies the von Neumann equation. Thus, whenever the quantum state is imperfectly known or prepared over a distribution of state vectors, there is generally no effective state vector that describes the system statistically. However, there is always a density matrix which does so.

The density matrix structure is also self-reproducing. If we consider a distribution of probabilities $p_n$ over possible density matrices $\hat{\rho}_n$ then we can still assign an effective density matrix to the statistics:

$$\mathbb{E}[G] = \sum_n p_n G_n = \sum_n p_n \mathrm{Tr}(\rho_n \hat{G}) = \mathrm{Tr}\left( \hat{G} \sum_n p_n \hat{\rho}_n \right) \overset{!}{=} \mathrm{Tr}(\hat{G}\hat{\rho}) \ \forall \hat{G} \implies \hat{\rho} = \sum_n p_n \hat{\rho}_n. \tag{2.1.80}$$

Again, short calculations reveal that $\hat{\rho}$ as defined here is self-adjoint, has trace 1, is nonnegative, and satisfies the von Neumann equation. Thus when combining quantum mechanics with statistical mechanics a density is absolutely necessary. Because probabilities themselves can be cast in the form of statistical expectations, we know that in a statistical context the expectation map $\mathbb{E}$ contains all statistical information. Evidently, this $\hat{\rho}$ is the statistical expectation of the system's single realization density matrix. Because the expectation of operator summaries is characterized by a density matrix, in quantum statistical mechanics, the state of the system remains a density matrix. Interestingly, whether considering a single quantum system or a distribution over quantum systems, the state is just a density matrix. (Technically, this is only true when the only quantities of interest are the expectations of operator summaries. The distribution $p_n$ and constituent density matrices $\hat{\rho}_n$ contain more information than $\hat{\rho}$ itself when considering quantities beyond the expectations of operator summaries. Expectations of operator summaries happen to be the experimentally relevant quantities of interest in quantum statistical mechanics, allowing $\hat{\rho}$ to be sufficiently descriptive for it to be the only thing we care about.)

It is extremely important to note here a subtle point of interpretation which is often gotten wrong. We see that when one has a probability distribution on the density matrix one can identify an effective density matrix for the statistical expectations of that distribution. That expected density matrix is the probability weighted sum of density matrices in the distribution where the weights are nonnegative and sum to 1. Such a sum is called a convex linear combination. Any given density matrix that is not a pure state can be written infinitely many ways as a convex sum. So, if one has a density matrix $\hat{\rho}$ that can be written in the form $\sum_n p_n \hat{\rho}_n$ where the $p_n$ as numbers form a mathematically valid probability distribution and the $\hat{\rho}_n$ are each a mathematically valid density matrix, that does **NOT** mean that $p_n$ is the actual probability the density matrix of a single instance of the system is $\hat{\rho}_n$. For example, on a two dimensional Hilbert space the density matrix $\hat{\rho} = \frac{1}{2}\mathbb{1}$ can be written as $\hat{\rho} = \frac{1}{2}(|+_z\rangle\langle+_z| + |-_z\rangle\langle-_z|)$ where $|+_z\rangle = \binom{1}{0}$ and $|-_z\rangle = \binom{0}{1}$. Naively interpreting in the way I am telling you not to would suggest that this means $\hat{\rho}$ comes from a source which in half of experiments is in the state $|+_z\rangle\langle+_z|$ and in the other half is in the state $|-_z\rangle\langle-_z|$. However, it is also possible to decompose $\hat{\rho}$ as $\hat{\rho} = \frac{1}{2}(|+_x\rangle\langle+_x| + |-_x\rangle\langle-_x|$ where $|+_x\rangle = \frac{|+_z\rangle + |-_z\rangle}{\sqrt{2}}$ and $|-_x\rangle = \frac{|+_z\rangle - |-_z\rangle}{2}$. Naively interpreting this decomposition instead would suggest that $\hat{\rho}$ comes from a source which in half of experiments is in the state $|+_x\rangle\langle+_x|$ and in the other half is in the state $|-_x\rangle\langle-_x|$. This interpretation is clearly inconsistent with the previous interpretation because none of these states match each other. So, absolutely plainly you can **NOT** reverse engineer the distribution $p_n$ and constituent density matrices $\hat{\rho}_n$ that

go into forming the average density matrix $\hat{\rho}$ that describes statistical expectations of operator summaries from the final average density matrix $\hat{\rho}$ alone.

It is a common misconception that density matrices are only useful or necessary for describing statistical ensembles of quantum systems. In truth, the density matrix is also necessary outside of quantum statistical mechanics. Suppose we have one quantum system which has operators acting on a Hilbert space $\mathbb{H}_1$ and another quantum system which has operators acting on another Hilbert space $\mathbb{H}_2$. Then, what is the appropriate Hilbert space of the combined interacting system? When we realized quantum field theory was necessary, we learned that the two particle Hilbert space is given by the tensor product of each of the individual particle Hilbert spaces (in the case that the particles were the same species, like photons, we also saw that indistinguishability reflected in the Hilbert space, in that case by symmetrization). This is because in general, when two quantum systems are combined, the combined Hilbert space is $\mathbb{H}_1 \otimes \mathbb{H}_2$. This is called a bipartite quantum system. This tensor product structure is automatically necessary when the ability to select two properties independently is combined with the linearity structure of quantum mechanics. Suppose the two parts of our system are allowed to interact. Then, the combined system will have some state on the space $\mathbb{H} = \mathbb{H}_1 \otimes \mathbb{H}_2$. If we were to commit to a state vector description, then letting $|n_1\rangle$ be an orthonormal basis for $\mathbb{H}_1$ and $|n_2\rangle$ be an orthonormal basis for $\mathbb{H}_2$, the state vector $|\Psi\rangle$ on $\mathbb{H}$ would take the form:

$$|\Psi\rangle = \sum_{mn} \Psi_{mn} |m_1\rangle \otimes |n_2\rangle. \tag{2.1.81}$$

Now consider the value of a quantity associated with the first system. Let $\hat{A}$ be a linear operator associated to some generator on $\mathbb{H}_1$. On $\mathbb{H}$ there is then some operator $\hat{F}$ to represent this quantity. Because this quantity is intended to only be a property of system 1, its value should be independent of system 2. As well, $\hat{F}$ should reduce to just $\hat{A}$ on system 1. So, $\hat{F}$ must take the form:

$$\hat{F} = \hat{A} \otimes \mathbb{1} \tag{2.1.82}$$

Then, the summary of $\hat{F}$ is given by:

$$\langle \Psi | \hat{F} | \Psi \rangle = \sum_{mn} \Psi^*_{mn} \langle m_1 n_2 | \hat{A} \otimes \mathbb{1} \sum_{m'n'} \Psi_{m'n'} | m'_1 n'_2 \rangle = \sum_{mnm'n'} \Psi^*_{mn} \Psi_{m'n'} \langle m_1 | \hat{A} | m'_1 \rangle \langle n_2 | \mathbb{1} | n'_2 \rangle \tag{2.1.83}$$

$$= \sum_{mnm'n'} \Psi^*_{mn} \Psi_{m'n'} A_{mm'} \delta_{nn'} = \sum_{mm'} A_{mm'} \sum_n \Psi^*_{mn} \Psi_{m'n}. \tag{2.1.84}$$

If we define the operator $\hat{\rho}_1$ on $\mathbb{H}_1$ by:

$$\langle m'_1 | \hat{\rho}_1 | m_1 \rangle = \sum_n \Psi_{m'n} \Psi^*_{mn} \tag{2.1.85}$$

$$\implies \langle \Psi | \hat{F} | \Psi \rangle = \sum_{mm'} A_{mm'} \langle m'_1 | \hat{\rho}_1 | m_1 \rangle = \sum_{mm'} \langle m_1 | \hat{A} | m'_1 \rangle \langle m'_1 | \hat{\rho}_1 | m_1 \rangle = \mathrm{Tr}(\hat{\rho}_1 \hat{A}). \tag{2.1.86}$$

There is in general no $\Psi_1$ so that $\langle \Psi | \hat{F} | \Psi \rangle$ is given by $\langle \Psi_1 | \hat{A} | \Psi_1 \rangle$ for all $\hat{A}$ and corresponding $\hat{F}$. But, there is a density matrix $\hat{\rho}_1$ so that the properties of system 1 are described by that density matrix no matter the state of the combined system (again the basic properties of trace 1, self-adjointness, and positivity can be checked directly). Even if a combined quantum system's full state is described by a state vector, there is no guarantee that any of its subsystems can be described by a state vector. In general, a subsystem of a quantum system *must* be described by a density matrix. This argument has nothing to do with the dynamics of a quantum subsystem, which we will discuss later. The density matrix is necessitated simply by considering a subsystem. However, every system in the universe other than the whole universe itself is a subsystem. Therefore, every system ever relevant experimentally demands a density matrix if treated carefully. It is essentially an occasional accident that the state vector description ever gives correct results; to consistently be correct a density matrix is mandatory at high precision.

In this subsystem description, the density matrix structure again reproduces itself. Consider a system with state density matrix $\hat{\rho}$ and Hilbert space $\mathbb{H}_1 \otimes \mathbb{H}_2$. Then, consider the same $\hat{F}$ and $\hat{A}$ structure as previously. The operator summary of $\hat{F}$ gives:

$$\langle \hat{F} \rangle = \mathrm{Tr}(\hat{F}\hat{\rho}) = \sum_{mn} \langle m_1 n_2 | (\hat{A} \otimes \hat{1}) \hat{\rho} | m_1 n_2 \rangle = \sum_{mm'n} \langle m_1 | \hat{A} | m'_1 \rangle \langle m'_1 n_2 | \hat{\rho} | m_1 n_2 \rangle. \tag{2.1.87}$$

If we define the operator $\hat{\rho}_1$ on $\mathbb{H}_1$ by:

$$\langle m'_1 | \hat{\rho}_1 | m_1 \rangle = \sum_n \langle m'_1 n_2 | \hat{\rho} | m_1 n_2 \rangle \tag{2.1.88}$$

$$\implies \langle \hat{F} \rangle = \sum_{mm'} \langle m_1 | \hat{A} | m_1' \rangle \langle m_1' | \hat{\rho}_1 | m_1 \rangle = \mathrm{Tr}(\hat{A}\rho_1). \tag{2.1.89}$$

Thus there remains an effective density matrix for the subsystem. By using the density matrix description, every subsystem of a quantum system has a well-defined state. We can see that the structure in (2.1.88) keeps the system 1 labels ($m_1$ and $m_1'$) fixed while summing over a basis of system 2 (the $n_2$ labels). This is called a partial trace. In particular, $\hat{\rho}_1$ is the partial trace of $\hat{\rho}$ when traced over system 2 and is called the reduced density matrix of system 1. This is written as (and defined properly by):

$$\hat{\rho}_1 = \mathrm{Tr}_2(\hat{\rho}) = \sum_{mm'n} |m_1\rangle \langle m_1 n_2 | \hat{\rho} | m_1' n_2 \rangle \langle m_1' |. \tag{2.1.90}$$

### 2.1.3 Entanglement

When the state of a quantum system is a state vector in a bipartite Hilbert space $\mathbb{H} = \mathbb{H}_1 \otimes \mathbb{H}_2$ yet it is not possible to describe the effective state on either $\mathbb{H}_1$ or $\mathbb{H}_2$ by a state vector, the two subsystems are said to be entangled (otherwise, they are separable). It is relatively simple to determine whether a state vector of a bipartite quantum system is entangled or separable. Consider a state vector of a bipartite quantum system:

$$|\Psi\rangle = \sum_{mn} \Psi_{mn} |m_1\rangle \otimes |n_2\rangle. \tag{2.1.91}$$

By definition if and only if the state is separable, the density matrix $\hat{\rho}_1$ is a pure state and we can assign a state vector $|\Psi_1\rangle$ to system 1. That is:

$$\hat{\rho}_1 = |\Psi_1\rangle\langle\Psi_1|, \qquad |\Psi_1\rangle = \sum_m \Psi_m^{(1)} |m_1\rangle \tag{2.1.92}$$

for some coefficients $\Psi_m^{(1)}$. Using (2.1.85), we have automatically that for a separable state:

$$\sum_n \Psi_{mn} \Psi_{m'n}^* = \Psi_m^{(1)} \Psi_{m'}^{(1)*}. \tag{2.1.93}$$

This is equation defines the $\Psi_m^{(1)}$ coefficients, if they exist, and as always for a wavefunction we can redefine all of the coefficients by multiplying by a single overall phase with no effect. For $m' = m$:

$$|\Psi_m^{(1)}| = \sqrt{\sum_n |\Psi_{mn}|^2}. \tag{2.1.94}$$

Let $m_0$ be a value of $m$ for which $|\Psi_m^{(1)}| \neq 0$. Such an $m$ must exist because $\hat{\rho}_1 \neq 0$. Choose the arbitrary overall phase so that $\Psi_{m_0}^{(1)} = |\Psi_{m_0}^{(1)}|$. Thus, using (2.1.93) with $m' = m_0$:

$$\Psi_m^{(1)} = \frac{\sum_n \Psi_{mn} \Psi_{m_0 n}^*}{\sqrt{\sum_n |\Psi_{m_0 n}|^2}}. \tag{2.1.95}$$

Returning this expression to (2.1.93), renaming $m_0$ to $m''$, and simplifying gives:

$$\left( \sum_n \Psi_{mn} \Psi_{m'n}^* \right) \left( \sum_n |\Psi_{m''n}|^2 \right) = \left( \sum_n \Psi_{mn} \Psi_{m''n}^* \right) \left( \sum_n \Psi_{m''n} \Psi_{m'n}^* \right). \tag{2.1.96}$$

While originally derived for $m''$ such that $\sum_n |\Psi_{m''n}|^2 \neq 0$, it is automatically also true if this condition is not met as then both sides are automatically 0. Thus, for and only for a separable state, this equation holds for all $m, m', m''$. Whether a state is separable can then be determined by computing and comparing the two sides of this equation. For a separable state, evaluate the above at $m' = m$ and then rename $m''$ to $m'$:

$$\left( \sum_n |\Psi_{mn}|^2 \right) \left( \sum_n |\Psi_{m'n}|^2 \right) = \left| \sum_n \Psi_{mn} \Psi_{m'n}^* \right|^2. \tag{2.1.97}$$

Define the vectors $|\Phi_m\rangle \in \mathbb{H}_2$ by:

$$|\Phi_m\rangle = \sum_n \Psi_{mn} |n_2\rangle. \tag{2.1.98}$$

Then (2.1.97) can be written simply as:

$$\langle \Phi_m | \Phi_m \rangle \langle \Phi_{m'} | \Phi_{m'} \rangle = |\langle \Phi_m | \Phi_{m'} \rangle|^2. \tag{2.1.99}$$

This is the saturated case of the Cauchy-Schwarz inequality for each $m$ and $m'$. Thus, all of the $|\Phi_m\rangle$ must be colinear and we can define a vector $|\Phi\rangle$ (which we take to be normalized by definition) and coefficients $A_m$ so that:

$$|\Phi_m\rangle = A_m|\Phi\rangle, \qquad \langle\Phi|\Phi\rangle = 1. \tag{2.1.100}$$

Let $|\Phi\rangle$ have coefficients $B_n$:

$$|\Phi\rangle = \sum_n B_n|n_2\rangle. \tag{2.1.101}$$

Then, the definition of $|\Phi_m\rangle$ yields:

$$\Psi_{mn} = A_m B_n. \tag{2.1.102}$$

Automatically, by (2.1.85):

$$\Psi^{(1)}_{m'}\Psi^{(1)*}_m = A_{m'}A^*_m. \tag{2.1.103}$$

Because this is the only restriction on $\Psi^{(1)}_m$, we can take $A_m = \Psi^{(1)}_m$. Then, it automatically follows similarly that $\hat{\rho}_2$ is a pure state with state vector $|\Phi\rangle$ and so we can take $B_n = \Psi^{(2)}_n$. Thus, a separable state vector $|\Psi\rangle$ is simply a tensor product of the form:

$$|\Psi\rangle = |\Psi_1\rangle \otimes |\Psi_2\rangle \tag{2.1.104}$$

for some $|\Psi_1\rangle$, $|\Psi_2\rangle$. The converse that a state of such form is separable follows by a one line calculation.

For a two component quantum system with a pure state density matrix $\hat{\rho} = |\Psi\rangle\langle\Psi|$, entanglement represents the extent to which the density matrices of the components of the system fail to contain all of the information of the density matrix of the full system. Let $\hat{F} = \hat{A} \otimes \hat{B}$ where $\hat{A}$ is an operator on $\mathbb{H}_1$ and $\mathbb{H}_2$ is an operator on $\mathbb{H}_2$. For $|\Psi\rangle$ separable, we can write $|\Psi\rangle = |\Psi_1\rangle \otimes |\Psi_2\rangle$ and so:

$$\langle\hat{F}\rangle = \langle\Psi|\hat{F}|\Psi\rangle = \langle\Psi_1|\hat{A}|\Psi_1\rangle\langle\Psi_2|\hat{B}|\Psi_2\rangle = \langle\hat{A}\rangle_1\langle\hat{B}\rangle_2. \tag{2.1.105}$$

Alternatively, if $|\Psi\rangle$ is entangled, then there is a some minimum number $K > 1$ of terms, normalized wavefunctions $|\Psi^{(k)}_1\rangle$, ..., $|\Psi^{(k)}_2\rangle$, and coefficients $c_k$ so that:

$$|\Psi\rangle = \sum_{k=1}^{K} c_k|\Psi^{(k)}_1\rangle \otimes |\Psi^{(k)}_2\rangle \tag{2.1.106}$$

$$\implies \langle\hat{F}\rangle = \sum_{k=1}^{K}\sum_{k'=1}^{K} c_k c^*_{k'}\langle\Psi^{(k')}_1|\hat{A}|\Psi^{(k)}_1\rangle\langle\Psi^{(k')}_2|\hat{B}|\Psi^{(k)}_2\rangle. \tag{2.1.107}$$

Thus for an entangled system it is not possible to reduce the summary of a joint operator to the subsystem summaries of terms in the joint operator. In this way, the density matrix $\hat{\rho}$ contains information about the joint properties of the two subsystem not contained in either of the subsystem density matrices $\hat{\rho}_1$, $\hat{\rho}_2$.

Entanglement is also defined for a mixed state density matrix by generalizing the above concept. We understood previously that if there is a classical probability distribution on the states of a system with density matrix $\hat{\rho}^{(n)}$ being appropriate with probability $p_n$, then the effective density matrix of the ensemble is:

$$\hat{\rho} = \sum_n p_n\hat{\rho}^{(n)}. \tag{2.1.108}$$

Thus, a linear combination of density matrices with positive coefficients that sum to 1 is called a "classical mixture", meaning equivalent in terms of summaries to a classical probability distribution on the state of the system. Any density matrix can be written as a classical mixture of pure states, because the diagonal form of $\hat{\rho}$ is automatically so. The density matrix of a two component quantum system is said to be separable if it is possible to write that density matrix as a convex linear combination of pure states which are themselves separable. Otherwise, that density matrix is called entangled. Unfortunately, it is a very difficult computational problem to determine whether a general density matrix on a large Hilbert space is entangled. There is no known simple criterion to check.

There is an analog to the factorization property (2.1.105) possible for density matrices, but it is stronger than separability. The subsystems of a bipartite quantum system are said to be *totally uncorrelated* if for all operators of the form $\hat{A} \otimes \hat{B}$:

$$\langle\hat{A} \otimes \hat{B}\rangle = \langle\hat{A}\rangle_1\langle\hat{B}\rangle_2. \tag{2.1.109}$$

Automatically it follows that by considering $\hat{A} = |n_1\rangle\langle m_1|, \hat{B} = |n_2'\rangle\langle m_2'|$:

$$\langle m_1 m_2'|\hat{\rho}|n_1 n_2'\rangle = \langle m_1|\hat{\rho}_1|n_1\rangle\langle m_2'|\hat{\rho}_2|n_2'\rangle \implies \hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2. \tag{2.1.110}$$

The converse follows by an equally short calculation. Thus the subsystems are totally uncorrelated if and only if the combined system density matrix factorizes as $\hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2$.

Much like interactions between subsystems create correlations in classical statistical mechanics, interactions between subsystems create correlations and entanglement in quantum mechanics. This is because the von Neumann equation only applies to an isolated whole system, not to an interacting subsystem. Consider a bipartite quantum system. Let $\hat{H}_1$ be the Hamiltonian operator of the first subsystem when it is free and $\hat{H}_2$ be the Hamiltonian operator of the second subsystem when it is free. Then, let $\hat{H}_{\text{int}}$ be the interaction term in the Hamiltonian operator between the two subsystems. So, the full Hamiltonian operator $\hat{H}$ of the bipartite system is:

$$\hat{H} = \hat{H}_1 \otimes \mathbb{1} + \mathbb{1} \otimes \hat{H}_2 + \hat{H}_{\text{int}}. \tag{2.1.111}$$

The full density matrix $\hat{\rho}$ evolves according to the von Neumann equation. Let us see how $\hat{\rho}_1$, the density matrix of the first subsystem, evolves. By definition:

$$\frac{d\hat{\rho}_1}{dt} = \frac{d}{dt}\text{Tr}_2(\hat{\rho}) = \sum_n \langle n_2|\frac{d\hat{\rho}}{dt}|n_2\rangle = \sum_n \frac{1}{i\hbar}\langle n_2|(\hat{H}\hat{\rho} - \hat{\rho}\hat{H})|n_2\rangle \tag{2.1.112}$$

$$= \frac{1}{i\hbar}\sum_n \langle n_2|((\hat{H}_1 \otimes \mathbb{1})\hat{\rho} + (\mathbb{1} \otimes \hat{H}_2)\hat{\rho} + \hat{H}_{\text{int}}\hat{\rho} - \hat{\rho}(\hat{H}_1 \otimes \mathbb{1}) - \hat{\rho}(\mathbb{1} \otimes \hat{H}_2) - \hat{\rho}\hat{H}_{\text{int}})|n_2\rangle. \tag{2.1.113}$$

By inserting complete sets of states, these terms can be simplified individually to give:

$$\sum_n \langle n_2|(\hat{H}_1 \otimes \mathbb{1})\hat{\rho}|n_2\rangle = \hat{H}_1\hat{\rho}_1 \tag{2.1.114}$$

$$\sum_n \langle n_2|\hat{\rho}(\hat{H}_1 \otimes \mathbb{1})|n_2\rangle = \hat{\rho}_1\hat{H}_1 \tag{2.1.115}$$

$$\sum_n \langle n_2|(\mathbb{1} \otimes \hat{H}_2)\hat{\rho}|n_2\rangle = \sum_{mm'nn'} |m_1\rangle\langle n_2|\hat{H}_2|n_2'\rangle\langle m_1 n_2'|\hat{\rho}|m_1' n_2\rangle\langle m_1'| \tag{2.1.116}$$

$$\sum_n \langle n_2|\hat{\rho}(\mathbb{1} \otimes \hat{H}_2)|n_2\rangle = \sum_n \langle n_2|(\mathbb{1} \otimes \hat{H}_2)\hat{\rho}|n_2\rangle. \tag{2.1.117}$$

Thus, after simplification:

$$\frac{d\hat{\rho}_1}{dt} + \frac{1}{i\hbar}[\hat{\rho}_1, \hat{H}_1] + \frac{1}{i\hbar}\text{Tr}_2([\hat{\rho}, \hat{H}_{\text{int}}]) = 0. \tag{2.1.118}$$

So, the density matrix $\hat{\rho}_1$ of the subsystem does not have an autonomous equation of motion. It is not possible in the case of interaction to determine the exact time evolution of $\hat{\rho}_1$ without knowing the full state $\hat{\rho}$ of the combined system (it is sometimes possible to come up with some useful approximations under special conditions). The forcing operator $\frac{1}{i\hbar}\text{Tr}_2([\hat{\rho}, \hat{H}_{\text{int}}])$ is essentially an arbitrary self-adjoint operator that preserves the trace and nonnegativity of $\hat{\rho}_1$. Because of this, it is not in general possible to find an operator $\hat{V}(t, t_0)$ so that $\hat{\rho}_1(t) = \hat{V}(t, t_0)\hat{\rho}(t_0)\hat{V}^\dagger(t, t_0)$ and consequently, even if $\hat{\rho}(t_0)$ is a pure state it is not in general possible to identify a time dependent state vector so that $\hat{\rho}_1(t)$ can be written as a pure state in terms of that vector. Thus, the presence of interactions causes pure states to become mixed states. From the subsystem perspective, this means that interactions cause the subsystems to become entangled. A note to be careful: partial traces only have the cyclicity property on operators entirely in the Hilbert space being traced over. By that I mean, although $\text{Tr}(\hat{\rho}\hat{H}_{\text{int}}) = \text{Tr}(\hat{H}_{\text{int}}\hat{\rho})$ is true, it is also true that $\text{Tr}_2(\hat{\rho}\hat{H}_{\text{int}}) \neq \text{Tr}_2(\hat{H}_{\text{int}}\hat{\rho})$, which can be confirmed by a short calculation putting these expressions in explicit bases.

### 2.1.4 Canonical Structure of Density Matrices

In section 1.4 we found it is possible to reconstruct the state vector up to an overall phase from operator summaries, thus making it observable up to that phase. Similarly and more generally, the entire density matrix can unambiguously be reconstructed from operator summaries. Thus quantum tomography applies to quantum mechanics in general. Consider a quantum system on a Hilbert space with orthonormal basis $|n\rangle$ and define the operators $\hat{S}_{mn}$ and $\hat{A}_{mn}$ by:

$$\hat{S}_{mn} = \frac{|m\rangle\langle n| + |n\rangle\langle m|}{2}, \qquad \hat{A}_{mn} = \frac{|m\rangle\langle n| - |n\rangle\langle m|}{2i}. \tag{2.1.119}$$

From these, $\hat{\rho}$ is explicitly given by:

$$\langle \hat{S}_{mn} \rangle + i\langle \hat{A}_{mn} \rangle = \text{Tr}(|m\rangle\langle n|\hat{\rho}) = \langle n|\hat{\rho}|m\rangle \implies \hat{\rho} = \sum_{mn} (\langle \hat{S}_{mn} \rangle + i\langle \hat{A}_{mn} \rangle)|n\rangle\langle m|. \tag{2.1.120}$$

The density matrix approach makes it extremely clear why the overall phase of the wavefunction/state vector is meaningless. The true state of a quantum system is a density matrix, and in the case that density matrix is a pure state it can be written in the form $\hat{\rho} = |\Psi\rangle\langle\Psi|$ thus allowing one to identify $|\Psi\rangle$ as the effective state. However, there is no unique way to decompose a pure state in this way as if and only if $|\Psi'\rangle = e^{i\varphi}|\Psi\rangle$ for some phase $\varphi$ then $\hat{\rho} = |\Psi\rangle\langle\Psi| = |\Psi'\rangle\langle\Psi'|$. **The density matrix itself is entirely physical, uniquely defined, and determinable by quantum tomography.** Of course, practically one still must complete many many measurements for this and can typically only determine the full density matrix in an asymptotic sense through the consequent enormity of data.

In this section we reorganized our entire operator formalism into the density matrix approach and found that it naturally reproduces the mathematical structure of quantum mechanics while remaining faithful as a representation to the structures introduced in Orig. section 1.3, unlike the Poisson representation. This new representation is called the Schrödinger representation. However, it is possible to describe all of quantum mechanics entirely within the Poisson representation and a classical Hamiltonian, as long as it is formulated on the correct phase space. This should not be at all surprising to you as we established in Orig. chapter 1 that any system of differential equations can be described by a Hamiltonian on a sufficiently large phase space. In the state vector description of quantum mechanics we know from (1.2.27), that the Hamiltonian is:

$$\mathcal{H} = \langle \hat{H} \rangle = \langle \Psi | \hat{H} | \Psi \rangle \tag{2.1.121}$$

where in an arbitrary orthonormal basis $|n\rangle$ the components $\Psi_n = \langle n|\Psi\rangle$ have the Poisson brackets:

$$\{\Psi_n, \Psi_m\} = 0, \qquad \{\Psi_n, \Psi_m^*\} = \frac{1}{i\hbar}\delta_{nm}. \tag{2.1.122}$$

We can rewrite the Hamiltonian in terms of the density matrix $\hat{\rho} = |\Psi\rangle\langle\Psi|$ as:

$$\mathcal{H} = \langle \hat{H} \rangle = \text{Tr}(\hat{\rho}\hat{H}) \tag{2.1.123}$$

and compute the Poisson brackets of the components $\rho_{mn} = \langle m|\hat{\rho}|n\rangle = \Psi_m \Psi_n^*$ using the product rule:

$$\{\rho_{mn}, \rho_{m'n'}\} = \{\Psi_m \Psi_n^*, \Psi_{m'} \Psi_{n'}^*\} = \Psi_m \Psi_{n'}^* \{\Psi_n^*, \Psi_{m'}\} + \Psi_{m'} \Psi_n^* \{\Psi_m, \Psi_{n'}^*\} \tag{2.1.124}$$

$$= -\frac{1}{i\hbar}\rho_{mn'}\delta_{m'n} + \frac{1}{i\hbar}\rho_{m'n}\delta_{mn'}. \tag{2.1.125}$$

From this Poisson bracket, the von Neumann immediately follows from the standard Hamilton's equations:

$$\frac{d\hat{\rho}}{dt} = \{\hat{\rho}, \mathcal{H}\} = \sum_{mn} |m\rangle\langle m|\{\hat{\rho}, \mathcal{H}\}|n\rangle\langle n| = \sum_{mn} \left\{ \rho_{mn}, \sum_{m'n'} \rho_{m'n'} \langle n'|\hat{H}|m'\rangle \right\} \langle n| \tag{2.1.126}$$

$$= \sum_{mm'nn'} |m\rangle\langle n'|\hat{H}|m'\rangle\langle n| \left( -\frac{1}{i\hbar}\rho_{mn'}\delta_{m'n} + \frac{1}{i\hbar}\rho_{m'n}\delta_{mn'} \right) = \frac{1}{i\hbar}\hat{H}\hat{\rho} - \frac{1}{i\hbar}\hat{\rho}\hat{H}. \tag{2.1.127}$$

Thus, even though we derived the form of the Hamiltonian and Poisson brackets for a pure state they give the correct results for a general density matrix.

The density matrix Poisson bracket can be written equivalently in terms of the operator summaries of generic operators. Let $\hat{A}$ and $\hat{B}$ be two operators on the Hilbert space. Then:

$$\{\langle \hat{A} \rangle, \langle \hat{B} \rangle\} = \sum_{mm'nn'} \langle n|\hat{A}|m\rangle\langle n'|\hat{B}|m'\rangle \{\rho_{mn}, \rho_{m'n'}\} \tag{2.1.128}$$

$$= \sum_{mm'nn'} \langle n|\hat{A}|m\rangle\langle n'|\hat{B}|m'\rangle \left( -\frac{1}{i\hbar}\rho_{mn'}\delta_{m'n} + \frac{1}{i\hbar}\rho_{m'n}\delta_{mn'} \right) = \langle \frac{1}{i\hbar}[\hat{A}, \hat{B}] \rangle. \tag{2.1.129}$$

Thus, Dirac's theorem is replicated for density matrices. This Poisson bracket determines the full density matrix Poisson bracket by simply considering $\hat{A} = |m\rangle\langle n|$ and $\hat{B} = |m'\rangle\langle n'|$.

**Summary 2.1: Mathematical Formulation of Quantum Mechanics**

In quantum mechanics, the state of the system is represented by a trace 1, self-adjoint, nonnegative operator $\hat{\rho}(t)$ (on a Hilbert space $\mathbb{H}$) known as the density matrix.

All of the properties of the state of the system are contained in the operator summary map $\langle\cdot\rangle$ which for a generic operator $\hat{A}$ on $\mathbb{H}$ is defined by:

$$\langle\hat{A}\rangle = \mathrm{Tr}(\hat{\rho}\hat{A}). \tag{2.1.130}$$

Real valued properties are the summaries of self-adjoint operators. This includes the Hamiltonian $\mathcal{H} = \langle\hat{H}\rangle$ where the self-adjoint operator $\hat{H}$ is called the Hamiltonian operator.

The Poisson bracket of the density matrix is captured by Dirac's theorem:

$$\{\langle\hat{A}\rangle, \langle\hat{B}\rangle\} = \langle\frac{1}{i\hbar}[\hat{A}, \hat{B}]\rangle \tag{2.1.131}$$

for arbitrary state independent operators $\hat{A}, \hat{B}$.

The resulting equation of motion for the density matrix is the von Neumann equation:

$$\frac{d\hat{\rho}}{dt} = \{\hat{\rho}, \mathcal{H}\} \implies \frac{d\hat{\rho}}{dt} + \frac{1}{i\hbar}[\hat{\rho}, \hat{H}] = 0. \tag{2.1.132}$$

The time evolution of the density matrix is solved by:

$$\hat{\rho}(t) = \hat{U}(t, t_0)\hat{\rho}(t_0)\hat{U}^\dagger(t, t_0) \tag{2.1.133}$$

where the time evolution operator $\hat{U}(t, t_0)$ is given by:

$$\hat{U}(t, t) = \mathbb{1}, \qquad i\hbar\frac{d\hat{U}(t, t_0)}{dt} = \hat{H}(t)\hat{U}(t, t_0) \implies \hat{U}(t, t_0) = \mathscr{T} : e^{-\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t')dt'}. \tag{2.1.134}$$

When two quantum subsystems on Hilbert spaces $\mathbb{H}_1$ and $\mathbb{H}_2$ are brought together to interact, the Hilbert space of the combined system is $\mathbb{H} = \mathbb{H}_1 \otimes \mathbb{H}_2$. The full system is described by a density matrix on $\mathbb{H}$. It is possible to identify an effective density matrix for either subsystem by using the partial trace over the other subsystem:

$$\hat{\rho}_1 = \mathrm{Tr}_2(\hat{\rho}), \qquad \hat{\rho}_2 = \mathrm{Tr}_1(\hat{\rho}). \tag{2.1.135}$$

Quantum mechanics as formulated above is done on an arbitrary Hilbert space. Very often, the Hilbert space of interest for us is of a particular form. Specifically, it is often the Hilbert space of square integrable functions of some real valued coordinates $Q^a$. Because the space of coordinate values is called configuration space, we call such a Hilbert space a **configuration Hilbert space**. There are several nice additional formulas that can be derived for quantum mechanics on a configuration Hilbert space that we will derive throughout this chapter. In our studies of single particle quantum mechanics, we already encountered a configuration Hilbert space as the Hilbert space of square integrable functions of spatial position, and then our hints of quantum field theory so far have also indicated that there is a configuration Hilbert space for the Hilbert space of square integrable gauge-invariant functionals of the electromagnetic vector potential. We have also encountered some non-configuration Hilbert spaces, such as the 2 dimensional Hilbert space of a single spinor. It is always possible to embed any Hilbert space into a sufficiently large configuration Hilbert space, though doing so may be inconvenient.

Importantly, whereas in single particle quantum mechanics the linearity of the Schrödinger equation followed from a systematic approximation of a generic stable Hamiltonian, here in proper fully general quantum mechanics, the linearity of the von Neumann equation is an exact property following from the existence and properties of the summary map. While there are effective nonlinear corrections to the Schrödinger equation for a single particle, we do not expect and have not observed any nonlinear corrections to full generic quantum mechanics.

The appearance (and arbitrariness) of $\hbar$ in the formalism of this section explains why $\hbar$ is universal, i.e. why wavefunctions of different particle species have the same value of $\hbar$. All of the Hilbert spaces of different species of particles exist as factors in a large tensor product forming the overall Hilbert space of the quantum state of the entire universe. The value of $\hbar$ in the overall full state von Neumann equation then gets inherited by every subspace. Thus, if we choose units that set $\hbar = 1$ for one species, it will be 1 everywhere.

## Exercises

**Problem 2.1: Pass without a trace.** An operator is trace class if its trace series in any orthonormal

basis is absolutely convergent. For such an operator, the trace is the same in all bases. Otherwise, it is possible to achieve any value of the trace. Here we will see the contradictions you find from taking traces of non-trace-class operators. Consider the Hilbert space of square integrable functions in a single spatial dimensions with operators $\hat{x}$ and $\hat{p}$ as usual. Beginning with the canonical commutation relation:

$$[\hat{x}, \hat{p}] = i\hbar \mathbb{1}$$

show that if we assume $\hat{x}$ and $\hat{p}$ to be trace class, then the left hand side is 0, and that if we assume $\mathbb{1}$ to be trace class then the right hand side is infinite, a contradiction! Show that the traces of $\hat{x}, \hat{p}, \mathbb{1}$ each fail to absolutely converge in the position and momentum bases, and thus none of these operators are trace class, so we're safe.

**Problem 2.2: Spread Theorems for Density Matrices.** For operators $\hat{A}, \hat{B}$ define the operator correlation in the state $\hat{\rho}$ between $\hat{A}$ and $\hat{B}$ exactly as in state vector quantum mechanics by:

$$\text{Cor}(\hat{A}, \hat{B}) = \frac{1}{2}\left(\langle(\hat{A}^\dagger - \langle\hat{A}^\dagger\rangle\mathbb{1})(\hat{B} - \langle\hat{B}\rangle\mathbb{1})\rangle + \langle(\hat{B} - \langle\hat{B}\rangle\mathbb{1})(\hat{A}^\dagger - \langle\hat{A}^\dagger\rangle\mathbb{1})\rangle\right).$$

as well define the operator spread of $\hat{\mathcal{O}}$ again by:

$$\text{Spr}(\hat{\mathcal{O}}) = \text{Cor}(\hat{\mathcal{O}}, \hat{\mathcal{O}})$$

using the summary map provided by the the density matrix $\langle\hat{\mathcal{O}}\rangle = \text{Tr}(\hat{\rho}\hat{\mathcal{O}})$.

(a) Let $\hat{\mathcal{O}}$ be a normal operator. Let $\mathcal{O}_n$ be the different eigenvalues of $\hat{\mathcal{O}}$. Using the spectral theorem, show that $\hat{\mathcal{O}}$ can be written as below in terms of projector operators $\hat{\pi}_n$, which project a vector in $\mathbb{H}$ into the subspace on which $\hat{\mathcal{O}}$ has eigenvalue $\mathcal{O}_n$ and that these projectors satisfy the identities below:

$$\hat{\mathcal{O}} = \sum_n \mathcal{O}_n \hat{\pi}_n, \qquad \hat{\pi}_n\hat{\pi}_k = \delta_{nk}\hat{\pi}_k, \qquad \hat{\pi}_n = \hat{\pi}_n^\dagger, \qquad \sum_n \hat{\pi}_n = \mathbb{1}.$$

(b) Let $\hat{\mathcal{O}}$ be normal and $\hat{\pi}_n$ be its eigenspace projectors as above. Show that the density matrix is in a state of 0 operator spread for $\hat{\mathcal{O}}$ ($\text{Spr}(\hat{\mathcal{O}}) = 0$) if and only if then there exists an $n$ so that:

$$\hat{\rho} = \hat{\pi}_n\hat{\rho}\hat{\pi}_n.$$

(c) By following our previous proof of the uncertainty theorem, show that its conclusion is completely unchanged when density matrices are used. That is, show that for any two operators $\hat{A}$ and $\hat{B}$:

$$\text{Spr}(\hat{A})\text{Spr}(\hat{B}) \geq \left|\frac{1}{2}\langle[\hat{A}^\dagger, \hat{B}\rangle + \text{Cor}(\hat{A}, \hat{B})\right|^2.$$

(d) Show that the spectral spread theorem generalizes precisely to density matrices as well. That is, show that for any density matrix $\hat{\rho}$ and normal operator $\hat{\mathcal{O}}$, there is always an eigenvalue of $\hat{\mathcal{O}}$ within a distance of at most $\sqrt{\text{Spr}(\hat{\mathcal{O}})}$ from $\langle\hat{\mathcal{O}}\rangle$.

**Problem 2.3: State Purification.** Consider a density matrix $\hat{\rho}$ for a system with Hilbert space $\mathbb{H}$.

(a) Explain that there exist nonnegative coefficients $p_n$ and linearly independent, but not necessary orthogonal, normalized state vectors $|\Psi_n\rangle$ so that:

$$\hat{\rho} = \sum_n p_n |\Psi_n\rangle\langle\Psi_n|, \qquad \sum_n p_n = 1.$$

Explain that if $\hat{\rho}$ is a mixed state then there are infinitely many decompositions of $\hat{\rho}$ of this form.

(b) Consider one particular decomposition of $\hat{\rho}$ in the form of the previous part. Consider a new Hilbert space $\mathbb{H}'$ of dimension at least as large as the number of terms in the decomposition of $\hat{\rho}$ and define the total Hilbert space $\mathbb{H}_{\text{total}} = \mathbb{H} \otimes \mathbb{H}'$. Let $|\Phi_n\rangle$ be orthonormal vectors in $\mathbb{H}'$ where the range of $n$ is equal to the number of terms in the decomposition of $\hat{\rho}$. Consider the state vector:

$$|\Psi_{\text{total}}\rangle = \sum_n \sqrt{p_n}|\Psi_n\rangle \otimes |\Phi_n\rangle.$$

Show that the reduced density matrix on $\mathbb{H}$ derived from this state vector is precisely $\hat{\rho}$. Thus, conclude that every density matrix can be understood as the reduced state of a pure state on a Hilbert space of sufficiently large dimension. Explain that there are infinitely many ways to "purify" a density matrix.

(c) Show that the usual state vector definitions of operator summary, correlation, and spread for operators that act only in $\mathbb{H}$ with this state vector $|\Psi_{\text{total}}\rangle$ reduce to the density matrix definitions we have given

for operator summary, correlation, and spread.

**Problem 2.4: Action for Quantum Mechanics.** Consider a density matrix $\hat{\rho}(t)$ and a unitary operator $\hat{U}(t)$ on a Hilbert space. (Do not assume $\hat{U}(t)$ a priori to be the time evolution operator).

(a) Define the velocity operator $\hat{V}(t)$ and the variation operator $\delta\hat{\theta}(t)$ by:

$$\hat{V}(t) = i\hbar \frac{d\hat{U}}{dt}\hat{U}^\dagger, \qquad \delta\hat{\theta}(t) = i\hbar \delta\hat{U}\hat{U}^\dagger \tag{2.1.136}$$

where $\delta\hat{U}$ represents a general variation of $\hat{U}$ that preserves unitarity. Show that $\hat{V}$ and $\delta\hat{\theta}$ are self-adjoint and that there are no other constraints placed on them from the unitarity of $\hat{U}$.

(b) Show that under variations of $\hat{U}$:

$$\delta\hat{V} = \frac{d}{dt}\delta\hat{\theta} + \frac{i}{\hbar}[\delta\hat{\theta}, \hat{V}]. \tag{2.1.137}$$

(c) Consider (for Hamiltonian operator $\hat{H}(t)$) the action:

$$S[\hat{\rho}, \hat{U}] = \int_{t_0}^{t_f} \mathrm{Tr}(\hat{\rho}(\hat{V} - \hat{H}))dt. \tag{2.1.138}$$

Show under variations of this action which preserve the initial and final values of $\hat{U}$:

$$\delta S = \int_{t_0}^{t_f} \mathrm{Tr}\left(\delta\hat{\rho}(\hat{V} - \hat{H}) - \delta\hat{\theta}\left(\frac{d\hat{\rho}}{dt} + \frac{1}{i\hbar}[\hat{\rho}, \hat{V}]\right)\right) dt. \tag{2.1.139}$$

Thus, demonstrate the equations of motion produced by this action are precisely:

$$i\hbar\frac{d\hat{U}}{dt} = \hat{H}\hat{U}, \qquad \frac{d\hat{\rho}}{dt} + \frac{1}{i\hbar}[\hat{\rho}, \hat{H}] = 0. \tag{2.1.140}$$

The first shows that $\hat{U}$ does indeed become the time evolution operator and the second that $\hat{\rho}$ satisfies the von Neumann equation as desired. Thus, $S$ is the action for quantum mechanics.

## 2.2 Classical Limit

Quantum mechanics is our status quo best fundamental formalism for physics. However, very often classical physics is in excellent agreement with experiment. In this section we will understand thoroughly how the classical world emerges out of the quantum one under appropriate conditions. Quantum dynamics is fully specified by the Hamiltonian operator $\hat{H}$ appearing in the von Neumann equation whereas classical dynamics is fully specified by a classical Hamiltonian function $\mathcal{H}_{\mathrm{cl}}$. Whatever $\hat{H}$ a quantum system has thus will imply a particular $\mathcal{H}_{\mathrm{cl}}$ whenever that quantum system is well approximated by classical mechanics.

Consider a quantum system on a configuration Hilbert space with $d$ coordinate operators $\hat{Q}^a$ and their conjugate momentum operators $\hat{P}_a$. Automatically these operators satisfy the canonical commutation relations in (2.1.9). By Dirac's theorem (2.1.131), their operator summaries are canonically conjugate variables in the classical sense:

$$\{\langle\hat{Q}^a\rangle, \langle\hat{Q}^b\rangle\} = 0, \qquad \{\langle\hat{Q}^a\rangle, \langle\hat{P}_b\rangle\} = \delta^a_b, \qquad \{\langle\hat{P}_a\rangle, \langle\hat{P}_b\rangle\} = 0. \tag{2.2.1}$$

Define $\bar{Q}^a = \langle\hat{Q}^a\rangle$ and $\bar{P}_a = \langle\hat{P}_a\rangle$ as shorthand for these quantities.

### 2.2.1 Weyl Transform

In order to get a classical Hamiltonian function out of the Hamiltonian operator, we need some reasonable way of mapping an operator on Hilbert space to a function on classical phase space. We can find a reasonable such map by studying the matrix elements of the operator. From the Hamiltonian operator $\hat{H}(t)$, define the complex valued function $\zeta(t, Q, P)$ by the matrix element combination:

$$\zeta(t, Q, P) = \frac{\langle P|\hat{H}(t)|Q\rangle}{\langle P|Q\rangle}. \tag{2.2.2}$$

The fact that $\hat{H}$ is self-adjoint imposes conditions on $\zeta$. In particular:

$$\zeta^*(t,Q,P) = \frac{\langle Q|\hat{H}^\dagger|P\rangle}{\langle Q|P\rangle} = \frac{1}{\langle Q|P\rangle}\langle Q|\int |P'\rangle\langle P'|Q'\rangle\langle Q'|\zeta(t,Q',P')d^dQ'd^dP'|P\rangle \tag{2.2.3}$$

$$= \int e^{-\frac{i}{\hbar}(Q^a-Q'^a)(P_a-P'_a)}\zeta(t,Q',P')\frac{d^dQ'd^dP'}{(2\pi\hbar)^d}. \tag{2.2.4}$$

$\zeta$ and $\zeta^*$ have a very complicated relationship with each other. We can gradually disentangle this relationship by Fourier transforming. First do so to both sides with respect to $P$:

$$\int e^{-i\lambda^a P_a}\zeta^*(t,Q,P)d^dP = \int e^{\frac{i}{\hbar}(Q^a-Q'^a)P'_a}e^{-\frac{i}{\hbar}(\hbar\lambda^b+Q^b-Q'^b)P_b}\zeta(t,Q',P')\frac{d^dQ'd^dP'}{(2\pi\hbar)^d}d^dP \tag{2.2.5}$$

$$= \int e^{-i\lambda^a P'_a}\zeta(t,Q+\hbar\lambda,P')d^dP'. \tag{2.2.6}$$

We can make these two sides look even more similar by shifting the value of $Q^a$ to $Q^a - \frac{\hbar\lambda^a}{2}$:

$$\int e^{-i\lambda^a P'_a}\zeta^*\left(t,Q-\frac{\hbar\lambda}{2},P'\right)d^dP' = \int e^{-i\lambda^a P'_a}\zeta\left(t,Q+\frac{\hbar\lambda}{2},P'\right)d^dP'. \tag{2.2.7}$$

Now inverse Fourier transforming this with respect to $\lambda$ back into $P$, we find:

$$\int e^{i\lambda^a(P_a-P'_a)}\zeta^*\left(t,Q-\frac{\hbar\lambda}{2},P'\right)\frac{d^d\lambda d^dP'}{(2\pi)^d} = \int e^{i\lambda^a(P_a-P'_a)}\zeta\left(t,Q+\frac{\hbar\lambda}{2},P'\right)\frac{d^d\lambda d^dP'}{(2\pi)^d}. \tag{2.2.8}$$

By renaming the integration variable $\lambda^a$ on the left-hand side to $-\lambda^a$, we find:

$$\left(\int e^{i\lambda^a(P_a-P'_a)}\zeta\left(t,Q+\frac{\hbar\lambda}{2},P'\right)\frac{d^d\lambda d^dP'}{(2\pi)^d}\right)^* = \int e^{i\lambda^a(P_a-P'_a)}\zeta\left(t,Q+\frac{\hbar\lambda}{2},P'\right)\frac{d^d\lambda d^dP'}{(2\pi)^d}. \tag{2.2.9}$$

Therefore, the self-adjointness of $\hat{H}$ ensures that this complicated expression is a real-valued function (as it is equal to its own complex conjugate). Define then the function $H(t,Q,P)$ by this combination:

$$H(t,Q,P) = \int e^{i\lambda^a(P_a-P'_a)}\zeta\left(t,Q+\frac{\hbar\lambda}{2},P'\right)\frac{d^d\lambda d^dP'}{(2\pi)^d} \implies H^*(t,Q,P) = H(t,Q,P). \tag{2.2.10}$$

By using the definition of $\zeta$ in terms of $\hat{H}$, we can give $H$ a simpler and more direct definition in terms of $\hat{H}$:

$$H(t,Q,P) = \int e^{i\lambda^a(P_a-P'_a)}\frac{\langle P'|\hat{H}|Q+\frac{\hbar\lambda}{2}\rangle}{\langle P'|Q+\frac{\hbar\lambda}{2}\rangle}\frac{d^d\lambda d^dP'}{(2\pi)^d} \tag{2.2.11}$$

$$= (2\pi\hbar)^{\frac{d}{2}}\int e^{i\lambda^a P_a+\frac{i}{\hbar}\left(Q^a-\frac{\hbar\lambda^a}{2}\right)P'_a}\langle P'|\hat{H}|Q+\frac{\hbar\lambda}{2}\rangle\frac{d^d\lambda d^dP'}{(2\pi)^d} \tag{2.2.12}$$

$$= \int e^{i\lambda^a P_a}\langle Q-\frac{\hbar\lambda}{2}|\hat{H}|Q+\frac{\hbar\lambda}{2}\rangle\hbar^d d^d\lambda. \tag{2.2.13}$$

Finally changing to the new integration variable $q^a = \hbar\lambda$, we have:

$$\boxed{H(t,Q,P) = \int e^{\frac{i}{\hbar}q^aP_a}\left\langle Q-\frac{q}{2}\right|\hat{H}(t)\left|Q+\frac{q}{2}\right\rangle d^dq.} \tag{2.2.14}$$

We can recover $\hat{H}$ from $H$ by again Fourier transforming with respect to $P$:

$$\int e^{-\frac{i}{\hbar}q^aP_a}H(t,Q,P)d^dP = (2\pi\hbar)^d\left\langle Q-\frac{q}{2}\right|\hat{H}\left|Q+\frac{q}{2}\right\rangle. \tag{2.2.15}$$

Now replace $Q$ with $\frac{Q+Q'}{2}$ and $q$ with $Q-Q'$:

$$\int e^{\frac{i}{\hbar}P_a(Q'^a-Q^a)}H\left(t,\frac{Q+Q'}{2},P\right)d^P = (2\pi\hbar)^d\langle Q'|\hat{H}|Q\rangle. \tag{2.2.16}$$

Finally multiplying $|Q'\rangle\langle Q|$ onto both sides and integrating over both $Q$ and $Q'$, using that that the $|Q\rangle$ form a complete set of states, we have:

$$\boxed{\hat{H} = \int e^{\frac{i}{\hbar}P_a(Q'^a-Q^a)}H\left(t,\frac{Q+Q'}{2},P\right)|Q'\rangle\langle Q|d^dQ'\frac{d^dQd^dP}{(2\pi\hbar)^d}.} \tag{2.2.17}$$

So, the relationship between $\hat{H}$ and $H$ is invertible. This is a very nice discovery! We now see that for any operator on Hilbert space (in this case we used $\hat{H}$), we can define a function on phase space associated to it by (2.2.14). The original operator can be recovered from that function by (2.2.17). In the case that operator is self-adjoint, the associated function is real-valued (the converse holds as well, from the inverse formula). This map relating functions on classical phase space and operators on quantum Hilbert space is known as the **Weyl transform**. We write $\hat{\mathcal{W}}$ to mean the Weyl transform that takes a function on classical phase space and returns an operator on Hilbert space and $\mathcal{W}^{-1}$ to mean the inverse Weyl transform:

$$\boxed{\hat{H} = \hat{\mathcal{W}}(H), \qquad H = \mathcal{W}^{-1}(\hat{H}).} \tag{2.2.18}$$

The function $H$ given by the inverse Weyl transform of $\hat{H}$ is called the **Wigner function** associated to $\hat{H}$. The Weyl transform has several nice properties we will uncover over the rest of the section. It should be stressed that we have **NOT** shown that $H(t, Q, P)$ is the classical Hamiltonian. In fact, we will show that such a claim is generally wrong. The classical Hamiltonian is related to $H(t, Q, P)$ but that relationship is unambiguously unequal.

As written in (2.2.17), the Weyl transform does not seem symmetric in how it treats the coordinates and their conjugate momenta. However, we can make it plainly so with a bit of effort. To begin, rename the integration variable $Q'$ to $Q - \hbar\lambda$ (with $\lambda$ the new integration variable):

$$\hat{H} = \int e^{-\frac{i}{\hbar}\lambda^a P_a} H\left(t, Q - \frac{\hbar\lambda}{2}, P\right) |Q - \hbar\lambda\rangle\langle Q| d^d\lambda \frac{d^dQ d^dP}{(2\pi)^d}. \tag{2.2.19}$$

We can rewrite $|Q - \hbar\lambda\rangle$ more simply using:

$$|Q - \hbar\lambda\rangle = \int |P\rangle\langle P|Q - \hbar\lambda\rangle d^dP = \int |P\rangle e^{i\lambda^a P_a} \frac{e^{-\frac{i}{\hbar}Q^a P_a}}{(2\pi\hbar)^{\frac{d}{2}}} d^dP = e^{i\lambda^a \hat{P}_a}|Q\rangle \tag{2.2.20}$$

$$\implies \hat{H} = \int e^{i\lambda^a \hat{P}_a} e^{-i\lambda^b P_b} H\left(t, Q - \frac{\hbar\lambda}{2}, P\right) |Q\rangle\langle Q| d^d\lambda \frac{d^dQ d^dP}{(2\pi)^d}. \tag{2.2.21}$$

We can rewrite $|Q\rangle\langle Q|$ nicely by noting:

$$\delta(\hat{Q} - Q) = \int \delta(\hat{Q} - Q)|Q'\rangle\langle Q'|d^dQ' = \int \delta(Q' - Q)|Q'\rangle\langle Q'|d^dQ' = |Q\rangle\langle Q| \tag{2.2.22}$$

$$\implies \hat{H} = \int e^{i\lambda^a \hat{P}_a} e^{-i\lambda^b P_b} \delta(\hat{Q} - Q) H\left(t, Q - \frac{\hbar\lambda}{2}, P\right) d^d\lambda \frac{d^dQ d^dP}{(2\pi)^d}. \tag{2.2.23}$$

We can turn this Dirac delta into another exponential by using the fundamental Fourier integral:

$$\delta(\hat{Q} - Q) = \int \delta(Q' - Q)|Q'\rangle\langle Q'|d^dQ' = \int \frac{e^{ik_a(Q'^a - Q^a)}}{(2\pi)^d}|Q'\rangle\langle Q'|d^dQ' d^dk = \int \frac{e^{ik_a(\hat{Q}^a - Q^a)}}{(2\pi)^d} d^dk \tag{2.2.24}$$

$$\implies \hat{H} = \int e^{i\lambda^a \hat{P}_a} e^{ik_b \hat{Q}^b} e^{-ik_c Q^c - i\lambda^c P_c} H\left(t, Q - \frac{\hbar\lambda}{2}, P\right) \frac{d^d\lambda d^dk d^dQ d^dP}{(2\pi)^{2d}}. \tag{2.2.25}$$

The two operator exponentials cannot be combined the way that numerical exponentials can be because $\hat{Q}^a$ and $\hat{P}_a$ do not commute. However, using the Baker-Campbell-Hausdorff formula, we can combine them in a more complicated way. By the result of problem 1.21(b) with, in the notation of that problem, $\hat{A} = i\lambda^a \hat{P}_a$ and $\hat{B} = ik_b \hat{Q}^b$ and using the canonical commutation relation between $\hat{Q}^a$ and $\hat{P}_b$, we have:

$$e^{i\lambda^a \hat{P}_a} e^{ik_b \hat{Q}^b} = e^{ik_a \hat{Q}^a + i\lambda^a \hat{P}_a} e^{\frac{i\hbar}{2}\lambda^b k_b} \tag{2.2.26}$$

$$\implies \hat{H} = \int e^{ik_a \hat{Q}^a + i\lambda^a \hat{P}_a} e^{-ik_b\left(Q^b - \frac{\hbar\lambda^b}{2}\right) - i\lambda^b P_b} H\left(t, Q - \frac{\hbar\lambda}{2}, P\right) \frac{d^d\lambda d^dk d^dQ d^dP}{(2\pi)^{2d}}. \tag{2.2.27}$$

For the final step, simply shift the integration variable $Q$ to $Q + \frac{\hbar\lambda}{2}$:

$$\boxed{\hat{H} = \int e^{ik_a(\hat{Q}^a - Q^a) + i\lambda^a(\hat{P}_a - P_a)} H(t, Q, P) \frac{d^d\lambda d^dk d^dQ d^dP}{(2\pi)^{2d}}.} \tag{2.2.28}$$

The formula (2.2.28) can help us understand precisely what the Weyl transform is really doing to the function $H$. First, it Fourier transforms the full phase space dependence of $H$ from $(Q, P)$ to $(k, \lambda)$ then it inverse Fourier transforms back from $(k, \lambda)$ into the operators $(\hat{Q}, \hat{P})$ in a way symmetric between $\hat{Q}$ and $\hat{P}$. Thus, it essentially replaces $Q$ and $P$ with $\hat{Q}$ and $\hat{P}$ symmetrically Let us now compute explicitly that

means. Let $(q,p)$ be a point in classical phase space. By shifting the definitions of $Q$ and $P$ in (2.2.28) we can rewrite it as:

$$\hat{H} = \int e^{ik_a(\hat{Q}^a - q^a) + i\lambda^a(\hat{P}_a - p_a)} e^{-ik_b Q^b - i\lambda^b P_b} H(t, q+Q, p+P) \frac{d^d k d^d \lambda d^d Q d^d P}{(2\pi)^{2d}}. \tag{2.2.29}$$

Next, Taylor expand the operator exponential:

$$e^{ik_a(\hat{Q}^a - q^a) + i\lambda^a(\hat{P}_a - p_a)} = \sum_{N=0}^{\infty} \frac{i^N}{N!} \Big( k_a(\hat{Q}^a - q^a) + \lambda^a(\hat{P}_a - p_a) \Big)^N. \tag{2.2.30}$$

If we expand out the $N^{\text{th}}$ power operator polynomial here, it will contain once each every possible order of factors $k_a(\hat{Q}^a - q^a)$ and $\lambda^a(\hat{P}_a - p_a)$ so that the total number of factors of both is $N$. Recall the definition of the symmetrized product of operators from (1.2.79). Then, the operator binomial theorem is:

$$(\hat{A} + \hat{B})^N = \sum_{n=0}^{N} \frac{N!}{n!(N-n)!} \text{Sym}(\hat{A}, ...\hat{A}(n \text{ copies}), \hat{B}, ..., \hat{B}(N-n \text{ copies})). \tag{2.2.31}$$

Using this in the exponential series:

$$e^{ik_a(\hat{Q}^a - q^a) + i\lambda^a(\hat{P}_a - p_a)} = \sum_{N=0}^{\infty} \sum_{n=0}^{N} \frac{i^N}{N!} \frac{N!}{n!(N-n)!} \text{Sym}(k_a(\hat{Q}^a - q^a)(n \text{ copies}), \lambda^b(\hat{P}_b - p_b)(N-n \text{ copies})) \tag{2.2.32}$$

$$= \sum_{\substack{m=0 \\ n=0}}^{\infty,\infty} \frac{i^{m+n}}{m!n!} k_{a_1} ... k_{a_m} \lambda^{b_1} ... \lambda^{b_n} \text{Sym}((\hat{Q}^{a_1} - q^{a_1}), ..., (\hat{Q}^{a_m} - q^{a_m}), (\hat{P}_{b_1} - p_{b_1}), ..., (\hat{P}_{b_n} - p_{b_n})). \tag{2.2.33}$$

For brevity, define the operators:

$$\hat{M}^{a_1 ... a_m}{}_{b_1 ... b_n} = \text{Sym}((\hat{Q}^{a_1} - q^{a_1}), ..., (\hat{Q}^{a_m} - q^{a_m}), (\hat{P}_{b_1} - p_{b_1}), ..., (\hat{P}_{b_n} - p_{b_n})). \tag{2.2.34}$$

Importantly, these operators do not depend on any of our integration variables, so we can freely move them around the integral. Returning to the integral, we have:

$$\hat{H} = \sum_{\substack{m=0 \\ n=0}}^{\infty,\infty} \frac{i^{m+n}}{m!n!} \hat{M}^{a_1 ... a_m}{}_{b_1 ... b_n} \int k_{a_1} ... k_{a_m} \lambda^{b_1} ... \lambda^{b_n} e^{-ik_c Q^c - i\lambda^c P_c} H(t, q+Q, p+P) \frac{d^d k d^d \lambda d^d Q d^d P}{(2\pi)^{2d}}. \tag{2.2.35}$$

Each factor of $k_a$ can be rewritten as coming from a $Q^a$ derivative of the exponential. Similarly, each factor of $\lambda^a$ can be rewritten as coming from a $P_a$ derivative of the exponential. Thus:

$$\hat{H} = \sum_{\substack{m=0 \\ n=0}}^{\infty,\infty} \frac{(-1)^{m+n}}{m!n!} \hat{M}^{a_1 ... a_m}{}_{b_1 ... b_n} \int H(t, q+Q, p+P) \frac{\partial}{\partial Q^{a_1}} ... \frac{\partial}{\partial Q^{a_m}} \frac{\partial}{\partial P_{b_1}} ... \frac{\partial}{\partial P_{b_n}} e^{-ik_c Q^c - i\lambda^c P_c} \frac{d^d k d^d \lambda d^d Q d^d P}{(2\pi)^{2d}} \tag{2.2.36}$$

We can now integrate all of these derivatives by parts all derivatives off of the exponential and onto $H$. Doing so we will acquire no surface terms because the regulator (exponential damping) used in the definition of the Fourier integral kills them. Therefore:

$$\hat{H} = \sum_{\substack{m=0 \\ n=0}}^{\infty,\infty} \frac{\hat{M}^{a_1 ... a_m}{}_{b_1 ... b_n}}{m!n!} \int e^{-ik_c Q^c - i\lambda^c P_c} \frac{\partial}{\partial Q^{a_1}} ... \frac{\partial}{\partial Q^{a_m}} \frac{\partial}{\partial P_{b_1}} ... \frac{\partial}{\partial P_{b_n}} H(t, q+Q, p+P) \frac{d^d k d^d \lambda d^d Q d^d P}{(2\pi)^{2d}}. \tag{2.2.37}$$

Because $H(t, q+Q, p+P)$ depends only on $Q$ through $q+Q$, differentiating it with respect to $Q$ is the same as differentiating with respect to $q$. Similarly, differentiating it with respect to $P$ is the same as differentiating it with respect to $p$ and so:

$$\hat{H} = \sum_{\substack{m=0 \\ n=0}}^{\infty,\infty} \frac{\hat{M}^{a_1 ... a_m}{}_{b_1 ... b_n}}{m!n!} \frac{\partial}{\partial q^{a_1}} ... \frac{\partial}{\partial q^{a_m}} \frac{\partial}{\partial p_{b_1}} ... \frac{\partial}{\partial p_{b_n}} \int e^{-ik_c Q^c - i\lambda^c P_c} H(t, q+Q, p+P) \frac{d^d k d^d \lambda d^d Q d^d P}{(2\pi)^{2d}}. \tag{2.2.38}$$

All of the remaining integrals are now easy, giving us ultimately:

$$\boxed{\hat{H} = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{1}{m!n!} \frac{\partial^{m+n} H(t,q,p)}{\partial q^{a_1}...\partial q^{a_m}\partial p_{b_1}...\partial p_{b_n}} \text{Sym}(\hat{Q}^{a_1} - q^{a_1}, ..., \hat{Q}^{a_m} - q^{a_m}, \hat{P}_{b_1} - p_{b_1}, ..., \hat{P}_{b_n} - p_{b_n}).} \tag{2.2.39}$$

The original point $(q,p)$ was arbitrary and so the right hand side is independent of the choice of $(q,p)$ even though it is not manifest. The right hand side here is just the Taylor series for the function $H(t,q,p)$ when one wishes to formally evaluate it at $(\hat{Q},\hat{P})$ where the prescription for how to resolve any ordering ambiguities is to fully symmetrize the operator products. Thus as a formal shorthand for (2.2.39), we write:

$$\hat{H}(t) = \text{Sym}(H(t,\hat{Q},\hat{P})). \tag{2.2.40}$$

### 2.2.2 Classicality Conditions

We now require that we choose the coordinates for the coordinate operators $\hat{Q}^a$ and their conjugate momenta $\hat{P}_a$ so that the variables $\bar{Q}^a = \langle \hat{Q}^a \rangle$ and $\bar{P}_a = \langle \hat{P}_a \rangle$ have finite classical limits. Then, in that limit $\bar{Q}^a$ and $\bar{P}_a$ should satisfy Hamilton's equations from some Hamiltonian $\mathcal{H}_{\text{cl}}(t,\bar{Q},\bar{P})$. Using Ehrenfest's theorem we know that their exact quantum equations of motion are:

$$\frac{d\langle \hat{Q}^a \rangle}{dt} = \frac{1}{i\hbar}\langle [\hat{Q}^a, \hat{H}] \rangle, \qquad \frac{d\langle \hat{P}_a \rangle}{dt} = \frac{1}{i\hbar}\langle [\hat{P}_a, \hat{H}] \rangle. \tag{2.2.41}$$

Using the Weyl transform, we can make these look more suggestive. By (2.2.17) and integration by parts, we find for the first of these:

$$\frac{1}{i\hbar}[\hat{Q}^a, \hat{H}] = \frac{1}{i\hbar}\int e^{\frac{i}{\hbar}P_b(Q'^b - Q^b)} H\left(t, \frac{Q+Q'}{2}, P\right) (\hat{Q}^a|Q'\rangle\langle Q| - |Q'\rangle\langle Q|\hat{Q}^a) d^dQ' \frac{d^dQ d^P}{(2\pi\hbar)^d} \tag{2.2.42}$$

$$= -\int H\left(t, \frac{Q+Q'}{2}, P\right) \frac{\partial}{\partial P_a} e^{\frac{i}{\hbar}P_b(Q'^b - Q^b)} |Q'\rangle\langle Q| d^dQ' \frac{d^dQ d^P}{(2\pi\hbar)^d} \tag{2.2.43}$$

$$= \int e^{\frac{i}{\hbar}P_b(Q'^b - Q^b)} \frac{\partial H}{\partial P_a}\left(t, \frac{Q+Q'}{2}, P\right) |Q'\rangle\langle Q| d^dQ' \frac{d^dQ d^P}{(2\pi\hbar)^d} = \hat{\mathcal{W}}\left(\frac{\partial H}{\partial P_a}\right). \tag{2.2.44}$$

To evaluate the second, it is useful to recall (2.1.8), from which we have (by taking the adjoint):

$$\langle Q|\hat{P}_a = -i\hbar \frac{\partial \langle Q|}{\partial Q^a} \implies \hat{P}_a|Q\rangle = i\hbar \frac{\partial |Q\rangle}{\partial Q^a}. \tag{2.2.45}$$

Then, again using (2.2.17):

$$\frac{1}{i\hbar}[\hat{P}_a, \hat{H}] = \frac{1}{i\hbar}\int e^{\frac{i}{\hbar}P_b(Q'^b - Q^b)} H\left(t, \frac{Q+Q'}{2}, P\right) \left(i\hbar \frac{\partial |Q'\rangle}{\partial Q'^a}\langle Q| + i\hbar |Q'\rangle \frac{\partial \langle Q|}{\partial Q^a}\right) d^dQ' \frac{d^dQ d^dP}{(2\pi\hbar)^d} \tag{2.2.46}$$

$$= -\int |Q'\rangle\langle Q| \left(\frac{\partial}{\partial Q'^a} + \frac{\partial}{\partial Q^a}\right) \left(e^{\frac{i}{\hbar}P_b(Q'^b - Q^b)} H\left(t, \frac{Q+Q'}{2}, P\right)\right) d^dQ' \frac{d^dQ d^dP}{(2\pi\hbar)^d} \tag{2.2.47}$$

$$= -\int |Q'\rangle\langle Q| e^{\frac{i}{\hbar}P_b(Q'^b - Q^b)} \left.\frac{\partial H(t,q,P)}{\partial q^a}\right|_{q=\frac{Q+Q'}{2}} d^dQ' \frac{d^dQ d^dP}{(2\pi\hbar)^d} = -\hat{\mathcal{W}}\left(\frac{\partial H}{\partial Q^a}\right). \tag{2.2.48}$$

Because of these nice results, it is instructive to use them to define the operator derivative:

$$\frac{\partial \hat{H}}{\partial \hat{Q}^a} \stackrel{\text{def}}{=} \hat{\mathcal{W}}\left(\frac{\partial}{\partial Q^a}\mathcal{W}^{-1}(\hat{H})\right), \qquad \frac{\partial \hat{H}}{\partial \hat{P}_a} \stackrel{\text{def}}{=} \hat{\mathcal{W}}\left(\frac{\partial}{\partial P_a}\mathcal{W}^{-1}(\hat{H})\right). \tag{2.2.49}$$

Then, Ehrenfest's theorem for $\langle \hat{Q}^a \rangle$ and $\langle \hat{P}_a \rangle$ gives the exact quantum equations:

$$\frac{d\langle \hat{Q}^a \rangle}{dt} = \left\langle \frac{\partial \hat{H}}{\partial \hat{P}_a} \right\rangle, \qquad \frac{d\langle \hat{P}_a \rangle}{dt} = -\left\langle \frac{\partial \hat{H}}{\partial \hat{Q}^a} \right\rangle. \tag{2.2.50}$$

These are obviously visually similar to Hamilton's equations of motion, however we still need to think about in what circumstances classical mechanics should be approximately valid to truly make the connection.

In order for $\bar{Q}^a$ and $\bar{P}_a$ to satisfy classical equations of motion in terms of a Hamiltonian $\mathcal{H}_{\text{cl}}(t,\bar{Q},\bar{P})$, they must (in the classical limit) decouple from all of the other degrees of freedom present in the density matrix and only interact with each other. One way we can package all of these other degrees of freedom is in the various moments of the density matrix (meaning the operator summaries of symmetrized products of $\hat{Q}$ and $\hat{P}$). The full quantum theory is always described by the full Hamiltonian $\mathcal{H}_{\text{qm}} = \text{Tr}(\hat{\rho}\hat{H})$ and so to have

this decoupling we must have in the classical limit $\mathcal{H}_{\text{qm}} = \mathcal{H}_{\text{cl}}+$ small corrections. Using (2.2.39) with the choice $(q,p) = (\bar{Q},\bar{P})$, we have:

$$\mathcal{H}_{\text{qm}} = \sum_{\substack{m=0\\n=0}}^{\infty,\infty} \frac{1}{m!n!} \frac{\partial^{m+n} H(t,\bar{Q},\bar{P})}{\partial \bar{Q}^{a_1}...\partial \bar{Q}^{a_m}\partial \bar{P}_{b_1}...\partial \bar{P}_{b_n}} \langle \text{Sym}(\hat{Q}^{a_1}-\bar{Q}^{a_1},...,\hat{Q}^{a_m}-\bar{Q}^{a_m},\hat{P}_{b_1}-\bar{P}_{b_1},...,\hat{P}_{b_n}-\bar{P}_{b_n})\rangle \tag{2.2.51}$$

$$= H(t,\bar{Q},\bar{P}) + \frac{1}{2}\frac{\partial^2 H}{\partial \bar{Q}^a \partial \bar{Q}^b}\text{Cor}(\hat{Q}^a,\hat{Q}^b) + \frac{\partial^2 H}{\partial \bar{Q}^a \partial \bar{P}_b}\text{Cor}(\hat{Q}^a,\hat{P}_b) + \frac{1}{2}\frac{\partial^2 H}{\partial \bar{P}_a \partial \bar{P}_b}\text{Cor}(\hat{P}_a,\hat{P}_b) + \text{Higher Order Terms.} \tag{2.2.52}$$

The choice $(q,p) = (\bar{Q},\bar{P})$ is the only one for which the $m=1,n=0$ and $m=0,n=1$ terms vanish and this causes the higher order terms to become moments evaluated about the classical variables $\bar{Q},\bar{P}$. Thus, the series terms correspond to summaries of increasingly higher powers of deviations of the density matrix from its central summary position and momentum values. When the density matrix is essentially localized (with respect to our experimental resolution) in both position and momentum, the higher order terms in this series will be suppressed.

For the rest of this and the next subsection, it is notationally simplest to consider a one dimensional configuration space $d=1$. None of the logic we give needs to be changed for $d>1$, but the indices become annoying to manage, and so for readability we will take $d=1$ for now. Let $w_Q$ be our experimental resolution with which we can distinguish coordinates and $w_P$ be the experimental resolution with which we can distinguish momenta. A density matrix is approximately classical when its central moments are small compared to the experimental resolution:

$$\text{Approximate Classicality} \rightarrow \frac{\langle \text{Sym}(\hat{Q}-\bar{Q}(m \text{ copies}),\hat{P}-\bar{P}(n \text{ copies}))\rangle}{w_Q^m w_P^n} \ll 1. \tag{2.2.53}$$

This is because if the density matrix is such that the operator spread of an operator $\hat{\mathcal{O}}$ is much smaller than our experimental resolution for that operator's summary $\langle\hat{\mathcal{O}}\rangle$, then the behavior of that operator on the density matrix is well approximated by simply $\langle\hat{\mathcal{O}}\rangle\mathbb{1}$. Therefore, if our experimental resolution for our coordinates and momenta are large compared to their spreads, the coordinate and momenta operators can effectively be identified with their summary values.

The $(m,n)$ order variation $\Delta^{(m,n)}G(q,p)$ of a quantity $G$ at $(q,p)$ is defined by the differential variation in $G$ over the region of standard experimental resolution:

$$\Delta^{(m,n)}G(q,p) = \frac{\partial^{m+n}G(q,p)}{\partial q^m \partial p^n} w_Q^m w_p^n. \tag{2.2.54}$$

Then, $\mathcal{H}_{\text{qm}}$ can be written exactly as:

$$\mathcal{H}_{\text{qm}} = \sum_{\substack{m=0\\n=0}}^{\infty,\infty} \frac{1}{m!n!}\Delta^{(m,n)}H(t,\bar{Q},\bar{P}) \frac{\langle \text{Sym}(\hat{Q}-\bar{Q}(m \text{ copies}),\hat{P}-\bar{P}(n \text{ copies}))\rangle}{w_Q^m w_P^n}. \tag{2.2.55}$$

So, if the density matrix is approximately classical and the function $H$ has small variations over the window of experimental resolution, then the value of $\mathcal{H}_{\text{qm}}$ is approximately given by the low order terms in this series. Because of the classicality conditions, approximate classicality only extends to quantities of interest satisfying the hierarchy of scales in which the distance in classical phase space one must travel to vary them significantly is much larger than the experimental resolution which is in turn much larger than the spread of the density matrix. The first term in the series is purely $H(t,\bar{Q},\bar{P})$ and thus has the correct classical degrees of freedom. All of the higher order terms depend on increased numbers of degrees of freedom. So, in the classical limit this first term should dominate. This domination is required in order for the necessary decoupling of $\bar{Q}$ and $\bar{P}$ from the other degrees of freedom in the density matrix to occur, as only that first term depends exclusively on $\bar{Q}$ and $\bar{P}$. Thus, in the classical limit, we have $\mathcal{H}_{\text{qm}} = H(t,\bar{Q},\bar{P})+$ small corrections. We argued previously that in the classical limit $\mathcal{H}_{\text{qm}} = \mathcal{H}_{\text{cl}}(t,\bar{Q},\bar{P})+$ small corrections. Together these give that the imprecise relationship between $H$ and $\mathcal{H}_{\text{cl}}$ is:

$$\boxed{H(t,\bar{Q},\bar{P}) = \mathcal{H}_{\text{cl}}(t,\bar{Q},\bar{P}) + \text{ small corrections}} \tag{2.2.56}$$

where next we need to understand the precise meaning of "small corrections".

It is impossible to exactly realize the classical limit because the density matrix cannot be simultaneously

exactly localized in both coordinate space and momentum space. This is analogous to the fact that in reality it is impossible to exactly realize the nonrelativistic limit for any system in which there is motion, as at any nonzero velocity there are relativistic corrections present, however small in consideration of experimental resolution. The inability to localize the density matrix simultaneously in coordinate and momentum space is captured by the uncertainty theorem:

$$\text{Spr}(\hat{Q})\text{Spr}(\hat{P}) \geq \frac{\hbar^2}{4}. \tag{2.2.57}$$

Through use of the uncertainty theorem and (2.2.53), we find that approximate classicality requires:

$$\frac{\hbar^2}{4w_P^2} \ll \text{Spr}(\hat{Q}) \ll w_Q^2, \qquad \frac{\hbar^2}{4w_Q^2} \ll \text{Spr}(\hat{P}) \ll w_P^2, \qquad w_Q w_P \gg \frac{\hbar}{2} \tag{2.2.58}$$

exactly along the lines of (1.5.114), (1.5.115). Thus, $\hbar$ formally acts as a parameter which controls how small the simultaneous spreads can be made, in the way that $\frac{1}{c}$ in the nonrelativistic limit formally acts as a parameter which controls velocity corrections. Just as $c$ is always the same finite value and the nonrelativistic theory is an approximate effective theory at low velocities, $\hbar$ is always the same finite value and the "classical theory" in which $\bar{Q}$ and $\bar{P}$ form a closed system of equations is an appropriate approximate effective theory when the above scaling hierarchy is satisfied. Because of this, formally $\hbar \to 0^+$ is necessary (but not sufficient) for the classical limit in the same way that $\frac{1}{c} \to 0$ is necessary for the nonrelativistic limit. $\hbar \to 0^+$ is required to be able to simultaneously localize the density matrix in position and momentum space. However, we must then also perform that localization.

In problem 1.15 we defined the dilation generator $\hat{G}$ and dilation transformation $\hat{D}$ by:

$$\hat{G} = \frac{(\hat{Q}-q)(\hat{P}-p)+(\hat{P}-p)(\hat{Q}-q)}{2}, \qquad \hat{D}(\varepsilon) = e^{-\frac{i}{\hbar}\varepsilon\hat{G}}. \tag{2.2.59}$$

From these, we showed:

$$\hat{D}^\dagger \hat{Q} \hat{D} = q + e^{\varepsilon}(\hat{Q}-q), \qquad \hat{D}^\dagger \hat{P} \hat{D} = p + e^{-\varepsilon}(\hat{P}-p). \tag{2.2.60}$$

We would like to localize the density matrix about $\bar{Q}$ in position space and $\bar{P}$ in position space and so we will use dilation with center $(q,p) = (\bar{Q},\bar{P})$. With this choice, the density matrix spread depend on $\varepsilon$ as:

$$\text{Spr}_\varepsilon(\hat{Q}) = e^{2\varepsilon}\text{Spr}_0(\hat{Q}), \qquad \text{Spr}_\varepsilon(\hat{P}) = e^{-2\varepsilon}\text{Spr}_0(\hat{P}) \tag{2.2.61}$$

where $\text{Spr}_\varepsilon$ is computed using the dilated density matrix:

$$\hat{\rho}(\varepsilon) = \hat{D}(\varepsilon)\hat{\rho}_0\hat{D}^\dagger(\varepsilon) \tag{2.2.62}$$

and $\hat{\rho}_0$ is the base density matrix of the system. In order to localize the density matrix in position space we need to take $s = e^\varepsilon \to 0^+$. Then, the moments of the density matrix depend on $s$ as:

$$\langle \text{Sym}(\hat{Q}-\bar{Q}(m \text{ copies}), \hat{P}-\bar{P}(n \text{ copies}))\rangle_\varepsilon = s^{m-n}\langle \text{Sym}(\hat{Q}-\bar{Q}(m \text{ copies}), \hat{P}-\bar{P}(n \text{ copies}))\rangle_0. \tag{2.2.63}$$

Localizing the density matrix in position space will antilocalize it in momentum space unless we first take $\hbar \to 0^+$, because their spreads are inversely dependent on $\varepsilon$. To implement this correctly we need to be slightly careful. We would like to keep the summary momentum $\bar{P}$ of the density matrix fixed as it should survive in the classical limit. We can strip off $\bar{P}$ and leave behind a density matrix of summary momentum 0 by recalling the momentum translation operator in (1.5.39) and (1.5.55):

$$e^{\frac{i}{\hbar}p\hat{Q}}|P\rangle = |P+p\rangle \implies e^{\frac{i}{\hbar}p\hat{Q}}\hat{P}e^{-\frac{i}{\hbar}p\hat{Q}} = \int_{-\infty}^{\infty} P e^{\frac{i}{\hbar}p\hat{Q}}|P\rangle\langle P|e^{-\frac{i}{\hbar}p\hat{Q}}dP = \int_{-\infty}^{\infty} P|P+p\rangle\langle P+p|dP = \hat{P} - p\mathbb{1}. \tag{2.2.64}$$

So, if we define the shifted density matrix $\hat{\sigma}$ by:

$$\hat{\sigma} = e^{-\frac{i}{\hbar}\bar{P}\hat{Q}}\hat{\rho}_0 e^{\frac{i}{\hbar}\bar{P}\hat{Q}} \implies \langle\hat{P}\rangle_\sigma = \text{Tr}(\hat{\sigma}\hat{P}) = \text{Tr}\left(\hat{\rho}_0 e^{\frac{i}{\hbar}\bar{P}\hat{Q}}\hat{P}e^{-\frac{i}{\hbar}\bar{P}\hat{Q}}\right) = \text{Tr}(\hat{\rho}_0(\hat{P}-\bar{P}\mathbb{1})) = 0. \tag{2.2.65}$$

The momentum operator contains a factor of $\hbar$ in its definition and so it is useful to define the $\hat{K}$ operator so that for a wavefunction $|\Psi\rangle$:

$$\hat{K}|\Psi\rangle = \frac{\hat{P}}{\hbar}|\Psi\rangle = -i\int_{-\infty}^{\infty}\frac{\partial\Psi}{\partial Q}|Q\rangle dQ \tag{2.2.66}$$

In terms of these ingredients, $\mathcal{H}_{\text{qm}}$ computed using our dilated density matrix becomes:

$$\mathcal{H}_{\text{qm}} = \langle \hat{H} \rangle_{\varepsilon} = \sum_{\substack{m=0 \\ n=0}}^{\infty,\infty} \frac{\hbar^n s^{m-n}}{m!n!} \frac{\partial^{m+n} H(t,\bar{Q},\bar{P})}{\partial \bar{Q}^m \partial \bar{P}^n} \langle \text{Sym}(\hat{Q} - \bar{Q}(m \text{ copies}), \hat{K}(n \text{ copies})) \rangle_{\sigma}. \tag{2.2.67}$$

Because this last summary is computed using $\sigma$, we need not worry about the scaling of momentum spoiling the value of $\bar{P}$, because its summary momentum 0 stays 0.

$H(t,\bar{Q},\bar{P})$ is an at this time completely unknown function, and so is fully allowed to depend on $\hbar$ as well. By (2.2.56), since we are using classically well-behaved coordinates, $H$ is well-behaved in the classical limit. By (2.2.58), the formal classical limit will require us to take $\hbar \to 0^+$. So, $H$ must depend smoothly on $\hbar$ near this limit. We write its expansion in terms of $\hbar$ as:

$$H(t,\bar{Q},\bar{P}) = \sum_{k=0}^{\infty} \hbar^k H^{(k)}(t,\bar{Q},\bar{P}) \tag{2.2.68}$$

where each $H^{(k)}(t,\bar{Q},\bar{P})$ is $\hbar$ independent. Returning this expansion to $\mathcal{H}_{\text{qm}}$ we have:

$$\mathcal{H}_{\text{qm}} = \sum_{\substack{m=0 \\ n=0 \\ k=0}}^{\infty,\infty,\infty} \frac{\hbar^{n+k} s^{m-n}}{m!n!} \frac{\partial^{m+n} H^{(k)}(t,\bar{Q},\bar{P})}{\partial \bar{Q}^m \partial \bar{P}^n} \langle \text{Sym}(\hat{Q} - \bar{Q}(m \text{ copies}), \hat{K}(n \text{ copies})) \rangle_{\sigma}. \tag{2.2.69}$$

Now taking $\hbar \to 0^+$ explicitly:

$$\lim_{\hbar \to 0^+} \mathcal{H}_{\text{qm}} = \sum_{m=0}^{\infty} \frac{s^m}{m!} \frac{\partial^m H^{(0)}(t,\bar{Q},\bar{P})}{\partial \bar{Q}^m} \langle (\hat{Q} - \bar{Q})^m \rangle_{\sigma}. \tag{2.2.70}$$

Finally, the $s \to 0$ limit completes the localization to give the full classical limit:

$$\boxed{\mathcal{H}_{\text{cl}}(t,\bar{Q},\bar{P}) = \lim_{s \to 0^+} \lim_{\hbar \to 0^+} \mathcal{H}_{\text{qm}} = H^{(0)}(t,\bar{Q},\bar{P}).} \tag{2.2.71}$$

This is the precise statement of (2.2.56), which we can now write as:

$$\boxed{H(t,Q,P) = \mathcal{H}_{\text{cl}}(t,Q,P) + \sum_{n=1}^{\infty} \hbar^n H^{(n)}(t,Q,P)} \tag{2.2.72}$$

with the understanding that the classical limit does nothing to fix $H^{(n)}$ for $n > 0$. The corrections are justified as being called small in the classical limit because formally the classical limit takes $\hbar \to 0^+$. All of this same logic can be repeated for other quantities of interest besides the Hamiltonian. Therefore, the **the classical limit fixes the $\hbar$ independent piece of each operator's Wigner function to be the corresponding classical function**. But, **the classical limit says nothing about the $\hbar$ dependent pieces of an operator's Wigner function.**

A scheme for taking a classical theory (and thus an $\mathcal{H}_{\text{cl}}(t,Q,P)$) and assigning it a set of $\hbar$-corrections $H^{(n)}$ to arrive at a canonical choice of quantum theory (meaning a specific Hamiltonian operator $\hat{H}$) is called a **quantization scheme**. We will see shortly, that **there is no such thing as canonical quantization**. Specifically, what we mean is there does not exist a quantization scheme which is self-consistent under canonical transformations. The underlying cause of this is that a canonical transformations of the $Q$ and $P$ coordinates mixes different orders of $\hbar$ corrections with each other and because classical and quantum mechanics are not equivalent as representations, this mixing cannot possibly preserve the scheme. The meaning of the fact that canonical quantization does **NOT** exist is that you should *always* think about the quantum theory as more fundamental. A quantum theory defines an unambiguous corresponding classical theory by the classical limit but a classical theory does **NOT** define an unambiguous corresponding quantum theory. There are generally infinitely many different quantum theories possessing the same classical limit. Classical mechanics alone is not enough to learn everything (this should not be surprising; we need quantum mechanics for a reason). In some special cases a given classical limit *together with additional structures required of the quantum theory specific to that case* can be enough to fix that quantum theory.

### 2.2.3 Groenewold's theorem

The specific result that canonical quantization does not exist is called Groenewold's theorem, which we will now prove. The method of proof will be by contradiction. Suppose we have a quantum theory on a

configuration Hilbert space and suppose further that there exists a map from generators in classical phase space into generators in the quantum theory which is preserved by canonical transformations. A generator in the classical theory is a function $G(Q,P)$ on the classical phase space. A generator in the quantum theory is the operator summary $\langle\hat{G}\rangle$ of some operator $\hat{G}$. Thus, such a map would be characterized by a map from classical functions into operators, which we write as $\hat{R}$ ($R$ standing for "representation"). That is, such a map would have:

$$\hat{G} = \hat{R}(G). \tag{2.2.73}$$

The unit constant function 1 on classical phase space is necessarily mapped to the identity operator $\mathbb{1}$ so that $\langle\hat{R}(1)\rangle = 1$ always:

$$\hat{R}(1) = \mathbb{1}. \tag{2.2.74}$$

The statement that $\langle\hat{Q}\rangle$ and $\langle\hat{P}\rangle$ are the classically conjugate canonical variables is just:

$$\hat{R}(Q) = \hat{Q}, \qquad \hat{R}(P) = \hat{P}. \tag{2.2.75}$$

In order for generic extensive quantities in the classical theory to remain extensive in the quantum theory, $\hat{R}$ must be linear:

$$\hat{R}(aF(Q,P) + bG(Q,P)) = a\hat{R}(F) + b\hat{R}(G) \tag{2.2.76}$$

for constants $a, b$ and functions $F, G$. In order for a real-valued classical quantity to be mapped to a real-valued quantum quantity, $\langle\hat{R}(G)\rangle$ must be real for real-valued $G$, and so $\hat{R}(G)$ must be self-adjoint:

$$G = G^* \implies \hat{R}(G) = (\hat{R}(G))^\dagger. \tag{2.2.77}$$

So far these conditions are all satisfied by the Weyl transform. It is the final condition that will spell death for canonical quantization. In order for the map $\hat{R}$ to be preserved under canonical transformation, we must have that whether $\hat{R}$ is applied before or after performing a canonical transformation the same result is achieved. Consider the canonical transformation generated in classical mechanics by $G$. The corresponding transformation in quantum mechanics is thus generated by $\langle\hat{R}(G)\rangle$ and so in order to preserve $\hat{R}$, we must have that the change in any generator $F$ or its quantum version $\langle\hat{R}(F)\rangle$ coincide. That change is, as always, implemented by the Poisson bracket:

$$\{\langle\hat{R}(F)\rangle, \langle\hat{R}(G)\rangle\}_{\text{qm}} \stackrel{!}{=} \langle\hat{R}(\{F,G\}_{\text{cl}})\rangle. \tag{2.2.78}$$

The left hand side is how the change in $\langle\hat{R}(F)\rangle$ would be computed purely in quantum mechanics while the right hand side is how the change in $\langle\hat{R}(F)\rangle$ would be computed by first working in classical mechanics and then applying the map. For the map $\hat{R}$ to be preserved by $G$, these two must agree. This must hold for all $F, G$ in order for the map to be preserved under generic canonical transformations.

We now proceed with direct computation to show that this last property leads to contradictions. First note that using Dirac's theorem we can express it more simply as:

$$\frac{1}{i\hbar}\langle[\hat{R}(F), \hat{R}(G)]\rangle = \langle\hat{R}(\{F,G\}_{\text{cl}})\rangle \tag{2.2.79}$$

which because this must hold in a generic state $\hat{\rho}$, means:

$$[\hat{R}(F), \hat{R}(G)] = i\hbar\hat{R}(\{F,G\}_{\text{cl}}). \tag{2.2.80}$$

Consider $\hat{R}(Q^2)$. By computing the classical Poisson brackets:

$$\{Q^2, Q\}_{\text{cl}} = 0, \quad \{Q^2, P\}_{\text{cl}} = 2Q \quad \implies \quad [\hat{R}(Q^2), \hat{Q}] = 0, \quad [\hat{R}(Q^2), \hat{P}] = 2i\hbar\hat{Q}. \tag{2.2.81}$$

These can be rewritten as:

$$[\hat{R}(Q^2) - \hat{Q}^2, \hat{Q}] = 0, \qquad [\hat{R}(Q^2) - \hat{Q}^2, \hat{P}] = 0 \tag{2.2.82}$$

and thus $\hat{R}(Q^2) - \hat{Q}^2$ must not be a function of $\hat{Q}$ or $\hat{P}$; all that is left is for it to be proportional to the identity operator. So, for some constant $a$:

$$\hat{R}(Q^2) = \hat{Q}^2 + a\mathbb{1}. \tag{2.2.83}$$

Consider now $\hat{R}(QP)$. By computing the classical Poisson brackets:

$$\{QP, Q\}_{\text{cl}} = -Q, \quad \{QP, P\}_{\text{cl}} = P \quad \implies \quad [\hat{R}(QP), \hat{Q}] = -i\hbar\hat{Q}, \quad [\hat{R}(QP), \hat{P}] = i\hbar\hat{P}. \tag{2.2.84}$$

These can be rewritten as:

$$[\hat{R}(QP) - \frac{\hat{Q}\hat{P}+\hat{P}\hat{Q}}{2}, \hat{Q}] = 0, \quad [\hat{R}(QP) - \frac{\hat{Q}\hat{P}+\hat{P}\hat{Q}}{2}, \hat{P}] = 0 \quad \implies \quad \hat{R}(QP) = \frac{\hat{Q}\hat{P}+\hat{P}\hat{Q}}{2} + b\mathbb{1} \tag{2.2.85}$$

for some constant $b$.

Consider now $\hat{R}(P^2)$. By computing the classical Poisson brackets:

$$\{P^2, Q\}_{\text{cl}} = -2P, \quad \{P^2, P\}_{\text{cl}} = 0 \quad \implies \quad [\hat{R}(P^2), \hat{Q}] = -2i\hbar\hat{P}, \quad [\hat{R}(P^2), \hat{P}] = 0. \tag{2.2.86}$$

These can be rewritten as:

$$[\hat{R}(P^2) - \hat{P}^2, \hat{Q}] = 0, \quad [\hat{R}(P^2) - \hat{P}^2, \hat{P}] = 0 \implies \hat{R}(P^2) = \hat{P}^2 + c\mathbb{1} \tag{2.2.87}$$

for some constant $c$.

We can now determine the constants $a, b, c$ by looking at the Poisson brackets of $Q^2$, $QP$, $P^2$ with each other. By computing the classical Poisson brackets:

$$\{Q^2, QP\}_{\text{cl}} = 2Q^2, \qquad \{Q^2, P^2\}_{\text{cl}} = 4QP, \qquad \{QP, P^2\}_{\text{cl}} = 2P^2 \tag{2.2.88}$$

we have then:

$$[\hat{R}(Q^2), \hat{R}(QP)] = 2i\hbar\hat{R}(Q^2), \qquad [\hat{R}(Q^2), \hat{R}(P^2)] = 4i\hbar\hat{R}(QP), \qquad [\hat{R}(QP), \hat{R}(P^2)] = 2i\hbar\hat{R}(P^2). \tag{2.2.89}$$

Simplifying these using our found expressions for $\hat{R}(Q^2), \hat{R}(QP), \hat{R}(P^2)$ gives $a = b = c = 0$ and so ultimately, we must have:

$$\hat{R}(Q^2) = \hat{Q}^2, \qquad \hat{R}(QP) = \frac{\hat{Q}\hat{P}+\hat{P}\hat{Q}}{2}, \qquad \hat{R}(P^2) = \hat{P}^2. \tag{2.2.90}$$

Cool. Everything has worked so far for polynomials up through quadratic order.

Let us move onto computing how $\hat{R}$ must affect cubic polynomials. Begin with $Q^3$. The classical Poisson bracket with lower order polynomials requires:

$$\{Q^3, Q\}_{\text{cl}} = 0, \quad \{Q^3, P\}_{\text{cl}} = 3Q^2, \quad \{Q^3, Q^2\}_{\text{cl}} = 0, \quad \{Q^3, QP\}_{\text{cl}} = 3Q^3, \quad \{Q^3, P^2\}_{\text{cl}} = 6Q^2P \tag{2.2.91}$$

The first two of these as commutators require (using $\hat{R}(Q^2) = \hat{Q}^2$)):

$$[\hat{R}(Q^3) - \hat{Q}^3, \hat{Q}] = 0, \qquad [\hat{R}(Q^3) - \hat{Q}^3, \hat{P}] = 0 \implies \hat{R}(Q^3) = \hat{Q}^3 + d\mathbb{1} \tag{2.2.92}$$

for some constant $d$. The third of (2.2.91) is automatically satisfied for all $d$, however the fourth gives:

$$[\hat{R}(Q^3), \frac{\hat{Q}\hat{P}+\hat{P}\hat{Q}}{2}] = \hat{R}(Q^3) \implies d = 0 \implies \hat{R}(Q^3) = \hat{Q}^3. \tag{2.2.93}$$

Then, the fifth and final of (2.2.91) gives a convenient way to compute $\hat{R}(Q^2P)$:

$$\hat{R}(Q^2P) = \frac{1}{6i\hbar}[\hat{Q}^3, \hat{P}^2] = \frac{\hat{Q}^2\hat{P}+\hat{P}\hat{Q}^2}{2}. \tag{2.2.94}$$

We can now study the Poisson bracket conditions on $\hat{R}(Q^2P)$. By direct computation:

$$\{Q^2P, Q\}_{\text{cl}} = -Q^2, \quad \{Q^2P, P\}_{\text{cl}} = 2QP, \quad \{Q^2P, Q^2\}_{\text{cl}} = -2Q^3,$$
$$\{Q^2P, QP\}_{\text{cl}} = Q^2P, \quad \{Q^2P, P^2\}_{\text{cl}} = 4QP^2. \tag{2.2.95}$$

The first three of these are automatically satisfied by (2.2.94). However, the fourth as a commutator requires:

$$\frac{1}{i\hbar}[\hat{R}(Q^2P), \hat{R}(QP)] = \hat{R}(Q^2P). \tag{2.2.96}$$

Using (2.2.90) and (2.2.94) the left hand side is given by:

$$\frac{1}{i\hbar}[\hat{R}(Q^2P), \hat{R}(QP)] = \hat{Q}\hat{P}\hat{Q} \tag{2.2.97}$$

however the right hand side is given by (2.2.94). Because $\hat{Q}$ and $\hat{P}$ do not commute, these two expressions do not agree. Therefore, there is no way to define $\hat{R}(Q^2P)$ consistent with the preceding conditions. And so, the canonical quantization map $\hat{R}$ does not exist. This is Groenewold's theorem.

Because the map $\hat{R}$ does not exist, if some generic scheme is used to map a classical theory into a quantum theory in one system of coordinates, then it is possible to perform a canonical transformation of the classical coordinates so that if the same scheme were applied in the the resulting quantum theory will be inequivalent to the first one. Thus, any scheme you might attempt to use to turn a classical theory into a quantum one necessarily picks a preferred coordinate system for the classical phase space. Groenewold's

theorem concretely proves that the Poisson representation (using the classical Poisson bracket for the phase space variables) is an unfaithful representation of the operator formalism put forward in Orig. section 1.3. This is in contrast to the Schrödinger representation used in quantum mechanics, which is manifestly faithful as it just adds an overall multiplicative factor to the operators.

**IMPORTANT:** It is very interesting to note that there were no problems with $\hat{R}$ when we considered only generators that were at most quadratic polynomials in $Q$ and $P$. Such quadratic polynomials generate affine canonical transformation (linear with a possible inhomogeneous term) and so inadvertently we have proven that affine canonical transformations in classical and quantum mechanics are perfectly well mapped onto each other with no issues. This will be important when we discuss symmetries in quantum field theory as very often there the fields will only have symmetries in which they transform affinely.

### 2.2.4 Quantum Mechanics in Classical Phase Space

While there is no map consistent with canonical transformations that maps classical mechanics into quantum mechanics, there plainly is a map consistent with canonical transformations that maps quantum mechanics into classical mechanics: the classical limit. By (2.2.71), from a self-adjoint operator $\hat{G}$ on Hilbert space, we can define a corresponding classical function $G_{\text{cl}}$ as:

$$G_{\text{cl}}(Q,P) = \lim_{\hbar \to 0^+} \mathcal{W}^{-1}(\hat{G}) = \mathcal{C}(\hat{G}) \tag{2.2.98}$$

where we use this to define the classical limit map $\mathcal{C}$. Similarly but conversely to the previous subsection, in order for this map to be preserved by canonical transformations, we need, for operators $\hat{F}$ and $\hat{G}$:

$$\{\mathcal{C}(\hat{F}), \mathcal{C}(\hat{G})\}_{\text{cl}} = \mathcal{C}\left(\frac{1}{i\hbar}[\hat{F},\hat{G}]\right). \tag{2.2.99}$$

Momentarily we will confirm that this is in fact the case. In order to do so, and also to help get some other nice results, let us study how the Weyl transform interacts with multiplication. Consider operators $\hat{F}$ and $\hat{G}$. Each of these can be written as the Weyl transform of a Wigner function $(F(Q,P), G(Q,P))$. However, the Weyl transform of a product of Wigner functions is not the product of their Weyl transforms. That is:

$$\hat{F} = \hat{\mathcal{W}}(F), \qquad \hat{G} = \hat{\mathcal{W}}(G), \qquad \hat{F}\hat{G} \neq \hat{\mathcal{W}}(FG). \tag{2.2.100}$$

Instead, the effective product induced on Wigner functions by the Weyl transform is called the star product:

$$F \star G = \mathcal{W}^{-1}(\hat{\mathcal{W}}(F)\hat{\mathcal{W}}(G)). \tag{2.2.101}$$

To actually compute the star product, we can use (2.2.14) and (2.2.17). Applying these gives, in fully glory:

$$\begin{aligned} F \star G = \int & e^{\frac{i}{\hbar}(q^a P_a + P'_a(Q'^a - \widetilde{Q}'^a) + P''_a(Q''^a - \widetilde{Q}''^a))} F\left(\frac{Q' + \widetilde{Q}'}{2}, P'\right) G\left(\frac{Q'' + \widetilde{Q}''}{2}, P''\right) \\ & \cdot \langle Q - q/2 | Q' \rangle \langle \widetilde{Q}' | Q'' \rangle \langle \widetilde{Q}'' | Q + q/2 \rangle d^d q d^d Q' d^d Q'' \frac{d^d \widetilde{Q}' d^d P'}{(2\pi\hbar)^d} \frac{d^d \widetilde{Q}'' d^d P''}{(2\pi\hbar)^d} \end{aligned} \tag{2.2.102}$$

$$\begin{aligned} = \int & e^{\frac{i}{\hbar}(q^a P_a + P'_a(Q^a - q^a/2 - Q''^a) + P''_a(Q''^a - Q^a - q^a/2))} \\ & \cdot F\left(\frac{Q - q/2 + Q''}{2}, P'\right) G\left(\frac{Q'' + Q + q/2}{2}, P''\right) \frac{d^d q d^d P'}{(2\pi\hbar)^d} \frac{d^d Q'' d^d P''}{(2\pi\hbar)^d} \end{aligned} \tag{2.2.103}$$

after simplifying the $Q$ state inner products into Dirac deltas and using those deltas to reduce the integrals. Now change from the integration variable $Q''$ to a new variable $\widetilde{q}$ defined so that the $Q$ argument of $F$ becomes $Q + \widetilde{q}$. Then:

$$F \star G = \int e^{\frac{i}{\hbar}(q^a P_a - P'_a(2\widetilde{q}^a + q^a) + 2P''_a \widetilde{q}^a)} F\left(Q + \widetilde{q}, P'\right) G\left(Q + \widetilde{q} + \frac{q}{2}, P''\right) \frac{d^d q d^d P'}{(2\pi\hbar)^d} \frac{d^d \widetilde{q} d^d P''}{(\pi\hbar)^d}. \tag{2.2.104}$$

Similarly changing from the integration variable $q$ to a new variable $\widetilde{q}'$ so that the $Q$ argument of $G$ becomes $Q + \widetilde{q}'$ gives:

$$F \star G = \int e^{\frac{2i}{\hbar}(\widetilde{q}'^a(P_a - P'_a) + \widetilde{q}^a(P'^a_a - P_a))} F\left(Q + \widetilde{q}, P'\right) G(Q + \widetilde{q}', P'') \frac{d^d \widetilde{q} d^d P'}{(\pi\hbar)^d} \frac{d^d \widetilde{q}' d^d P''}{(\pi\hbar)^d}. \tag{2.2.105}$$

We can now just rename $\widetilde{q}$ to $q$ and $\widetilde{q}'$ to $q'$. For our last change of variables, change the integration variables $P'$ and $P''$ into $p$ and $p'$ by $P' = P + p$ and $P'' = P + p'$:

$$\boxed{F \star G = \int e^{\frac{2i}{\hbar}(q^a p'_a - q'^a p_a)} F(Q+q, P+p) G(Q+q', P+p') \frac{d^d q d^d p}{(\pi\hbar)^d} \frac{d^d q' d^d p'}{(\pi\hbar)^d}.} \tag{2.2.106}$$

This is the proper integral expression for the star product and is as simple as it can be made without any abuse of notation.

With a little bit of evil, we find a simpler and still correct (if applied correctly) formula for the star product. By using the operator exponential of the derivative operator, we can write (formally):

$$F(Q+q, P+p) = e^{q^a \frac{\partial}{\partial Q^a} + p_a \frac{\partial}{\partial P_a}} F(Q,P) \tag{2.2.107}$$

and analogously so for $G(Q+q', P+p')$ where the meaning of the exponential is as always made concrete by its series representation. It is useful to define the "left-derivative" as a derivative that acts only on expressions to its left, in contrast with the usual derivative which by standard convention acts on expressions to its right. When working with the left-derivative, we will also carefully note when we mean a right-derivative. In either case, the direction a derivative is supposed to act is marked by the direction of the arrow we draw above it:

$$F(Q,P) \frac{\overleftarrow{\partial}}{\partial Q^a} G(Q,P) = \frac{\partial F}{\partial Q^a} G, \qquad F(Q,P) \frac{\overrightarrow{\partial}}{\partial Q^a} G(Q,P) = F \frac{\partial G}{\partial Q^a}. \tag{2.2.108}$$

Then, we can rewrite (2.2.106) as:

$$F \star G = \int F(Q,P) e^{q^a \frac{\overleftarrow{\partial}}{\partial Q^a} + p_a \frac{\overleftarrow{\partial}}{\partial P_a}} e^{\frac{2i}{\hbar}(q^b p'_b - q'^b p_b)} e^{q^c \frac{\overrightarrow{\partial}}{\partial Q^c} + p_c \frac{\overrightarrow{\partial}}{\partial P_c}} G(Q,P) \frac{d^d q d^d p}{(\pi\hbar)^d} \frac{d^d q' d^d p'}{(\pi\hbar)^d}. \tag{2.2.109}$$

With regard to the $q, q', p, p'$ integrals, the $Q$ and $P$ derivatives can be treated as if they were constants (this can be made more proper by writing this in terms of the Fourier transforms of $F$ and $G$, in which case these derivatives literally become replaced by the Fourier transform parameters up to $i$'s). Performing the $q$ and $p$ integration (with this understanding of how to treat derivatives inspired by the Fourier transform) using the fundamental Fourier integral gives:

$$F \star G = \int F(Q,P) \delta\left(p' - \frac{i\hbar}{2} \frac{\overleftarrow{\partial}}{\partial Q}\right) \delta\left(q' + \frac{i\hbar}{2} \frac{\overleftarrow{\partial}}{\partial P}\right) e^{q'^a \frac{\overrightarrow{\partial}}{\partial Q^a} + p'_a \frac{\overrightarrow{\partial}}{\partial P_a}} G(Q,P) d^d q' d^d p'. \tag{2.2.110}$$

Completing then the final integration (and thankfully leaving behind the worst atrocity of this notation):

$$F \star G = F e^{\frac{i\hbar}{2}\left(\frac{\overleftarrow{\partial}}{\partial Q^a}\frac{\overrightarrow{\partial}}{\partial P_a} - \frac{\overleftarrow{\partial}}{\partial P_a}\frac{\overrightarrow{\partial}}{\partial Q^a}\right)} G = F e^{\frac{i\hbar}{2}\{\leftarrow,\rightarrow\}_{\text{cl}}} G \tag{2.2.111}$$

where we have noticed that the operator in the exponent is simply the classical Poisson bracket and so defined:

$$\{\leftarrow,\rightarrow\}_{\text{cl}} \stackrel{\text{def}}{=} \frac{\overleftarrow{\partial}}{\partial Q^a}\frac{\overrightarrow{\partial}}{\partial P_a} - \frac{\overleftarrow{\partial}}{\partial P_a}\frac{\overrightarrow{\partial}}{\partial Q^a} \implies F\{\leftarrow,\rightarrow\}_{\text{cl}} G = \{F,G\}_{\text{cl}}. \tag{2.2.112}$$

Note that for real $F$ and $G$, $F \star G$ is not necessarily real. This should not be surprising as for self-adjoint $\hat{F}$ and $\hat{G}$, $\hat{F}\hat{G}$ is not necessarily self-adjoint. The fact that the classical Poisson bracket is antisymmetric implies:

$$G \star F = G e^{\frac{i\hbar}{2}\{\leftarrow,\rightarrow\}_{\text{cl}}} F = F e^{-\frac{i\hbar}{2}\{\leftarrow,\rightarrow\}_{\text{cl}}} G \tag{2.2.113}$$

as at each order in the series swapping $F$ and $G$ gives a minus sign for each power of the Poisson bracket.

By Dirac's theorem, the quantum Poisson bracket of operator summaries simply reduces to the operator summary of $\frac{1}{i\hbar}$ times their commutator. Thus, it is an interesting question of what does the operator commutator do to Wigner functions? From the star product, it quickly follows that:

$$\mathcal{W}^{-1}\left(\frac{1}{i\hbar}[\hat{F},\hat{G}]\right) = \frac{1}{i\hbar}(F \star G - G \star F) = F \frac{e^{\frac{i\hbar}{2}\{\leftarrow,\rightarrow\}_{\text{cl}}} - e^{-\frac{i\hbar}{2}\{\leftarrow,\rightarrow\}_{\text{cl}}}}{i\hbar} G = F \frac{2}{\hbar} \sin\left(\frac{\hbar}{2}\{\leftarrow,\rightarrow\}_{\text{cl}}\right) G. \tag{2.2.114}$$

This motivates the definition of the **Moyal bracket**:

$$\boxed{\{\!\{F,G\}\!\} \stackrel{\text{def}}{=} F \frac{2}{\hbar} \sin\left(\frac{\hbar}{2}\{\leftarrow,\rightarrow\}_{\text{cl}}\right) G} \implies \frac{1}{i\hbar}[\hat{F},\hat{G}] = \hat{\mathcal{W}}\left(\left\{\!\!\left\{\mathcal{W}^{-1}(\hat{F}), \mathcal{W}^{-1}(\hat{G})\right\}\!\!\right\}\right). \tag{2.2.115}$$

Thus, the Weyl transform and the induced Moyal bracket on the Wigner functions provides a completely

invertible representation of all of the quantum operators in terms of functions on the classical phase space. There is no contradiction here with Groenewold's theorem, because the Moyal bracket is **NOT** the classical Poisson bracket.

From the Moyal bracket, it is very easy to see how the classical Poisson bracket emerges classically. Simply series expanding the Moyal bracket in $\hbar$ and defining $F^{(0)}$ and $G^{(0)}$ by the $\hbar \to 0^+$ limits of $F$ and $G$:

$$\{\!\{F, G\}\!\} = F \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \left(\frac{\hbar}{2}\right)^{2n} \{\leftarrow, \rightarrow\}_{\text{cl}}^{2n+1} G = \{F^{(0)}, G^{(0)}\}_{\text{cl}} + \mathcal{O}(\hbar). \tag{2.2.116}$$

The great comedy of this is that evidently classical mechanics is the $\sin x \approx x$ approximation of quantum mechanics. It is a very common error when people first learn about the Moyal bracket to write the above formula with $\mathcal{O}(\hbar^2)$ instead of $\mathcal{O}(\hbar)$. The origin of this error is that if $F$ and $G$ are not $\hbar$ expanded, then the Moyal bracket itself's first corrections are $\mathcal{O}(\hbar^2)$. The $F$ and $G$ are often not expanded in other sources because it is common to incorrectly believe that $F = F^{(0)}$ and $G = G^{(0)}$ because of a lack of appreciation for the nonexistence of a covariant meaning of quantization. Even if in a particular choice of coordinates $F = F^{(0)}$ for some quantity of interest, there is no reason to believe that holds for other quantities of interest in those coordinates, and such a correspondence is guaranteed to fail under generic changes of coordinates. From (2.2.115) it immediately follows that:

$$\lim_{\hbar \to 0^+} \mathcal{W}^{-1}\left(\frac{1}{i\hbar}[\hat{F}, \hat{G}]\right) = \lim_{\hbar \to 0^+} \left\{\!\!\left\{\mathcal{W}^{-1}(\hat{F}), \mathcal{W}^{-1}(\hat{G})\right\}\!\!\right\} \tag{2.2.117}$$

$$\implies \mathcal{C}\left(\frac{1}{i\hbar}[\hat{F}, \hat{G}]\right) = \{\lim_{\hbar \to 0^+} \mathcal{W}^{-1}(\hat{F}), \lim_{\hbar \to 0^+} \mathcal{W}^{-1}(\hat{G})\}_{\text{cl}} \tag{2.2.118}$$

which just simplifies to (2.2.99) as promised.

The existence of the Moyal bracket is perhaps initially surprising given a very old result of ours, namely theorem 1.2, in which we showed that the classical Poisson bracket is the only possible bracket of phase space functions satisfying bilinearity, antisymmetry, the Jacobi identity, and the Leibniz property. The operator commutator certainly satisfies all four of these properties, so shouldn't the Moyal bracket? But then, shouldn't the Moyal bracket be the classical Poisson bracket? There is actually no issue, because while yes the Moyal bracket does satisfy all four of these properties, it satisfies one of them *slightly* differently than we intended in Orig. section 1.4 and so avoids the logic of Orig. theorem 1.2. The subtlety is that in Orig. section 1.4 we assumed that the bracket we were developing would satisfy the Leibniz property with respect to ordinary multiplication. If that is required (in addition to the other three properties), the only possibility is the classical Poisson bracket. However, the Moyal bracket satisfies the Leibniz property not with respect to ordinary multiplication, but with respect to the star product instead:

$$\{\!\{F, GH\}\!\} = \{\!\{F, G\}\!\} \star H + G \star \{\!\{F, H\}\!\}. \tag{2.2.119}$$

This property follows automatically from the Leibniz property of operator commutators:

$$[\hat{F}, \hat{G}\hat{H}] = [\hat{F}, \hat{G}]\hat{H} + \hat{G}[\hat{F}, \hat{H}]. \tag{2.2.120}$$

Using the Moyal bracket, star product, and Weyl transform it is possible to express all of the equations of quantum mechanics purely in terms of Wigner functions, with no operators to be seen. This formulation is typically called "quantum mechanics in (classical) phase space". To begin this formulation, let us compute the trace of an operator product in terms of Wigner functions (using (2.2.17)):

$$\text{Tr}(\hat{F}\hat{G}) = \int \langle Q|\hat{F}\hat{G}|Q\rangle d^d Q \tag{2.2.121}$$

$$= \int e^{\frac{i}{\hbar}(P'_a - P_a)(Q''^a - Q^a)} F\left(\frac{Q+Q''}{2}, P\right) G\left(\frac{Q''+Q}{2}, P'\right) \frac{d^d Q d^d P}{(2\pi\hbar)^d} \frac{d^d Q'' d^d P'}{(2\pi\hbar)^d}. \tag{2.2.122}$$

Shift the integration variable $Q''$ to $Q'' + Q$. Then, shift the integration variable $Q$ to $Q - Q''/2$ and complete the the $Q''$ and $P'$ integrals. Simplifying, we arrive at the extremely convenient fact:

$$\text{Tr}(\hat{F}\hat{G}) = \int F(Q,P) G(Q,P) \frac{d^d Q d^d P}{(2\pi\hbar)^d}. \tag{2.2.123}$$

This does **NOT** generalize in the obvious way to a product of three operators, because of the distinction between the star product and the ordinary product.

Let $\rho(t, Q, P)$, called the Wigner density, be the Wigner function associated to the density matrix $\hat{\rho}(t)$.

Using (2.2.123) with $\hat{F} = \mathbb{1}$ and $\hat{G} = \hat{\rho}$, we have that the normalization condition on $\rho$ becomes:

$$\text{Tr}(\hat{\rho}) = 1 \implies \int \rho(t, q, P) \frac{d^d Q d^d P}{(2\pi\hbar)^d} = 1. \tag{2.2.124}$$

The fact that $\hat{\rho}$ is self-adjoint ensures that $\rho(t, Q, P)$ is real-valued. You would be reasonable to hope that since $\hat{\rho}$ is a nonnegative operator we could conclude that $\rho(t, Q, P)$ is a nonnegative function. I am sorry to disappoint you. There is no guarantee that $\rho(t, Q, P)$ is nonnegative (in fact in almost all settings it will be negative in some region of phase space). We will understand conceptually why this had to be true in section 2.5.5. Instead, the nonnegativity of $\hat{\rho}$ simply ensures that for any wavefunction $\psi(Q)$ (using (2.2.17)):

$$\langle \psi | \hat{\rho} | \psi \rangle = \int e^{\frac{i}{\hbar} P_a (Q'^a - Q^a)} \rho \left( t, \frac{Q + Q'}{2}, P \right) \psi^*(Q') \psi(Q) d^d Q' \frac{d^d Q d^d P}{(2\pi\hbar)^d} \geq 0 \tag{2.2.125}$$

which is not a particularly easy to use property. However, it does ensure that $\rho(t, Q, P)$ is only negative in a very restricted and contained way. Vaguely, $\rho$ is *mostly nonnegative*.

The quantities of interest in quantum mechanics are operator summaries $\langle \hat{F} \rangle = \text{Tr}(\hat{\rho}\hat{F})$. Using (2.2.123) these operator summaries can be expressed very nicely as:

$$\langle \hat{F} \rangle = \text{Tr}(\hat{\rho}\hat{F}) = \int \rho(t, Q, P) F(t, Q, P) \frac{d^d Q d^d P}{(2\pi\hbar)^d}. \tag{2.2.126}$$

In classical statistical mechanics, the dynamical variable is the phase space density $f(t, Q, P)$ which is a real-valued, nonnegative function which integrates to 1; the quantities of interest are statistical expectations computed using that density as the probability distribution on phase space. The Wigner density is similar to the phase space density in how it is working here. It is a real-valued function which integrates to 1 and which can only be negative in a limited way, and the quantities of interest are integrals of the same form as a statistical expectation. Continuing with this restricted analogy, we define the Moyalization or "M-expectation" of a Wigner function by:

$$\mathbb{M}(F) = \int \rho(t, Q, P) F(t, Q, P) \frac{d^d Q d^d P}{(2\pi\hbar)^d} \implies \boxed{\langle \hat{F} \rangle = \mathbb{M}(F).} \tag{2.2.127}$$

This Moyalization is not a genuine expectation in the sense of probability theory because the Wigner density is not a valid probability density (because it is not nonnegative). From (2.2.115) it immediately follows that:

$$\boxed{\{\mathbb{M}(F), \mathbb{M}(G)\} = \mathbb{M}(\{\!\{F, G\}\!\}).} \tag{2.2.128}$$

This formula establishes precisely that the Moyal bracket is the correct bracket to use for Wigner functions under a continuous transformation in quantum mechanics.

The analogy of the Wigner density to the classical phase space density is not accidental. Suppose we have an uncertain density matrix and so have some probability distribution on it. By (2.1.80), the statistical expectation of operator summaries is still described by standard density matrix formulas purely in terms of the statistical expectation of the density matrix. Let $\hat{\rho}$ be the statistical expectation of the density matrix and let $\rho(t, Q, P)$ be its associated Wigner density. It follows then that for generic $\hat{F}$:

$$\mathbb{E}[\langle \hat{F} \rangle] = \mathbb{M}(F) \tag{2.2.129}$$

where the Moyalization is computed with respect to $\rho(t, Q, P)$ and $F$ is the Wigner function associated to $\hat{F}$. In the classical limit, $\langle \hat{F} \rangle$ will approach some quantity of interest $F_{\text{cl}}$. In the case $\rho(t, Q, P)$ was simultaneously localized in position and momentum space in the classical limit, the resulting right hand side would just be $F_{\text{cl}}$ evaluated at the classical coordinates and momenta. However, in the case that in the classical limit $\rho(t, Q, P)$ is some nontrivial distribution $\rho^{(0)}(t, Q, P)$, the classical limit of this formula gives:

$$\mathbb{E}[F_{\text{cl}}] = \int \rho^{(0)}(t, Q, P) F^{(0)}(t, Q, P) d\mathbb{P} \tag{2.2.130}$$

where $d\mathbb{P}$ is the same phase space measure introduced in Orig. chapter 5. This is precisely the classical expectation formula from classical statistical mechanics, and so we can see that the classical phase space density $f(t, Q, P)$ is the classical limit of $\rho(t, Q, P)$:

$$f(t, Q, P) = \lim_{\hbar \to 0^+} \rho(t, Q, P). \tag{2.2.131}$$

This finally explains why $\hbar$ was introduced in the definition of the statistical mechanical phase space measure Orig. (5.2.4). I put it in there precisely so that this above equation would hold without any extra normalization.

The classical phase space density satisfies the Liouville equation. The Wigner density satisfies a similar

but appropriately more complicated equation of motion. From the von Neumann equation with $H$ the Wigner function associated to $\hat{H}$ and using (2.2.115), the Wigner density satisfies the Liouville-Moyal equation:

$$\mathcal{W}^{-1}\left(\frac{d\hat{\rho}}{dt}+\frac{1}{i\hbar}[\hat{\rho},\hat{H}]\right)=0 \implies \boxed{\frac{\partial\rho}{\partial t}+\{\!\{\rho,H\}\!\}=0.} \tag{2.2.132}$$

While there is no such thing as canonical quantization as a generic map for turning classical theories into quantum ones, the Moyal bracket can help constrain a quantum theory based on desired symmetries motivated by the classical theory. Suppose we have studied a system classically and found that it has some symmetry algebra, meaning that there is some set of classical quantities $T^a_{\text{cl}}$ which are a basis of generators for the symmetries of the theory. In classical mechanics these generators will satisfy some algebra:

$$\{T^a_{\text{cl}},T^b_{\text{cl}}\}_{\text{cl}}=f^{ab}{}_cT^c_{\text{cl}} \tag{2.2.133}$$

with some structure constants $f^{ab}{}_c$. Typically the symmetries of interest are motivated by geometric considerations, and so we typically expect that whatever the quantum version of the system fundamentally is, it has the same symmetries. Then, there must be operators $\hat{T}^a$ on the Hilbert space so that:

$$[\hat{T}^a,\hat{T}^b]=i\hbar f^{ab}{}_c\hat{T}^c \tag{2.2.134}$$

by theorem 1.5. Let $T^a$ be the Wigner functions associated to the $\hat{T}^a$. These means we need:

$$\{\!\{T^a,T^b\}\!\}=f^{ab}{}_cT^c, \qquad T^a=T^a_{\text{cl}}+\sum_{n=1}^{\infty}\hbar^nT^{(n)a}. \tag{2.2.135}$$

The first of these is the Wigner function statement of (2.2.135) while the second is necessary for (2.2.133) to be recovered in the classical limit. Enforcing both of these provides a set of conditions on the $\hbar$ corrections of each $T^a$. Depending on the precise algebra of interest, these conditions may be trivially satisfied by the classical $T^a_{\text{cl}}$, they may provide loose but nontrivial constraints on $T^{(n)a}$, they may fully determine the $T^{(n)a}$, or they may have no solution at all (in which case, one cannot extend the classical symmetry unchanged to the quantum theory). This is the limited extent in which classical mechanics plus some information about what symmetries we demand of the quantum theory can provide some restrictions on that quantum theory.

## Exercises

**Problem 2.5: WKB Approximation in general.** Consider a quantum system in a pure state on configuration Hilbert space. Define the real-valued fields $\rho(t,Q),S(t,Q)$ by:

$$\Psi(t,Q)=e^{\frac{i}{\hbar}S(t,Q)}\sqrt{\rho(t,Q)}.$$

(a) Show, from the Poisson brackets satisfied by $\Psi$ and $\Psi^*$ that $\rho$ and $S$ satisfy the Poisson brackets:

$$\{\rho(Q),\rho(Q')\}=0, \qquad \{\rho(Q),S(Q')\}=\delta(Q-Q'), \qquad \{S(Q),S(Q')\}=0.$$

Thus, conclude that $S$ is the conjugate momentum field to $\rho$.

(b) Show that in terms of $\rho$, $S$, and $H$ (Wigner function associated to the Hamiltonian operator $\hat{H}$), the quantum Hamiltonian becomes exactly:

$$\mathcal{H}=\int e^{i\left(\lambda^aP_a-\frac{1}{\hbar}\left(S\left(t,Q+\frac{\hbar\lambda}{2}\right)-S\left(t,Q+\frac{\hbar\lambda}{2}\right)\right)\right)}\sqrt{\rho\left(t,Q+\frac{\hbar\lambda}{2}\right)\rho\left(t,Q-\frac{\hbar\lambda}{2}\right)}H\left(t,Q,P\right)d^d\lambda\frac{d^dQd^dP}{(2\pi)^d}.$$

Explain how this may be systematically expanded in $\hbar$ to develop a semiclassical approximation to quantum mechanics. (Sadly, this expansion often is poorly behaved beyond low orders.)

(c) Show that in the $\hbar\to0^+$ the above Hamiltonian becomes:

$$\mathcal{H}=\int\rho(t,Q)\mathcal{H}_{\text{cl}}\left(t,Q,\frac{\partial S}{\partial Q}\right)d^dQ.$$

Show that the equations of motion from this Hamiltonian are precisely:

$$\frac{\partial S}{\partial t}+\mathcal{H}_{\text{cl}}\left(t,Q,\frac{\partial S}{\partial Q}\right)=0, \qquad \frac{\partial\rho}{\partial t}+\frac{\partial}{\partial Q^a}\left(\rho\left.\frac{\partial\mathcal{H}_{\text{cl}}(t,Q,P)}{\partial P_a}\right|_{P=\frac{\partial S}{\partial Q}}\right)=0.$$

Explain that these are exactly the Hamilton-Jacobi equation for $S$ and a continuity equation for $\rho$ in which the flow velocity is the classical velocity field.

**Problem 2.6: ⋆ Affinely Realized Symmetries.** Consider a classical system with Hamiltonian $\mathcal{H}(Q,P)$ so that all the continuous symmetries (other than time translation) of the theory are given in terms of a basis of generators $T_a(Q,P)$.

(a) Explain that for some structure constants $f_{ab}{}^c$:

$$\{\mathcal{H}, T_a\} = 0, \qquad \{T_a, T_b\} = f_{ab}{}^c T_c.$$

(b) Assume $\mathcal{H}$ is an analytic function of $(Q,P)$ and that all the $T_a$ are polynomials of at most quadratic order in $(Q,P)$. Show that under a general transformation generated by a combination of the $T_a$, the phase space variables $\Phi = (Q,P)$ before and after the transformation are related affinely:

$$\Phi'^a = M^a{}_b \Phi^b + C^a$$

for some constants $M^a{}_b$ and $C^a$ independent of $\Phi$ but dependent on the particular transformation. Such a system is said to have affinely realized symmetries and the coordinates in which these transformations take this form are called affine coordinates.

(c) Assume our system as above has affinely realized symmetries and that we are in affine coordinates. Show, using the Moyal bracket, that:

$$\{\!\{\mathcal{H}, T_a\}\!\} = 0, \qquad \{\!\{T_a, T_b\}\!\} = f_{ab}{}^c T_c.$$

Therefore, for such a system the symmetry algebra is unchanged when using the Moyal bracket.

(d) From a classical system with affinely realized symmetries, consider the quantum system defined by using the Hamiltonian operator $\hat{H}$ as below. Show that that system has symmetry generating operators precisely given by $\hat{T}_a$ as below:

$$\hat{H} = \hat{\mathcal{W}}(\mathcal{H}), \qquad \hat{T}_a = \hat{\mathcal{W}}(T_a) \implies [\hat{H}, \hat{T}_a], \qquad [\hat{T}_a, \hat{T}_b] = i\hbar f_{ab}{}^c \hat{T}_c.$$

Therefore, for a classical system with affinely realized symmetries and in affine coordinates, it is simple to define a corresponding quantum system in those affine coordinates which has all of the same affine symmetries as the classical theory. The is the case of "quantization" that is most relevant in applications to quantum field theory because very often all of our fields will only have affinely realized symmetries (for example, the behavior of tensors under coordinate transformation).

## 2.3 Lindblad Equation

This section will largely follow [31] quite closely. The time evolution of an isolated quantum system is always given by the von Neumann equation. But, quite often systems are interest and not reasonably approximated as isolated. So, it is natural to wonder how the subsystem density matrix of a larger quantum system evolves when that subsystem is not isolated from the remainder of the system. In general, that subsystem density matrix does not satisfy a closed equation of motion in terms of itself, which is totally unsurprising since it is interacting with other degrees of freedom. However, just as how in classical mechanics a mass moving on a table (after a little bit of averaging of the interacted with degrees of freedom) approximately satisfies a closed equation of motion with dissipation from friction, it is often possible and useful to approximately describe the time-averaged (over a fast time scale) dynamics of a subsystem density matrix with a closed equation of motion. What form can such an equation of motion take? That is out question this section.

When the subsystem is strongly coupled to other degrees of freedom, even if it can be given effectively closed dynamics, those dynamics are generally nonlinear (see for example [7]). This doesn't violate the linearity of quantum mechanics, because such linearity only applies to the full isolated system (for which it still holds). Once degrees of freedom are solved for and eliminated the result can be nonlinear, because the coefficients of the density matrix of one system depend on the degrees of freedom integrated out, which in turn needed to be solved for in terms of that same density matrix. In this section we will always assume that the failure of isolation is due to only a very weak coupling with other degrees of freedom, and this will allow us to still find an effective linear equation of motion for the subsystem density matrix. First, for now we will assume that we set out to look for a linear equation. Later on this section, we will justify that linearity in the case of weak coupling by a more direct derivation.

Since for now we are only concerned with our subsystem, we will just think about it as an open system with density matrix $\hat{\rho}(t)$ and not directly reference the degrees of freedom it is coupled to that we are silently integrating out. We assume that after that integrating out there is a linear effective equation of motion for this density matrix, i.e. that there is be some linear map $\mathcal{V}$ that relates the density matrix at time $t$ to the

density matrix at time 0:

$$\hat{\rho}(t) = \mathcal{V}(\hat{\rho}(0)). \tag{2.3.1}$$

Linearity is of course captured by:

$$\mathcal{V}(a\hat{\rho}_1 + b\hat{\rho}_2) = a\mathcal{V}(\hat{\rho}_1) + b\mathcal{V}(\hat{\rho}_2) \tag{2.3.2}$$

and this map must take the space of valid density matrices into itself.

A valid density matrix necessarily is self-adjoint, nonnegative, and of trace 1, so the outputs of $\mathcal{V}$ must be too. Because of this, $\mathcal{V}$ must be a so-called Completely Positive Trace Preserving Map (CPTP map). A map is called trace preserving if for any operator $\hat{A}$ on which it is defined it satisfies:

$$\mathrm{Tr}(\mathcal{V}(\hat{A})) = \mathrm{Tr}(\hat{A}). \tag{2.3.3}$$

This will protect the trace of our density matrix. Complete positivity is a bit more complicated. A linear map on operators is called positive if, for any nonnegative operator $\hat{A}$ on which it is defined, the operator $\mathcal{V}(\hat{A})$ is also nonnegative. A linear map $\mathcal{V}$ is called completely positive if the linear map $\mathcal{V} \otimes \mathbb{1}$ is positive for any size of identity matrix $\mathbb{1}$. It is plain that we need $\mathcal{V}$ to be positive so that the density matrix remains positive under time evolution. We additionally need complete positivity because if our open system is embedded in a system with some other degrees of freedom that do not evolve, the combined system density matrix must also be a valid density matrix. Such a system would evolve by the map $\mathcal{V} \otimes \mathbb{1}$, and thus this map must also be positive. It is surprising at first that complete positivity and positivity are distinct, so let us look at the standard simplest example where they differ. Consider a system of two spin-1/2 particles and the wavefunction:

$$|\psi\rangle = \frac{|+-\rangle + |-+\rangle}{\sqrt{2}}. \tag{2.3.4}$$

We will represent wavefunctions and operators as explicit vectors as:

$$|+\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \quad |-\rangle = \begin{pmatrix}0\\1\end{pmatrix}, \quad |++\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad |+-\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \quad |-+\rangle = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad |--\rangle = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{2.3.5}$$

The resulting density matrix for this system is:

$$\hat{\rho} = |\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix}0&0&0&0\\0&1&1&0\\0&1&1&0\\0&0&0&0\end{pmatrix}. \tag{2.3.6}$$

Now consider the transpose only of the first particle. Suppose $\hat{A}$ is nonnegative. For any $|\phi\rangle$:

$$0 \leq \langle\phi|\hat{A}|\phi\rangle = \langle\phi^*|\hat{A}^T|\phi^*\rangle. \tag{2.3.7}$$

Because this is true for any $|\phi\rangle$ and the complex conjugate is invertible (by just conjugating again), we have equivalently that for all $|\phi\rangle$:

$$\langle\phi|\hat{A}^T|\phi\rangle \geq 0. \tag{2.3.8}$$

Therefore the transpose is a positive linear map. We will now see that it is not completely positive because $T \otimes \mathbb{1}$ is not positive. Acting $T \otimes \mathbb{1}$ on $\hat{\rho}$ we have:

$$(T \otimes \mathbb{1})\hat{\rho} = \frac{1}{2}\begin{pmatrix}0&0&0&1\\0&1&0&0\\0&0&1&0\\1&0&0&0\end{pmatrix}. \tag{2.3.9}$$

This operator has $-\frac{1}{2}$ as an eigenvalue (under the eigenvector $\frac{|++\rangle - |--\rangle}{\sqrt{2}}$) and so it isn't positive. So, the transpose is not completely positive and could not possibly be a valid $\mathcal{V}$. This is why $\mathcal{V}$ must be CPTP.

Let us now find the general form of a CPTP map. Consider an operator $\hat{V}$. Plainly, the map $\mathcal{V}(\hat{\rho}) = \hat{V}\hat{\rho}\hat{V}^\dagger$ is positive since for any $|\psi\rangle$ we can define the vector $|\phi\rangle = \hat{V}^\dagger|\psi\rangle$ and thus have that for all $|\psi\rangle$:

$$\langle\psi|\mathcal{V}(\hat{\rho})|\psi\rangle = \langle\psi|\hat{V}\hat{\rho}\hat{V}^\dagger|\psi\rangle = \langle\phi|\hat{\rho}|\phi\rangle \geq 0 \tag{2.3.10}$$

by the nonnegativity of the density matrix. It is also true that a $\mathcal{V}$ of this form is completely positive since

$\mathcal{V} \otimes \mathbb{1}$ can be written on a tensor product space simply as:

$$(\mathcal{V} \otimes \mathbb{1})\hat{\rho} = (\hat{V} \otimes \mathbb{1})\hat{\rho}(\hat{V}^\dagger \otimes \mathbb{1}) \tag{2.3.11}$$

which is also of the same form as $\mathcal{V}$ with $\hat{V}$ replaced by $\hat{V} \otimes \mathbb{1}$. While every map of this form is completely positive, not every completely positive map takes this form.

Suppose the Hilbert space $\mathbb{H}$ for our system has dimension $d$ and let $|n\rangle$ be an arbitrary orthonormal basis of $\mathbb{H}$. For helping with calculation, we introduce a second ancillary Hilbert space $\mathbb{H}_A$ of the same dimension with arbitrary orthonormal basis $|n_A\rangle$. Now define the vector $|\psi\rangle$:

$$|\psi\rangle = \sum_n |n\rangle \otimes |n_A\rangle. \tag{2.3.12}$$

Of course, $|\psi\rangle\langle\psi|$ is a positive operator, because that form is positive for any nonzero vector. If $\mathcal{V}$ is a completely positive map acting on our system, then it must be a positive map when we lift it to act on the combined system as:

$$\mathcal{V}' = \mathcal{V} \otimes \mathbb{1}_A. \tag{2.3.13}$$

We can immediately compute:

$$\mathcal{V}'(|\psi\rangle\langle\psi|) = \sum_{mn} \mathcal{V}(|m\rangle\langle n|) \otimes |m_A\rangle\langle n_A| \implies \mathcal{V}(|m\rangle\langle n|) = \langle m_A|\mathcal{V}'(|\psi\rangle\langle\psi|)|n_A\rangle. \tag{2.3.14}$$

Therefore, we can recover how $\mathcal{V}$ acts on our original system solely from the matrix elements of how $\mathcal{V}'$ acts on $|\psi\rangle$. Because $\mathcal{V}'$ is positive, the operator $\mathcal{V}'(|\psi\rangle\langle\psi|)$ is nonnegative and thus can be diagonalized with an orthonormal basis of eigenvectors and nonnegative eigenvalues. That is, there exists an orthonormal basis $|\ell_\mathcal{V}\rangle$ so that for some nonnegative coefficients $a_\ell$:

$$\mathcal{V}'(|\psi\rangle\langle\psi|) = \sum_\ell a_\ell |\ell_\mathcal{V}\rangle\langle\ell_\mathcal{V}|. \tag{2.3.15}$$

If for a given $\ell$ the coefficient $a_\ell$ is nonzero, define the (unnormalized) vector:

$$|v_\ell\rangle = \sqrt{a_\ell}|b_\ell\rangle. \tag{2.3.16}$$

Then, we have that there exist vectors $|v_\ell\rangle$ so that $\mathcal{V}'(|\psi\rangle\langle\psi|)$ can be written as:

$$\mathcal{V}'(|\psi\rangle\langle\psi|) = \sum_\ell |v_\ell\rangle\langle v_\ell|. \tag{2.3.17}$$

Each vector $|v_\ell\rangle$ can be decomposed into the basis $|m\rangle \otimes |n_A\rangle$ with some coefficients:

$$|v_\ell\rangle = \sum_{mn} V_{\ell mn}|m\rangle \otimes |n_A\rangle. \tag{2.3.18}$$

From these coefficients we define the operators $\hat{V}_\ell$:

$$\hat{V}_\ell = \sum_{mn} V_{\ell mn}|m\rangle\langle n|. \tag{2.3.19}$$

From this definition, it immediately follows that:

$$(\hat{V}_\ell \otimes \mathbb{1}_A)|\psi\rangle = |v_\ell\rangle. \tag{2.3.20}$$

And from this we have:

$$\langle n_A|v_\ell\rangle = \hat{V}_\ell|n\rangle. \tag{2.3.21}$$

Now by combining (2.3.14), (2.3.17), and (2.3.21):

$$\mathcal{V}(|m\rangle\langle n|) = \sum_\ell \langle m_A|v_\ell\rangle\langle v_\ell|n_A\rangle = \sum_\ell \hat{V}_\ell|m\rangle\langle n|\hat{V}_\ell^\dagger. \tag{2.3.22}$$

Any operator $\hat{A}$ on our system is a linear combination of the $|m\rangle\langle n|$ and so by the linearity of $\mathcal{V}$:

$$\mathcal{V}(\hat{A}) = \sum_\ell \hat{V}_\ell \hat{A} \hat{V}_\ell^\dagger \tag{2.3.23}$$

meaning any completely positive linear map can be written in this form. Similarly, anything of this form is plainly completely positive because each term in the sum is. Thus, we have Choi's theorem: a linear map is completely positive if and only if it takes the form of (2.3.23) for some operators $\hat{V}_\ell$.

Now that we understand completely positive maps, let us see how to enforce trace preservation (2.3.3)

upon them. Suppose $\mathcal{V}$ is a completely positive map and we have expressed it in the form of (2.3.23) and consider (2.3.3) with $\hat{A} = |m\rangle\langle n|$. Simplifying both sides while using (2.3.23) yields:

$$\langle n| \sum_\ell \hat{V}_\ell^\dagger \hat{V}_\ell |m\rangle = \delta_{mn}. \tag{2.3.24}$$

If we now multiply $|n\rangle\langle m|$ onto both sides and sum over $m$ and $n$ we find:

$$\sum_\ell \hat{V}_\ell^\dagger \hat{V}_\ell = \mathbb{1}. \tag{2.3.25}$$

Thus the $\hat{V}_\ell$ operators must satisfy this in order for the map to be trace preserving. Further, if they do satisfy this the trace is in fact preserved by using the cyclicity of the trace:

$$\mathrm{Tr}(\mathcal{V}(\hat{A})) = \sum_\ell \mathrm{Tr}(\hat{V}_\ell^\dagger \hat{V}_\ell \hat{A}) = \mathrm{Tr}(\hat{A}). \tag{2.3.26}$$

So, we have the **Choi-Krauss theorem**: a map $\mathcal{V}$ is linear completely positive trace preserving if and only if for some operators $\hat{V}_\ell$ (called the Krauss operators of the map), it can be put into the form:

$$\mathcal{V}(\hat{A}) = \sum_\ell \hat{V}_\ell \hat{A} \hat{V}_\ell^\dagger \qquad \text{where} \qquad \sum_\ell \hat{V}_\ell^\dagger \hat{V}_\ell = \mathbb{1}. \tag{2.3.27}$$

Returning the Choi-Kraus theorem to (2.3.1), by assuming our open quantum system has a closed linear equation of motion, it must be the case that for some operators $\hat{V}_\ell(t, t_0)$:

$$\hat{\rho}(t) = \sum_\ell \hat{V}_\ell(t, t_0) \hat{\rho}(t_0) \hat{V}_\ell^\dagger(t, t_0) \tag{2.3.28}$$

where the $\hat{V}_\ell(t, t_0)$ operators satisfy:

$$\sum_\ell \hat{V}_\ell^\dagger(t, t_0) \hat{V}_\ell(t, t_0) = \mathbb{1}. \tag{2.3.29}$$

To proceed, we first introduce a useful basis of operators. On a $d$ dimensional Hilbert space, a generic operator has $d^2$ independent matrix elements and the set of such operators forms a vector space (of dimension $d^2$). We introduce the inner product on operators $\hat{A}, \hat{B}$ by:

$$(\hat{A}, \hat{B}) = \mathrm{Tr}(\hat{A}^\dagger \hat{B}). \tag{2.3.30}$$

This inner product satisfies all of the usual conditions on the definition of an inner product. Let $\hat{f}_\ell$ be a list of $d^2$ operators which are orthonormal under this inner product:

$$\mathrm{Tr}(\hat{f}_k^\dagger \hat{f}_\ell) = \delta_{k\ell}. \tag{2.3.31}$$

Thus, this set of $d^2$ operators form an orthonormal basis of the vector space of operators on our Hilbert space. So, there must exist coefficients $\alpha_{\ell k}(t, t_0)$ so that:

$$\hat{V}_\ell(t, t_0) = \sum_k \alpha_{\ell k}(t, t_0) \hat{f}_k. \tag{2.3.32}$$

Returning this form to (2.3.28), we have that:

$$\hat{\rho}(t) = \sum_{kk'} \left( \sum_\ell \alpha_{\ell k}(t, t_0) \alpha_{\ell k'}^*(t, t_0) \right) \hat{f}_k \hat{\rho}(t_0) \hat{f}_{k'}^\dagger = \sum_{k\ell} \beta_{k\ell}(t, t_0) \hat{f}_k \hat{\rho}(t_0) \hat{f}_\ell^\dagger \tag{2.3.33}$$

where we have introduced the coefficients:

$$\beta_{kk'}(t, t_0) = \sum_\ell \alpha_{\ell k}(t, t_0) \alpha_{\ell k'}^*(t, t_0). \tag{2.3.34}$$

Immediately, these coefficients satisfy:

$$\beta_{kk'}^*(t, t_0) = \beta_{k'k}(t, t_0). \tag{2.3.35}$$

We can simplify further by choosing the $\hat{f}_k$ intelligently. We can take by definition one of the $\hat{f}_k$ to be proportional to the identity operator (call it $\hat{f}_0 = \frac{1}{\sqrt{d}}\mathbb{1}$). Then, orthonormality requires all of the other $\hat{f}_k$ to be traceless. So, now write $\hat{F}_k$ for the traceless $\hat{f}_k$ and when the $k$ index of $\hat{F}_k$ is summed over it is understood that it is covering all of the $\hat{f}_k$ except the one proportional to the identity operator. Expanding

in this way, (2.3.33) becomes:

$$\hat{\rho}(t) = \sum_{k\ell} \beta_{k\ell}(t,t_0)\hat{F}_k\hat{\rho}(t_0)\hat{F}_\ell^\dagger + \frac{1}{\sqrt{d}}\sum_k \beta_{k0}(t,t_0)\hat{F}_k\hat{\rho}(t_0) + \frac{1}{\sqrt{d}}\sum_k \beta_{0k}(t,t_0)\hat{\rho}(t_0)\hat{F}_k^\dagger + \frac{1}{d}\beta_{00}(t,t_0)\hat{\rho}(t_0). \tag{2.3.36}$$

The limit $t_0 \to t$ gives that necessarily:

$$\beta_{k\ell}(t,t) = 0 \quad (k \text{ or } \ell \text{ nonzero}), \qquad \beta_{00}(t,t) = d. \tag{2.3.37}$$

Now differentiate both sides of (2.3.36) with respect to $t$, multiply by $i\hbar$, and evaluate in the limit $t_0 \to t$. After doing so:

$$i\hbar\dot{\hat{\rho}}(t) = \sum_{k\ell} ig_{k\ell}(t)\hat{F}_k\hat{\rho}(t)\hat{F}_\ell^\dagger + \frac{1}{\sqrt{d}}\sum_k \mathfrak{i}_k(t)\hat{F}_k\hat{\rho}(t) - \frac{1}{\sqrt{d}}\sum_k \mathfrak{i}_k^*(t)\hat{\rho}(t)\hat{F}_k^\dagger + \frac{i}{d}h(t)\hat{\rho}(t) \tag{2.3.38}$$

where we have introduced the coefficients:

$$g_{k\ell}(t) = \hbar\left.\frac{\partial\beta_{k\ell}(t',t)}{\partial t'}\right|_{t'=t}, \qquad \mathfrak{i}_k(t) = i\hbar\left.\frac{\partial\beta_{k0}(t',t)}{\partial t'}\right|_{t'=t}, \qquad h(t) = \hbar\left.\frac{\partial\beta_{00}(t',t)}{\partial t'}\right|_{t'=t} \tag{2.3.39}$$

and in the definition of $g_{k\ell}$ both $k$ and $\ell$ are nonzero. Yes, we intentionally only included the $i$ in the definition of $\mathfrak{i}_k(t)$. Let $\hat{F}(t)$ be defined by:

$$\hat{F}(t) = \frac{1}{\sqrt{d}}\sum_k \mathfrak{i}_k(t)\hat{F}_k. \tag{2.3.40}$$

Then:

$$i\hbar\dot{\hat{\rho}}(t) = \sum_{k\ell} ig_{k\ell}(t)\hat{F}_k\hat{\rho}(t)\hat{F}_\ell^\dagger + \hat{F}(t)\hat{\rho}(t) - \hat{\rho}(t)\hat{F}^\dagger(t) + \frac{i}{d}h(t)\hat{\rho}(t) \tag{2.3.41}$$

Decompose $\hat{F}(t)$ into its self-adjoint and anti-self-adjoint parts:

$$\hat{H}(t) = \frac{\hat{F}(t) + \hat{F}^\dagger(t)}{2}, \qquad \hat{G}(t) = \frac{\hat{F}(t) - \hat{F}^\dagger(t)}{2i} \implies \hat{F}(t) = \hat{H}(t) + i\hat{G}(t) \tag{2.3.42}$$

where $\hat{G}(t)$ and $\hat{H}(t)$ are now automatically self-adjoint. From these, we can put (2.3.41) into the form:

$$i\hbar\dot{\hat{\rho}}(t) = [\hat{H}(t),\hat{\rho}(t)] + \sum_{k\ell} ig_{k\ell}(t)\hat{F}_k\hat{\rho}(t)\hat{F}_\ell^\dagger + i\hat{G}(t)\hat{\rho}(t) + i\hat{\rho}(t)\hat{G}(t) + \frac{i}{d}h(t)\hat{\rho}(t). \tag{2.3.43}$$

We haven't yet imposed that our time evolution should be trace preserving. Doing so requires here that the trace of both sides is 0, so that $\hat{\rho}(t)$ maintains a constant trace of 1. Taking the trace and simplifying gives:

$$\mathrm{Tr}\left(\hat{\rho}\left(2\hat{G} + \frac{h}{d}\mathbb{1} + \sum_{k\ell} g_{k\ell}\hat{F}_\ell^\dagger\hat{F}_k\right)\right) \stackrel{!}{=} 0 \tag{2.3.44}$$

As this must hold for any density matrix, we must have:

$$\hat{G}(t) = -\frac{h(t)}{2d}\mathbb{1} - \frac{1}{2}\sum_{k\ell} g_{k\ell}(t)\hat{F}_\ell^\dagger\hat{F}_k. \tag{2.3.45}$$

Using this to simplify (2.3.43), we have:

$$i\hbar\frac{d\hat{\rho}(t)}{dt} = [\hat{H}(t),\hat{\rho}(t)] + i\sum_{k\ell} g_{k\ell}(t)(\hat{F}_k\hat{\rho}(t)\hat{F}_\ell^\dagger - \mathrm{Sym}(\hat{F}_\ell^\dagger\hat{F}_k,\hat{\rho}(t))). \tag{2.3.46}$$

By using the definition of $g_{k\ell}(t)$, we can continue to simplify. Because of (2.3.35), it follows from the definition of $g_{k\ell}(t)$ that:

$$g_{\ell k}^*(t) = g_{\ell k}(t). \tag{2.3.47}$$

Therefore, the matrix formed by the $g_{k\ell}(t)$ coefficients is self-adjoint and so at each $t$ it can be diagonalized by an orthonormal basis of eigenvectors $\iota_b(t)$ (with $b$ labeling which eigenvector is considered) with real eigenvalues $\lambda_b(t)$. We refer to the components of the eigenvector $\iota_b$ by $(\iota_b)_k$. So, $g_{k\ell}(t)$ is given by:

$$g_{k\ell}(t) = \sum_b \lambda_b(t)(\iota_b(t))_k(\iota_b(t))_\ell^*. \tag{2.3.48}$$

We can go even further using the properties of $\beta_{k\ell}(t,t_0)$. Consider a list of complex coefficients $\phi_k$ with $k$

taking the same index range as the indices of $\beta_{k\ell}$ (except for the value 0). Then, by (2.3.34):

$$\sum_{kk'} \beta_{kk'}(t,t_0)\phi_k^* \phi_{k'} = \sum_{\ell} \left| \sum_k \alpha_{\ell k}(t,t_0)\phi_k \right|^2 \geq 0. \tag{2.3.49}$$

So, the matrix formed by the $\beta_{kk'}(t,t_0)$ coefficients is nonnegative. By (2.3.37) at $t_0 = t$ we have $\beta_{kk'} = 0$ and so by the definition of $g_{k\ell}$:

$$g_{k\ell}(t) = \lim_{\varepsilon \to 0} \frac{\beta_{k\ell}(t+\varepsilon, t)}{\varepsilon}. \tag{2.3.50}$$

The right hand side forms the entries of a nonnegative matrix for all $\varepsilon$ and therefore so do the $g_{k\ell}(t)$. Thus, the eigenvalues $\lambda_b(t)$ are nonnegative allowing us to unambiguously take their positive square roots. Now define the "Jump Operators" $\hat{L}_k(t)$ by:

$$\hat{L}_k(t) = \sum_{\ell} \sqrt{\lambda_k(t)}(\iota_k(t))_\ell \hat{F}_\ell. \tag{2.3.51}$$

The jump operators are automatically orthogonal to each other under the operator inner product by using the orthonormality of the $\hat{F}_k$ and the $\iota_b$:

$$\mathrm{Tr}(\hat{L}_\ell^\dagger(t)\hat{L}_k(t)) = 0 \quad (k \neq \ell). \tag{2.3.52}$$

With these, we can now reexpress (2.3.46) as:

$$\boxed{i\hbar \frac{d\hat{\rho}(t)}{dt} = [\hat{H}(t), \hat{\rho}(t)] + i\sum_k (\hat{L}_k(t)\hat{\rho}(t)\hat{L}_k^\dagger(t) - \mathrm{Sym}(\hat{L}_k^\dagger(t)\hat{L}_k(t), \hat{\rho}(t)))}. \tag{2.3.53}$$

This result is known as the **Lindblad Equation** in jump diagonal form. It is the most general possible linear closed equation of motion for an open quantum system. Here $\hat{H}(t)$ some self-adjoint operator (which we call as always the Hamiltonian operator of the system) and the jump operators $\hat{L}_k(t)$ are some orthogonal set of traceless operators. These jump operators contain all of the dissipative effects on the open quantum system.

While the Lindblad equation as derived is general and exact (in the case that a linear closed equation of motion exists to begin with), it leaves many important questions unanswered. Should the Hamiltonian operator simply be the free Hamiltonian of the subsystem? (No.) How do you determine what the jump operators should be for the system? (It tells you nothing.) When should a closed linear equation of motion approximately exist? (Not usually.) To gain insight on these questions we need to approach the Lindblad equation differently, as a proper approximation of an interacting subsystem in a more hands on way. Consider a total system with total Hamiltonian operator $\hat{H}_T$ which consists of two interacting subsystems, the first being the open system we care about with free Hamiltonian operator $\hat{H}_0$ and the second being a set of environmental degrees of freedom with free Hamiltonian operator $\hat{H}_E$. The two interact with some interaction term $\hat{H}_1$. So, we write the full Hamiltonian operator as:

$$\hat{H}_T = \hat{H}_0 \otimes \mathbb{1} + \mathbb{1} \otimes \hat{H}_E + \hat{H}_1 \tag{2.3.54}$$

where as usual we will typically be lazy about the distinction between $\hat{H}_0 \otimes \mathbb{1}$ and $\hat{H}_0$ going forward. We will find later that it is convenient to slightly shift some pieces of $\hat{H}_0$ into $\hat{H}_1$ so for now we will actually use:

$$\hat{H}_T = \hat{H}_0' \otimes \mathbb{1} + \mathbb{1} \otimes \hat{H}_E + \hat{H}_1' \tag{2.3.55}$$

where $\hat{H}_0'$ will differ slightly from $\hat{H}_0$ in a way we will make precise very soon and $\hat{H}_1'$ differs from $\hat{H}_1$ precisely to compensate so that $\hat{H}_T$ is unchanged. The full passive picture density matrix $\hat{\rho}_T(t)$ of the total system then evolves according to the von Neumann equation as:

$$i\hbar \frac{d\hat{\rho}_T(t)}{dt} = [\hat{H}_T, \hat{\rho}_T(t)]. \tag{2.3.56}$$

Let $\hat{U}_T(t)$ be the time evolution operator for the total time evolution and assume the total system is isolated:

$$\hat{U}_T(t) = e^{-\frac{i}{\hbar}\hat{H}_T t}. \tag{2.3.57}$$

Then as always the total density matrix depends on time as:

$$\hat{\rho}_T(t) = \hat{U}_T(t)\hat{\rho}_T(0)\hat{U}_T^\dagger(t). \tag{2.3.58}$$

For actually performing our analysis it is useful to go to the interaction picture and so let $\hat{U}_S(t)$ be the time evolution operator for the subsystem and $\hat{U}_E(t)$ for the environment when the interaction is between the two

is neglected. Let $\hat{U}_0$ be combined free system's time evolution operator:

$$\hat{U}_S(t) = e^{-\frac{i}{\hbar}\hat{H}_0' t}, \qquad \hat{U}_E(t) = e^{-\frac{i}{\hbar}\hat{H}_E t}, \qquad \hat{U}_0(t) = \hat{U}_S(t) \otimes \hat{U}_E(t). \tag{2.3.59}$$

The interaction picture density matrix $\hat{\rho}_{TI}(t)$ of the total system is then defined by:

$$\hat{\rho}_{TI}(t) = \hat{U}_0^\dagger(t)\hat{\rho}_T(t)\hat{U}_0(t). \tag{2.3.60}$$

Differentiating this definition and using the von Neumann equation yields the interaction picture von Neumann equation:

$$i\hbar\frac{d\hat{\rho}_{TI}}{dt} = [\hat{H}_I, \hat{\rho}_{TI}] \tag{2.3.61}$$

where we have defined the interaction picture Hamiltonian operator $\hat{H}_I(t)$:

$$\hat{H}_I(t) = \hat{U}_0^\dagger(t)\hat{H}_1'(t)\hat{U}_0(t). \tag{2.3.62}$$

As usual, the interaction picture is only useful in-so-far as the interaction terms can be treated with perturbation theory. Exactly integrating the interaction picture von Neumann equation from 0 to $t$ gives:

$$\hat{\rho}_{TI}(t) = \hat{\rho}_T(0) - \frac{i}{\hbar}\int_0^t [\hat{H}_I(t'), \hat{\rho}_{TI}(t')]dt'. \tag{2.3.63}$$

Using this back in the von Neumann equation turns it into:

$$i\hbar\frac{d\hat{\rho}_{TI}}{dt} = [\hat{H}_I(t), \hat{\rho}_T(0)] - \frac{i}{\hbar}\int_0^t [\hat{H}_I(t), [\hat{H}_I(t'), \hat{\rho}_{TI}(t')]]dt'. \tag{2.3.64}$$

$\hat{\rho}_{TI}(t)$ and $\hat{\rho}_{TI}(t')$ differ on the order of $\hat{H}_I$ and so if we now approximate the interaction as small we can approximate $\hat{\rho}_{TI}(t')$ appearing in the integral by $\hat{\rho}_{TI}(t)$:

$$i\hbar\frac{d\hat{\rho}_{TI}}{dt} \approx [\hat{H}_I(t), \hat{\rho}_T(0)] - \frac{i}{\hbar}\int_0^t [\hat{H}_I(t), [\hat{H}_I(t'), \hat{\rho}_{TI}(t)]]dt'. \tag{2.3.65}$$

We don't want an equation of motion for $\hat{\rho}_{TI}$, we want one for the subsystem density matrix. As always, we define the subsystem density matrix $\hat{\rho}(t)$ by tracing over the environmental degrees of freedom:

$$\hat{\rho}(t) = \mathrm{Tr}_E(\hat{\rho}_T(t)). \tag{2.3.66}$$

The interaction picture subsystem density matrix is given by again factoring off its free evolution:

$$\hat{\rho}_I(t) = \hat{U}_S^\dagger(t)\hat{\rho}(t)\hat{U}_S(t). \tag{2.3.67}$$

Cyclicity of the partial trace (which is valid for operators acting only on the traced over degrees of freedom) together with the unitarity of $\hat{U}_E(t)$ allow us to express $\hat{\rho}_I(t)$ nicely in terms of $\hat{\rho}_{TI}(t)$:

$$\hat{\rho}_I(t) = \hat{U}_S^\dagger(t)\mathrm{Tr}_E(\hat{U}_E(t)\hat{U}_E^\dagger(t)\hat{\rho}_T(t))\hat{U}_S(t) = \mathrm{Tr}_E(\hat{U}_0^\dagger(t)\hat{\rho}_T(t)\hat{U}_0(t)) = \mathrm{Tr}_E(\hat{\rho}_{TI}(t)). \tag{2.3.68}$$

Tracing over (2.3.65) then gives:

$$i\hbar\frac{d\hat{\rho}_I(t)}{dt} \approx \mathrm{Tr}_E([\hat{H}_I(t), \hat{\rho}_T(0)]) - \frac{i}{\hbar}\int_0^t \mathrm{Tr}_E([\hat{H}_I(t), [\hat{H}_I(t'), \hat{\rho}_{TI}(t)]])dt'. \tag{2.3.69}$$

In order to proceed further we need to make further reasonable assumptions on the relationship between our open subsystem and its environment and to intelligently represent our interaction term. We assume that at time $t = 0$ they are completely independent and so:

$$\hat{\rho}_T(0) = \hat{\rho}(0) \otimes \hat{\rho}_E(0). \tag{2.3.70}$$

We further assume that at $t = 0$ the environment is in equilibrium so that under free time evolution by the environment Hamiltonian operator it remains stationary:

$$[\hat{H}_E, \hat{\rho}_E(0)] = 0. \tag{2.3.71}$$

The space of operators acting on the subsystem's Hilbert space is itself a vector space. Therefore, we can find a basis of operators $\hat{S}_k$ (not necessarily self-adjoint) so that any operator on the subsystem's Hilbert space can be written as a linear combination of these. Let $|n_E\rangle$ be an orthonormal basis for the environment's Hilbert space. Then, the combination $\langle m_E|\hat{H}_1|n_E\rangle$ is an operator on the subsystem's Hilbert space and thus can be written, for some coefficients $h_{kmn}$ as a linear combination of the $\hat{S}_k$:

$$\langle m_E|\hat{H}_1|n_E\rangle = \sum_k h_{kmn}\hat{S}_k. \tag{2.3.72}$$

By inserting complete sets of states on the environment, this allows us to write the full $\hat{H}_1$ as:

$$\hat{H}_1 = \sum_{kmn} \hat{h}_{kmn} \hat{S}_k \otimes |m_E\rangle\langle n_E|. \tag{2.3.73}$$

So, define the operators $\hat{E}_k$ on the environment's Hilbert space as:

$$\hat{E}_k = \sum_{mn} h_{kmn} |m_E\rangle\langle n_E|. \tag{2.3.74}$$

With these, $\hat{H}_1$ now takes the form:

$$\hat{H}_1 = \sum_k \hat{S}_k \otimes \hat{E}_k. \tag{2.3.75}$$

It is now time to specify how $\hat{H}_0'$ and $\hat{H}_0$ are related (and consequently how $\hat{H}_1'$ and $\hat{H}_1$ are related). $\hat{H}_0' - \hat{H}_0$ is some operator on the subsystem Hilbert space, and so there exist some constants $c_k$ so that:

$$\hat{H}_0' = \hat{H}_0 + \sum_k c_k \hat{S}_k. \tag{2.3.76}$$

To compensate the inclusion of this difference, we need:

$$\hat{H}_1' = \hat{H}_1 - \sum_k c_k \hat{S}_k \tag{2.3.77}$$

so that the total Hamiltonian operator is unchanged. In terms of this decomposition, the first term on the right hand side of (2.3.69) becomes:

$$\mathrm{Tr}_E([\hat{H}_I(t), \hat{\rho}_T(0)]) = \sum_k [\hat{U}_S^\dagger(t) \hat{S}_k \hat{U}_S(t), \hat{\rho}_S(0)] \mathrm{Tr}_E((\hat{E}_k - c_k \mathbb{1}) \hat{U}_E(t) \hat{\rho}_E(0) \hat{U}_E^\dagger(t)) \tag{2.3.78}$$

Because of (2.3.71):

$$\hat{U}_E(t) \hat{\rho}_E(0) \hat{U}_E^\dagger(t) = \hat{\rho}_E(0) \tag{2.3.79}$$

and so:

$$\mathrm{Tr}_E([\hat{H}_I(t), \hat{\rho}_T(0)]) = \sum_k [\hat{U}_S^\dagger(t) \hat{S}_k \hat{U}_S(t), \hat{\rho}_S(0)] (\mathrm{Tr}_E(\hat{E}_k \hat{\rho}_E(0)) - c_k). \tag{2.3.80}$$

We now define $\hat{H}_0'$ by choosing the constants $c_k$ as:

$$c_k = \mathrm{Tr}_E(\hat{E}_k \hat{\rho}_E(0)) \implies \mathrm{Tr}_E([\hat{H}_I(t), \hat{\rho}_T(0)]) = 0. \tag{2.3.81}$$

The reason $\hat{H}_0'$ is useful over $\hat{H}_0$ is that we want to approximate the interaction term $\hat{H}_1$ as having a small effect for perturbation theory; our choice of $c_k$ is precisely such that the expectation value of $\hat{H}_1$ is 0 in the initial state which prevents a constant accumulating drift that would happen if the $c_k$ were not included. It just helps perturbation theory be more accurate. Essentially, the presence of the environment causes a static shift in the effective Hamiltonian operator of the subsystem embedded in that environment.

As the subsystem and the environment interact, they will necessarily develop correlations and the total density matrix will no longer cleanly factor into the density matrix of the system and the density matrix of the environment. However, if we assume the environment is macroscopically large compared to the subsystem (containing many degrees of freedom), in statistical mechanical equilibrium, and very difficult to perturb from that equilibrium (due to its macroscopic scale), then we can approximate the total density matrix as factoring:

$$\hat{\rho}_{TI}(t) \approx \hat{\rho}_I(t) \otimes \hat{\rho}_E(0) \tag{2.3.82}$$

where $\hat{\rho}_E$ always remains at its $t = 0$ value because of the equilibrium of the environment. This intentional discarding of correlations is precisely the information loss necessary to explain the change in apparent entropy during the system's evolution. With these considerations, (2.3.69) now becomes:

$$i\hbar \frac{d\hat{\rho}_I(t)}{dt} \approx -\frac{i}{\hbar} \int_0^t \mathrm{Tr}_E([\hat{H}_I(t), [\hat{H}_I(t - t'), \hat{\rho}_I(t) \otimes \hat{\rho}_E(0)]]) dt' \tag{2.3.83}$$

where I have redefined the integration variable from $t'$ to $t - t'$ for later convenience.

So far we have allowed the basis $\hat{S}_k$ of subsystem operators to be generic. It is now time to make an intelligent choice for them. $\hat{H}_0'$ is a self-adjoint operator on the subsystem's Hilbert space. Consider two operators $\hat{A}$ and $\hat{B}$ (not necessarily self-adjoint) on that same Hilbert space. A one line calculation shows

and the cyclicity of the trace show that:

$$[\hat{H}_0', \hat{A}]^\dagger = [\hat{A}^\dagger, \hat{H}_0'] \implies \mathrm{Tr}(\hat{B}^\dagger[\hat{H}_0', \hat{A}]) = \mathrm{Tr}([\hat{H}_0', \hat{B}]^\dagger \hat{A}). \tag{2.3.84}$$

The space of operators on the subsystem's Hilbert space is itself a vector space and itself becomes a Hilbert space when equipped with the inner product defined by (2.3.30). The map $[\hat{H}_0', \cdot]$ is then a linear operator on that Hilbert space and this calculation shows that it is in fact a self-adjoint operator on that Hilbert space. Therefore, it can be diagonalized by an orthonormal basis of eigenvectors (which are now eigenoperators when regarding the original vector space of wavefunctions) with real eigenvalues. That is, there exists a basis $\hat{S}_k$ for the vector space of operators on the subsystem's Hilbert space so that:

$$[\hat{H}_0', \hat{S}_k] = -\hbar\omega_k \hat{S}_k \tag{2.3.85}$$

with real eigenvalues $-\hbar\omega_k$. Because $\hat{H}_0'$ acts so simply on this basis, we can explicitly compute $\hat{U}_S^\dagger \hat{S}_k \hat{U}_S$ as:

$$\frac{d}{dt}(\hat{U}_S^\dagger \hat{S}_k \hat{U}_S) = \frac{i}{\hbar}\hat{U}_S^\dagger[\hat{H}_0', \hat{S}_k]\hat{U}_S = -i\omega_k \hat{U}_S^\dagger \hat{S}_k \hat{U}_S \implies \hat{U}_S^\dagger \hat{S}_k \hat{U}_S = e^{-i\omega_k t}\hat{S}_k. \tag{2.3.86}$$

It is actually quite easy to find the $\hat{S}_k$. With $|n_S\rangle$ an orthonormal eigenbasis of $\hat{H}_0'$ on the subsystem's Hilbert space with eigenvalues $E_n'$, $|n_S\rangle\langle m_S|$ over the range of all pairs $(n, m)$ forms a basis for operators on the subsystem's Hilbert space and:

$$[\hat{H}_0', |n_S\rangle\langle m_S|] = (E_n' - E_m')|n_S\rangle\langle m_S| \tag{2.3.87}$$

and so we can identify each $(n, m)$ with a $k$ so that $\hat{S}_k = |n_S\rangle\langle m_S|$ and $\hbar\omega_k = E_m' - E_n'$. Now the interaction Hamiltonian operator can be computed quite simply as:

$$\hat{H}_I(t) = \sum_k e^{-i\omega_k t}\hat{S}_k \otimes \hat{U}_E^\dagger \delta\hat{E}_k \hat{U}_E = \sum_k e^{i\omega_k t}\hat{S}_k^\dagger \otimes \hat{U}_E^\dagger \delta\hat{E}_k^\dagger \hat{U}_E \tag{2.3.88}$$

where the second equation follows because $\hat{H}_I(t)$ is self-adjoint and we have defined for convenience:

$$\delta\hat{E}_k = \hat{E}_k - \mathrm{Tr}_E(\hat{E}_k \hat{\rho}_E(0))\mathbb{1}. \tag{2.3.89}$$

We now expand out both commutators in (2.3.83) and use the expression for $\hat{H}_I$ in terms of $\hat{E}_k^\dagger$ for terms which (cyclically under the trace) are to the right of $\hat{\rho}_E(0)$ (and separated only by factors of $\hat{U}_E, \hat{U}_E^\dagger$) and the corresponding expression in terms of $\hat{E}_k$ for terms which (cyclically under the trace) are to the left of $\hat{\rho}_E(0)$ (and separated only be factors of $\hat{U}_E, \hat{U}_E^\dagger$). Doing so yields:

$$\int_0^t \mathrm{Tr}_E([\hat{H}_I(t), [\hat{H}_I(t-t'), \hat{\rho}_I(t) \otimes \hat{\rho}_E(0)]])dt' = \sum_{k\ell} e^{i(\omega_k - \omega_\ell)t}\Big(\gamma_{k\ell}(t, \omega_\ell)[\hat{S}_k^\dagger, \hat{S}_\ell \hat{\rho}_I(t)] + \gamma_{\ell k}^*(t, \omega_k)[\hat{\rho}_I(t)\hat{S}_k^\dagger, \hat{S}_\ell]\Big) \tag{2.3.90}$$

where we have defined the $\gamma_{k\ell}(t, \omega)$ coefficients by:

$$\gamma_{k\ell}(t, \omega) = \int_0^t e^{i\omega t'}\,\mathrm{Tr}_E\left(\hat{\rho}_E(0)\delta\hat{E}_k^\dagger \hat{U}_E(t')\delta\hat{E}_\ell \hat{U}_E^\dagger(t')\right)dt'. \tag{2.3.91}$$

We still have more approximations to make. We now approximate that in (2.3.90) for $\omega_k \neq \omega_\ell$ the phase $e^{i(\omega_k - \omega_\ell)t}$ oscillates very rapidly on the time scale we are sensitive to. This is justified by considering that we are looking at the evolution of the interaction picture $\hat{\rho}_I(t)$, which should evolve on the time scale of the interaction, which being with a macroscopic environment we expect to be larger than the intrinsic time scales of the subsystem (which are contained in $\omega_k, \omega_\ell$). If this phase oscillates much more rapidly than anything else in this equation, we can time average it to vanish for $\omega_k \neq \omega_\ell$, giving us:

$$\int_0^t \mathrm{Tr}_E([\hat{H}_I(t), [\hat{H}_I(t-t'), \hat{\rho}_I(t) \otimes \hat{\rho}_E(0)]])dt' \approx \sum_{\substack{k\ell \\ \omega_k = \omega_\ell}} \left(\gamma_{k\ell}(t, \omega_\ell)[\hat{S}_k^\dagger, \hat{S}_\ell \hat{\rho}_I(t)] + \gamma_{\ell k}^*(t, \omega_\ell)[\hat{\rho}_I(t)\hat{S}_k^\dagger, \hat{S}_\ell]\right). \tag{2.3.92}$$

The integrand of (2.3.91) is the operator correlation of $\hat{E}_\ell$ at the initial time and $\hat{E}_k^\dagger$ at the time $t'$ (evolved in the active picture). If the environment is very macroscopic, its degrees of freedom will quickly become uncorrelated with each other and so we approximate the above integrand as falling to 0 very quickly for large $t'$. Because of this, we can effectively extend the upper integration limit to $\infty$ provided $t$ is not extremely

small. That is, we approximate:

$$\gamma_{k\ell}(t,\omega) \approx \lim_{t\to\infty} \gamma_{k\ell}(t,\omega) = \int_0^\infty e^{i\omega t'} \mathrm{Tr}_E \left(\hat{\rho}_E(0)\delta\hat{E}_k^\dagger \hat{U}_E(t')\delta\hat{E}_\ell \hat{U}_E^\dagger(t')\right) dt'. \tag{2.3.93}$$

Viewing these coefficients as a matrix, we break them into self-adjoint and anti-self-adjoint parts as:

$$\Gamma_{k\ell}(\omega) = \lim_{t\to\infty} \frac{\gamma_{k\ell}(t,\omega) + \gamma_{\ell k}^*(t,\omega)}{\hbar}, \qquad \pi_{k\ell}(\omega) = \lim_{t\to\infty} \frac{\gamma_{k\ell}(t,\omega) - \gamma_{\ell k}^*(t,\omega)}{2i\hbar}. \tag{2.3.94}$$

In terms of this decomposition, returning everything to (2.3.83) gives:

$$i\hbar \frac{d\hat{\rho}_I(t)}{dt} \approx \sum_{\substack{k\ell \\ \omega_k=\omega_\ell}} \pi_{k\ell}(\omega_k)[\hat{S}_k^\dagger \hat{S}_\ell, \hat{\rho}_I] + i \sum_{\substack{k\ell \\ \omega_k=\omega_\ell}} \Gamma_{k\ell}(\omega_k)(\hat{S}_\ell \hat{\rho}_I \hat{S}_k^\dagger - \mathrm{Sym}(\hat{S}_k^\dagger \hat{S}_\ell, \hat{\rho}_I)). \tag{2.3.95}$$

Now this is looking like the Lindblad equation! Because the matrix formed by the components $\Gamma_{k\ell}$ is self-adjoint, we can diagonalize it by some orthonormal basis of eigenvectors $\iota_b$ (with components $(\iota_b)_k$) and eigenvalues $\lambda_b$:

$$\Gamma_{k\ell}(\omega) = \sum_b \lambda_b(\omega)(\iota_b(\omega))_k (\iota_b(\omega))_\ell^*. \tag{2.3.96}$$

Using the explicit definition of $\Gamma_{k\ell}(\omega)$ and simplifying gives that it is precisely:

$$\Gamma_{k\ell}(\omega) = \frac{1}{\hbar}\int_{-\infty}^\infty e^{i\omega t} \mathrm{Tr}_E(\hat{\rho}_E(0)\delta\hat{E}_k^\dagger \hat{U}_E(t)\delta\hat{E}_\ell \hat{U}_E^\dagger(t))dt \tag{2.3.97}$$

which is (up to convention) the Fourier transform of the operator correlation of $\hat{E}_\ell$ at the initial time and $\hat{E}_k^\dagger$ at the time $t'$ (evolved in the active picture). Let $|n_E\rangle$ be an orthonormal simultaneous eigenbasis of $\hat{H}_E$ and $\hat{\rho}_E(0)$ (which is possible because they commute). In this basis we have;

$$\hat{\rho}_E(0) = \sum_n p_n |n_E\rangle\langle n_E|, \qquad \hat{H}_E = \sum_n E_n |n_E\rangle\langle n_E| \tag{2.3.98}$$

where $p_n$ and $E_n$ are real numbers, $p_n \geq 0$, and $\sum_n p_n = 1$. If we now contract this with some components $\phi_k$ as we would for a matrix summary we find (after inserting complete sets of states):

$$\sum_{k\ell} \Gamma_{k\ell}\phi_k^*\phi_\ell = \sum_{mn} 2\pi\delta(\hbar\omega + E_n - E_m)\left|\sum_k \langle m|\delta\hat{E}_k|n\rangle\phi_k\right|^2 \geq 0. \tag{2.3.99}$$

Thus, $\Gamma_{k\ell}$ forms a nonnegative matrix and its eigenvalues $\lambda_b$ are consequently nonnegative. Define then the jump operators:

$$\hat{L}_b = \sum_k \sqrt{\lambda_b(\omega_k)}(\iota_b(\omega_k))_k \hat{S}_k. \tag{2.3.100}$$

In terms of these, (2.3.95) becomes:

$$i\hbar \frac{d\hat{\rho}_I(t)}{dt} \approx \sum_{\substack{k\ell \\ \omega_k=\omega_\ell}} \pi_{k\ell}(\omega_k)[\hat{S}_k^\dagger \hat{S}_\ell, \hat{\rho}_I] + i\sum_k (\hat{L}_k \hat{\rho}_I(t)\hat{L}_k^\dagger - \mathrm{Sym}(\hat{L}_k^\dagger \hat{L}_k, \hat{\rho}_I(t))). \tag{2.3.101}$$

Now this is really looking like the Lindblad equation. All we have to do to finish it is return to the passive picture. Doing so simply gives:

$$i\hbar \frac{d\hat{\rho}(t)}{dt} \approx [\hat{H}, \hat{\rho}(t)] + i\sum_k (\hat{L}_k \hat{\rho}(t)\hat{L}_k^\dagger - \mathrm{Sym}(\hat{L}_k^\dagger \hat{L}_k, \hat{\rho}(t))) \tag{2.3.102}$$

where we have defined the effective subsystem Hamiltonian as:

$$\hat{H} = \hat{H}_0 + \sum_k \mathrm{Tr}_E(\hat{E}_k \hat{\rho}_E(0))\hat{S}_k + \sum_{\substack{k\ell \\ \omega_k=\omega_\ell}} \pi_{k\ell}(\omega_k)\hat{S}_k^\dagger \hat{S}_\ell. \tag{2.3.103}$$

We have recovered the Lindblad equation through a series of (often reasonable) approximations of a full theory! Because of this, by following the chain of definitions for all the quantities appearing here it is possible to go from a full theory description to this reduced effective description while being able to calculate all of the ingredients of the final equation. This has a significant advantage over just asserting (2.3.53) directly as that description gave no assistance in selecting $\hat{H}$ or the jump operators.We can now also understand why the Lindblad equation might ever be useful. If you have a subsystem of interest weakly coupled to a large

number of environmental degrees of freedom that can be taken as always in equilibrium (essentially due to the thermodynamic limit), then that subsystem will often approximately satisfy the Lindblad equation. If instead you had strong coupling or a nonthermodynamically large number of degrees of freedom, or significant environmental entanglements, the Lindblad equation would not hold. A closed equation may not exist for such a subsystem, and even if an approximate one does, it will typically not be linear.

### Exercises

**Problem 2.7: Most General Two-Dimensional Lindblad System.** Consider a spin-1/2 state evolving under the most general possible Lindblad equation (as might happen for the spin state of a charged single spin-1/2 particle like the electron interacting with electromagnetism at a nonzero temperature). Most general jump operators, full solution, entropy, description of what happens.

(a) Show, similarly to problem 1.62(f), that for some vector $\vec{\xi}(t)$ with $|\vec{\xi}| \leq 1$ and some real valued functions $\Omega_0(t)$, $\Omega_1(t)$, $\Omega_2(t)$, $\Omega_3(t)$ any valid spin-1/2 density matrix and Hamiltonian operator can be written in terms of the Pauli matrices as:

$$\hat{\rho}(t) = \frac{1}{2}(\mathbb{1} + \xi^a(t)\sigma_a), \qquad \hat{H}(t) = \frac{\hbar}{2}(\Omega_0(t)\mathbb{1} + \Omega^a(t)\sigma_a).$$

Show as well that we can take the $\hat{F}_k = \frac{\sigma_k}{\sqrt{2}}$.

(b) Show from the Lindblad equation as expressed in (2.3.46) that the spin polarization $\vec{\xi}(t)$ satisfies:

$$\dot{\xi}^a = \epsilon^a{}_{bc}\Omega^b\xi^c - M^a{}_b\xi^b \qquad \text{where} \qquad M_{k\ell}(t) = \mathrm{Tr}(g)\delta_{kl} - g_{k\ell}.$$

Show that since the only condition on $g_{k\ell}$ is that it be a nonnegative matrix, similarly the only condition on $\Gamma_{k\ell}$ is that it be nonnegative.

(c) Suppose that the jump operators do not depend on time, as would be appropriate for a system experiencing a steady-state bath. Assume $M_{ab}$ is not identically 0. Show that $|\vec{\xi}|$ decreases monotonically with time. Show that $M_{ab}$ has a least eigenvalue $m$ and that $|\vec{\xi}|$ decays in time at least as quickly as $e^{-mt}$. Thus, any nonzero form of dissipation asymptotically results in $\vec{\xi} \to 0$, corresponding to a purely unpolarized density matrix for the system.

## 2.4 Measurement Problem

This section is based on [1,38], and conversations with John P. Ralston. As presented so far, in quantum mechanics the state of the system is a well-defined physically real quantity known as the density matrix which evolves deterministically according to the von Neumann equation and which has well-defined properties given by operator summaries. This seems to be in stark contrast to the usual presentation of quantum mechanics, in which the Born rule and probability are abound. In this section we will explain why I have the view that the Born rule is a derived concept in quantum mechanics that originates from the statistical mechanics of interacting systems. Neumaier [38] calls this the **Thermal Interpretation of Quantum Mechanics** and is the most thoroughly descriptive source I know of on this way of thinking (the name thermal coming from the statistical mechanical origin of the peculiarities of quantum measurement). I also call this way of thinking the Unitary Interpretation of Quantum Mechanics, because the key point in my eyes is that quantum mechanical dynamics is always deterministically unitary, even during measurement. I have encountered several physicists who think this way, and there is historical precedent for this perspective throughout the entire century of quantum mechanics' history. For a basic example, in 1929 Mott [36] showed that once a charged particle has ionized a single atom in a bubble chamber, it will only ionize further particles in a nearly straight line, thus leading to the apparent point-particle-like trajectories in such a chamber using purely deterministic wave-mechanics and in 2021 Schonfeld [54] showed how the fact Born rule distribution of the first such ionization event can be derived from statistical mechanics.

Consider a microscopic quantum system of interest for observation which we will call the System (S). Because it is microscopic, we cannot access its properties directly. Instead, in order to measure anything about it we need to have it interact with some macroscopic measurement device we call the Detector (D) so that some property of the System becomes correlated with a macroscopic property of the Detector. Then, because that property is macroscopic, it is available to us to be directly observed. The detector will include not only the meter that we read off but many many more environmental degrees of freedom (in say the double-slit experiment, the detector is the entire physical apparatus of the experiment including the air and

the space the waves travel through while the meter is just the screen or any equipment capable of measuring at the slits). The combined Hilbert space of the System+Detector system we call $\mathbb{H} = \mathbb{H}_S \otimes \mathbb{H}_D$. In any single experimental trial, at time $t = 0$ before the System and Detector have interacted, the System and Detector are totally uncorrelated and so:

$$\hat{\rho}_0 = \hat{\rho}_S \otimes \hat{\rho}_D. \tag{2.4.1}$$

The System and Detector will interact through some total Hamiltonian operator $\hat{H}$ so that $\hat{\rho}(t)$ of the combined system satisfies the von Neumann equation:

$$\frac{d\hat{\rho}}{dt} + \frac{1}{i\hbar}[\hat{\rho}, \hat{H}] = 0, \qquad \hat{\rho}(0) = \hat{\rho}_0. \tag{2.4.2}$$

During the interaction the combined system is isolated and so $\hat{H}$ has no explicit time dependence. Consequently, the time evolution is solved by:

$$\hat{\rho}(t) = \hat{U}(t)\hat{\rho}_0\hat{U}^\dagger(t), \qquad \hat{U}(t) = e^{-\frac{i}{\hbar}\hat{H}t}. \tag{2.4.3}$$

With $|n_S\rangle$ an orthonormal basis of states of $\mathbb{H}_S$ and $|n_D\rangle$ an orthonormal basis of states of $\mathbb{H}_D$, we can define the operators $\hat{U}_{mn}(t)$ which are operators on $\mathbb{H}_D$ by:

$$\hat{U}_{mn}(t) = \sum_{m'n'} |m'_D\rangle\langle m_S m'_D|\hat{U}(t)|n_S n'_D\rangle\langle n'_D|. \tag{2.4.4}$$

Automatically then:

$$\hat{U}(t) = \sum_{mn} |m_S\rangle\langle n_S| \otimes \hat{U}_{mn}(t). \tag{2.4.5}$$

Suppose we wish to measure some real-valued property $\langle\hat{s}\rangle_S$ of the System. (Because that property is real-valued, $\hat{s}$ is self-adjoint.) By definition we only actually have direct access to some part of our detector. So, there will be some generator $\langle\hat{A}\rangle_D$ on the Detector's phase space, for some associated operator $\hat{A}$ on the Detector's Hilbert space, which we will see experimentally (the $D$ subscript indicates the operator summary is taken using the Detector's density matrix). We would like to design the measurement interaction so that $\langle\hat{A}\rangle_D$ tells us as much as possible about $\langle\hat{s}\rangle_S$. The corresponding operator $\hat{X}$ to $A$ on the full Hilbert space is:

$$\hat{X} = \mathbb{1}_S \otimes \hat{A} \tag{2.4.6}$$

and the property given by $\hat{X}$ on the space of the combined system is:

$$\langle\hat{X}\rangle(t) = \mathrm{Tr}(\hat{\rho}(t)\hat{X}) = \mathrm{Tr}(\hat{U}(t)\hat{\rho}_0\hat{U}^\dagger(t)\hat{X}) = \mathrm{Tr}(\hat{\rho}_0\hat{U}^\dagger(t)\hat{X}\hat{U}(t)) \tag{2.4.7}$$

$$= \sum_{mm'nn'} \mathrm{Tr}\left(\hat{\rho}_0(|n'_S\rangle\langle m'_S| \otimes \hat{U}^\dagger_{m'n'}(t))(\mathbb{1}_S \otimes \hat{A})(|m_S\rangle\langle n_S| \otimes \hat{U}_{mn}(t))\right) \tag{2.4.8}$$

$$= \sum_{lmn} \mathrm{Tr}\left((\hat{\rho}_S \otimes \hat{\rho}_D)(|m_S\rangle\langle n_S| \otimes (\hat{U}^\dagger_{lm}(t)\hat{A}\hat{U}_{ln}(t)))\right) \tag{2.4.9}$$

$$= \mathrm{Tr}_S\left(\hat{\rho}_S \sum_{klmn} |m_S\rangle\langle k_D|\hat{\rho}_D\hat{U}^\dagger_{lm}(t)\hat{A}\hat{U}_{ln}(t)|k_D\rangle\langle n_S|\right) \tag{2.4.10}$$

by standard steps of calculation. If we now define the operator $\hat{A}_S(t)$ by:

$$\hat{A}_S(t) = \sum_{lmn} |m_S\rangle\mathrm{Tr}_D(\hat{\rho}_D\hat{U}^\dagger_{lm}(t)\hat{A}\hat{U}_{ln}(t))\langle n_S| \implies \langle\hat{X}\rangle(t) = \langle\hat{A}\rangle_D(t) = \langle\hat{A}_S(t)\rangle_S. \tag{2.4.11}$$

Therefore, if we measure the value of $\langle\hat{A}\rangle_D(t)$ by looking at the Detector, we are equivalently observing the value of $\langle\hat{A}_S(t)\rangle_S$ and thus able to learn something about the state of the System. Thus, in a well designed experiment, $\langle\hat{A}_S\rangle_S$ will resemble $\langle\hat{s}\rangle_S$ as closely as possible.

We must note something extremely important about the definition of $\hat{A}_S(t)$; it depends on $\hat{\rho}_D$! At the very least, this means it cannot actually identically be $\hat{s}$. Because we are able to measure the property $\langle\hat{A}\rangle_D$ directly, it must be macroscopic. Thus, the Detector itself must be macroscopic and so consist of a large number of microscopic degrees of freedom. Because the Detector is macroscopic, we are unable to control its microscopic state exactly and so we must treat the detector's state using statistical mechanics. Consequently, **in different experimental trials with the same macroscopic experimental setup the microscopic state of the Detector will be different**. The full initial microscopic state of the Detector is contained in $\hat{\rho}_D$ and thus we must treat $\hat{\rho}_D$ as a random variable distributed according to statistical mechanics in whatever

distribution appropriately describes our experimental setup procedure.

Because for different experimental trials $\hat{\rho}_D$ is in fact different, the resulting property $\langle \hat{A}_S \rangle_S$ is different in different experimental trials, with distribution determined by the distribution of $\hat{\rho}_D$ and the precise coupling between the System and Detector. Thus, **even when the system of interest is prepared in the exact same microscopic state in each experimental trial and even though the dynamics itself is fundamentally deterministic, because we are only able to access that state through a macroscopic interaction, the results of our experiments are apparently random. This apparent randomness is exactly like in statistical mechanics**. The only remaining questions are then what distribution may be expected for $\langle \hat{A}_S \rangle_S$ in an idealized measurement scheme and how can we set the interaction between S and D so that $\langle \hat{A}_S \rangle_S$ approximates $\langle \hat{s} \rangle_S$.

In a measurement the System will interact with the Detector for some period of time according to a Hamiltonian operator $\hat{H}$ which we engineer to correlate some part of the final state of the Detector as strongly as possible with a property of the incoming System. In principle, at all times the Hamiltonian operator can as always be written in the form:

$$\hat{H} = \hat{H}_S \otimes \mathbb{1}_D + \mathbb{1}_S \otimes \hat{H}_D + \hat{H}_{\text{int}} \tag{2.4.12}$$

where $\hat{H}_S$ and $\hat{H}_D$ are the free Hamiltonian operators of the System and Detector respectively and $\hat{H}_{\text{int}}$ contains their interactions. The interaction term $\hat{H}_{\text{int}}$ is all that is under our control in designing the experimental apparatus. Before the System reaches the Detector, the interaction between the two is negligible and so the interaction term can be neglected. Once the System is sufficiently close to the Detector, we assume in an ideal measurement that the interaction between them is very strong (so the system can properly drive the Detector during that time). Thus, during the interaction we approximate:

$$\hat{H} \approx \hat{H}_{\text{int}} \tag{2.4.13}$$

Going forward in our present analysis we will simply neglect the contributions from $\hat{H}_S, \hat{H}_D$.

Let us now try to determine the form of $\hat{H}$ most suited to an "ideal measurement". In an ideal measurement of some real-valued generator $\langle \hat{s} \rangle_S$ of the System (and thus self-adjoint $\hat{s}$), we require that the behavior of the operator $\hat{s}$ on the System is perturbed as little as possible and that there is no finer structure than what is provided by $\hat{s}$ which is also disturbed as little as possible. To make this more precise it will be helpful to introduce a little notation. Let $s_n$ be the eigenvalues of $\hat{s}$ and let $\hat{\pi}_n$ be its eigenspace projectors so that for all $|\Psi\rangle \in \mathbb{H}_S$:

$$\hat{\pi}_n|\Psi\rangle = |\Psi\rangle \iff \hat{s}|\Psi\rangle = s_n|\Psi\rangle \tag{2.4.14}$$

$$\implies \hat{\pi}_m\hat{\pi}_n = \delta_{mn}\hat{\pi}_n, \qquad \sum_n \hat{\pi}_n = \mathbb{1}_S, \qquad \hat{s} = \sum_n s_n\hat{\pi}_n. \tag{2.4.15}$$

It will be convenient to have an explicit orthonormal eigenbasis for $\hat{s}$ on $\mathbb{H}_S$. Let $|n, a_S\rangle$ be an orthonormal basis of $\mathbb{H}_S$ with labels $n$, $a$ where $n$ labels the eigenvalue $s_n$ of this vector under $\hat{s}$ and $a$ is some label (with in general range dependent on $n$) that is necessary because the eigenvalue $s_n$ may be degenerate:

$$\hat{s}|n, a_S\rangle = s_n|n, a_S\rangle, \qquad \langle n, a_S|n', a'_S\rangle = \delta_{nn'}\delta_{aa'} \implies \hat{\pi}_n = \sum_a |n, a_S\rangle\langle n, a_S|. \tag{2.4.16}$$

We can now precisely state our conditions on $\hat{H}$ to "ideally measure" $\langle \hat{s} \rangle_S$. Suppose that the System is initially in a state of 0 operator spread with respect to $\hat{s}$: $\text{Spr}_S(\hat{s}) = 0$. Then, for some $n$, $\hat{\pi}_n\hat{\rho}_S\hat{\pi}_n = \hat{\rho}_S$ and $\langle \hat{s} \rangle_S = s_n$. In order for $\langle \hat{s} \rangle_S$ to be disturbed as little as possible, we require that it remains in a state of 0 operator spread with respect to $\hat{s}$ with the same value of $\langle \hat{s} \rangle_S = s_n$. Further, if the eigenspace projected into by $\hat{\pi}_n$ has dimension greater than 1, our experiment should not be able to resolve any of this degeneracy because if it could then it would not be $\langle \hat{s} \rangle_S$ that we are measuring, it would be some other generator $\langle \hat{s}' \rangle_S$ where $\hat{s}'$ is less degenerate than $\hat{s}$. Thus, if for some $n$, $\hat{\pi}_n\hat{\rho}_S\hat{\pi}_n = \hat{\rho}_S$, then not only should the System stay in a state of $\hat{s}$-spread 0, it should stay in the *same* state of $\hat{s}$-spread 0. Recalling that $\text{Tr}_D(\hat{\rho})$ is the reduced density matrix of the system at all times, this means (using (2.1.118)) that:

$$\hat{\pi}_n\hat{\rho}_S\hat{\pi}_n = \hat{\rho}_S \implies \frac{d}{dt}(\text{Tr}_D(\hat{\rho})) = 0 \implies \text{Tr}_D([\hat{\rho}, \hat{H}]) = 0. \tag{2.4.17}$$

To see specifically what this requires of $\hat{H}$, it is useful to decompose it into a sum of tensor products along the lines of how we decomposed $\hat{U}$. Define the operators $\hat{H}_{m'n'}$ by:

$$\hat{H}_{m'n'} = \sum_{mnab} |m, a_S\rangle\langle m, a_S m'_D|\hat{H}|n, b_S n'_D\rangle\langle n, b_S|. \tag{2.4.18}$$

In terms of these $\hat{H}_{m'n'}$, we can express $\hat{H}$ as:

$$\hat{H} = \sum_{mm'nn'ab} |m, a_S m'_D\rangle\langle m, a_S m'_D|\hat{H}|n, b_S n'_D\rangle\langle n, b_S n'_D| = \sum_{m'n'} \hat{H}_{m'n'} \otimes |m'_D\rangle\langle n'_D|. \tag{2.4.19}$$

Letting $\rho_{Dn'm'} = \langle n'_D|\hat{\rho}_D|m'_D\rangle$ be the matrix elements of $\hat{\rho}_D$, we can now compute:

$$\mathrm{Tr}_D(\hat{\rho}\hat{H}) = \sum_{m'n'} \rho_{Dn'm'}\hat{\rho}_S\hat{H}_{m'n'}, \qquad \mathrm{Tr}_D(\hat{H}\hat{\rho}) = \sum_{m'n'} \rho_{Dn'm'}\hat{H}_{m'n'}\hat{\rho}_S. \tag{2.4.20}$$

Therefore (2.4.17) becomes:

$$\hat{\pi}_n\hat{\rho}_S\hat{\pi}_n = \hat{\rho}_S \implies \sum_{m'n'} \rho_{Dn'm'}[\hat{\rho}_S, \hat{H}_{m'n'}] = 0. \tag{2.4.21}$$

This must hold for all valid density matrices $\hat{\rho}_D$, and so this requires each term in the sum to vanish independently[2]:

$$\hat{\pi}_n\hat{\rho}_S\hat{\pi}_n = \hat{\rho}_S \implies [\hat{\rho}_S, \hat{H}_{m'n'}] = 0. \tag{2.4.22}$$

Consider a valid density matrix $\hat{\rho}_S$ which satisfies no particular conditions (is not necessarily a state of 0 $\hat{s}$-spread). Then, we can define another valid density matrix $\hat{\rho}_S^{(n)}$ by:

$$\hat{\rho}_S^{(n)} = \frac{\hat{\pi}_n\hat{\rho}_S\hat{\pi}_n}{\mathrm{Tr}(\hat{\pi}_n\hat{\rho}_S)}. \tag{2.4.23}$$

The denominator ensures that $\hat{\rho}_S^{(n)}$ has trace 1 and the projectors ensure that $\hat{\rho}_S^{(n)}$ is a state of 0 $\hat{s}$-spread in which $\langle\hat{s}\rangle_S = s_n$. Thus, for all $\hat{\rho}_S$, by (2.4.22) using $\hat{\rho}_S^{(n)}$ in place of $\hat{\rho}_S$ in that equation:

$$[\hat{\rho}_S^{(n)}, \hat{H}_{m'n'}] = 0 \implies \hat{\pi}_n\hat{\rho}_S\hat{\pi}_n\hat{H}_{m'n'} = \hat{H}_{m'n'}\hat{\pi}_n\hat{\rho}_S\hat{\pi}_n. \tag{2.4.24}$$

This last condition now holds without restriction on $\hat{\rho}_S$.

Let us put this last condition in components and use (2.4.16):

$$\langle m, a_S|\hat{\pi}_k\hat{\rho}_S\hat{\pi}_k\hat{H}_{m'n'}|n, b_S\rangle = \langle m, a_S|\hat{H}_{m'n'}\hat{\pi}_k\hat{\rho}_S\hat{\pi}_k|n, b_S\rangle \tag{2.4.25}$$

$$\implies \sum_c \delta_{mk}\langle m, a_S|\hat{\rho}_S|k, c_S\rangle\langle k, c_S|\hat{H}_{m'n'}|n, b_S\rangle = \sum_c \delta_{nk}\langle m, a_S|\hat{H}_{m'n'}|k, c_S\rangle\langle k, c_S|\hat{\rho}_S|n, b_S\rangle. \tag{2.4.26}$$

We can now differentiate both sides with respect to the matrix element $\langle m, a_S|\hat{\rho}_S|k, c_S\rangle$ as those matrix elements may be freely specified[3]:

$$\implies \delta_{mk}\langle k, c_S|\hat{H}_{m'n'}|n, b_S\rangle = \sum_d \delta_{nk}\langle m, a_S|\hat{H}_{m'n'}|k, d_S\rangle\delta_{mk}\delta_{ad}\delta_{kn}\delta_{cb} \tag{2.4.27}$$

$$\implies \langle m, a_S|\hat{H}_{m'n'}|n, b_S\rangle = \delta_{mn}\delta_{ab}\langle n, c_S|\hat{H}_{m'n'}|n, c_S\rangle \tag{2.4.28}$$

where to get to this last line we summed the previous line over all $m$ and then relabeled various indices. Note that the right hand side of the last expression looks like it depends on $c$, but the left hand side explicitly does not. Therefore, the right hand side does not either and so there exists numbers $h_{nm'n'}$ so that:

$$h_{nm'n'} = \langle n, c_S|\hat{H}_{m'n'}|n, c_S\rangle \implies \langle m, a_S|\hat{H}_{m'n'}|n, b_S\rangle = \delta_{mn}\delta_{ab}h_{nm'n'}. \tag{2.4.29}$$

From this we explicitly have:

$$\hat{H}_{m'n'} = \sum_{mnab} = |m, a_S\rangle\langle m, a_S|\hat{H}_{m'n'}|n, b_S\rangle\langle n, b_S| = \sum_n h_{nm'n'}\hat{\pi}_n. \tag{2.4.30}$$

Define then the operators $\hat{h}_n$ on $\mathbb{H}_D$ by:

$$\hat{h}_n = \sum_{m'n'} |m'_D\rangle h_{nm'n'}\langle n'_D|. \tag{2.4.31}$$

Returning these to $\hat{H}$, we find:

$$\hat{H} = \sum_n \hat{\pi}_n \otimes \hat{h}_n. \tag{2.4.32}$$

Thus, this is the general form required of the Hamiltonian operator for when our System interacts with the

[2] A careful reader may worry that this isn't quite true, because a density matrix must still satisfy some basics conditions (self-adjointness, positivity, trace-1). However, a short component calculation reveals that if $\mathrm{Tr}(\hat{\rho}\hat{\mathcal{O}}) = 0$ for some operator $\hat{\mathcal{O}}$ and all density matrices $\hat{\rho}$, then in fact $\hat{\mathcal{O}} = 0$.

[3] This differentiation argument is legal by the same logic as the previous footnote.

Detector in an ideal measurement. Because $\hat{H}$ is necessarily self-adjoint, and the projectors $\hat{\pi}_n$ are all linearly independent and self-adjoint, it automatically follows that each $\hat{h}_n$ is self-adjoint. Often in discussions of measurement as a dynamical process, you may see the suggested interaction Hamiltonian $\hat{H} = \hat{s} \otimes \hat{d}$ for some operator $\hat{d}$ on $\mathbb{H}_D$. By rewriting $\hat{s} = \sum_n s_n \hat{\pi}_n$, you can see that $\hat{s} \otimes \hat{d}$ is actually of the form (2.4.32) with $\hat{h}_n = s_n \hat{d}$ and so this is a legal special case of our formula.

Using the form (2.4.32), we can explicitly compute the time evolution operator $\hat{U}(t)$. For a natural number $N$ (because $\hat{\pi}_m \hat{\pi}_n = \delta_{mn} \hat{\pi}_n$):

$$\hat{H}(\hat{\pi}_n \otimes \hat{h}_n^N) = \hat{\pi}_n \otimes \hat{h}_n^{N+1} \implies \hat{H}^N(\hat{\pi}_n \otimes \mathbb{1}_D) = \hat{\pi}_n \otimes \hat{h}_n^N \tag{2.4.33}$$

by $N$ fold application of the previous fact. Summing over all $n$ reveals:

$$\hat{H}^N = \sum_n \hat{\pi}_n \otimes \hat{h}_n^N. \tag{2.4.34}$$

Thus, computing the exponential for $\hat{U}(t)$ we have:

$$\hat{U}(t) = \sum_n \hat{\pi}_n \otimes e^{-\frac{i}{\hbar} t \hat{h}_n}. \tag{2.4.35}$$

Using our expression for $\hat{U}(t)$, we can now repeat the calculation leading to (2.4.11) to find the appropriate expression for $\hat{A}_S$ in the case of an ideal measurement:

$$\langle \hat{X} \rangle(t) = \mathrm{Tr}(\hat{\rho}_0 \hat{U}^\dagger \hat{X} \hat{U}) \tag{2.4.36}$$

$$= \sum_{mn} \mathrm{Tr}\left( (\hat{\rho}_S \otimes \hat{\rho}_D) \left( \hat{\pi}_m \otimes e^{\frac{i}{\hbar} t \hat{h}_m} \right) (\mathbb{1}_S \otimes \hat{A}) \left( \hat{\pi}_n \otimes e^{-\frac{i}{\hbar} t \hat{h}_n} \right) \right) \tag{2.4.37}$$

$$= \sum_{mn} \mathrm{Tr}\left( (\hat{\rho}_S \hat{\pi}_m \hat{\pi}_n) \otimes \left( \hat{\rho}_D e^{\frac{i}{\hbar} t \hat{h}_m} \hat{A} e^{-\frac{i}{\hbar} t \hat{h}_n} \right) \right) = \sum_n \mathrm{Tr}_S(\hat{\rho}_S \hat{\pi}_n) \mathrm{Tr}_D\left( \hat{\rho}_D e^{\frac{i}{\hbar} t \hat{h}_n} \hat{A} e^{-\frac{i}{\hbar} t \hat{h}_n} \right). \tag{2.4.38}$$

So, in an ideal measurement, we identify:

$$\hat{A}_S(t) = \sum_n \hat{\pi}_n \mathrm{Tr}_D\left( \hat{\rho}_D e^{\frac{i}{\hbar} t \hat{h}_n} \hat{A} e^{-\frac{i}{\hbar} t \hat{h}_n} \right) \implies \langle \hat{X} \rangle(t) = \langle \hat{A}_S(t) \rangle_S. \tag{2.4.39}$$

The time dependent quantity $\langle \hat{X} \rangle(t)$ is **NOT** fully accessible in experiment, again due to the macroscopic nature of the Detector. The typical time scales on which quantum systems evolve are extremely microscopic compared to experimental resolution and so time-averaging and relaxation to equilibrium are inherent to how we as macroscopic beings access individual quantum systems. As well, we cannot access the exact real number value of the time-averaged relaxed version of $\langle \hat{X} \rangle$. Real data has finite precision; it is necessarily binned into bins of finite width. Each bin then defines an equivalence class of values of $\langle \hat{X} \rangle$ that are the considered the same with respect to the measurement. This is also inherent from the macroscopic nature of the Detector, because a bin essentially amounts to a "macrostate" with all of the different $\hat{\rho}_D$ that for a given $\hat{\rho}_S$ end up with $\langle \hat{X} \rangle$ in the same bin amounting to the very many "microstates" that we cannot distinguish between.

For each different trial in an experiment, $\hat{\rho}_D$ will be different due to the Detector having many microstates consistent with our macroscopic preparation. Because of this, we let $\omega$ label the microstate of the detector (in probability terms, be a point in the sample space) and $\hat{\rho}_D(\omega)$ be the resulting state of the detector giving that microstate (thus $\hat{\rho}_D(\omega)$ is a random variable in the standard sense of classical probability). The operator $\hat{A}_S$ depends on the detector through the time-dependent, $\omega$ dependent combinations $\mathrm{Tr}_D\left( \hat{\rho}_D e^{\frac{i}{\hbar} t \hat{h}_n} \hat{A} e^{-\frac{i}{\hbar} t \hat{h}_n} \right)$. After time-averaging/relaxation to equilibrium this quantity will approach *some* stable state which may depend on $\omega$ and $n$, because the original does. Thus, letting $\mathbb{T}$ indicate time averaging as usual, for some real numbers $A_n(\omega)$ we have:

$$\mathbb{T}\left( \mathrm{Tr}_D\left( \hat{\rho}_D(\omega) e^{\frac{i}{\hbar} t \hat{h}_n} \hat{A} e^{-\frac{i}{\hbar} t \hat{h}_n} \right) \right) = A_n(\omega). \tag{2.4.40}$$

The resulting time-averaged and relaxed value of $\langle \hat{X} \rangle$ is then also a random variable dependent on $\omega$:

$$\mathbb{T}(\langle \hat{X} \rangle) = \sum_n \mathrm{Tr}_S(\hat{\rho}_S \hat{\pi}_n) A_n(\omega). \tag{2.4.41}$$

When relating this time-relaxed value to the reading read out on a macroscopic detector, there is some sort of binning procedure which we represent mathematically by a function $\mathrm{Bin}(\cdot)$ which takes a real number and spits out the corresponding bin value given to that real number. Thus, the final experimentally measured

value $X$ in a single trial is given by:

$$X(\omega) = \text{Bin}(\mathbb{T}(\langle \hat{X} \rangle)) = \text{Bin}\left( \sum_n \text{Tr}_S(\hat{\rho}_S \hat{\pi}_n) A_n(\omega) \right). \tag{2.4.42}$$

Because real data is discretely binned, the binning function cannot possibly be linear over the real numbers.

What properties does the binning function have? The binning function is determined by how the experiment is calibrated, and that calibration is necessary to manage the imperfectly prepared nature of each different experimental trial for $\hat{\rho}_D(\omega)$. There are two criteria used to calibrate the binning of an ideal quantum experiment. Suppose the initial state of the System $\hat{\rho}_S$ is such that $\hat{s}$ has operator spread zero in that state: $\text{Spr}(\hat{s})_S = 0$. In such a state, $\hat{s}$ acts purely as $s_n \mathbb{1}$ and so unambiguously should be assigned the value $s_n$. Thus, in a perfectly calibrated measurement, the binning procedure would always assign the value $s_n$ to the measurement of such $\langle \hat{s} \rangle_S$:

$$\hat{\pi}_n \hat{\rho}_S \hat{\pi}_n = \hat{\rho}_S \implies X(\omega) = s_n \tag{2.4.43}$$

independent of $\omega$. This criterion essentially the defines the bins to which $X(\omega)$ may be assigned. Because each eigenspace projected onto by some $\hat{\pi}_n$ is intended to be distinguishable in our experiment (because that is what we mean by being able to measure $\langle \hat{s} \rangle_S$, and not the summary of some less refined operator), we must have a separate bin for each eigenspace. Because states within each eigenspace are intended to be indistinguishable (because that is what we mean by being able to measure $\langle \hat{s} \rangle_S$, and not the summary of some more refined operator), we must not have more than one bin for each eigenspace. Therefore, for a general $\hat{\rho}_S$, $X(\omega)$ will be assigned one of the values $s_n$.

Because $X(\omega)$ is intended to measure the value of $\langle \hat{s} \rangle_S$, the second criterion on a good binning procedure is that $X(\omega)$ is an unbiased estimator of $\langle \hat{s} \rangle_S$. That is, its statistical expectation should match $\langle \hat{s} \rangle_S$:

$$\mathbb{E}[X] = \langle \hat{s} \rangle_S. \tag{2.4.44}$$

Let $p_n$ be the probability that $X$ is sorted into the bin with value $s_n$ for a given System state $\hat{\rho}_S$. Expanding both sides of this equation (using the definition of statistical expectation and (2.4.15), we have:

$$\sum_n s_n p_n = \sum_n s_n \text{Tr}(\hat{\rho}_S \hat{\pi}_n). \tag{2.4.45}$$

The eigenvalues $s_n$ of $\hat{s}$ never entered the dynamics of the problem, only the eigenspace projectors $\hat{\pi}_n$. Therefore, as long as the $s_n$ are sufficiently distinct from each other for the bins surrounded them to also remain distinct, the precise values of $s_n$ are somewhat arbitrary. Essentially, a measurement of the summary of $\hat{s} = \sum_n s_n \hat{\pi}_n$ could equally well be called a measurement of the summary of $\hat{s}' = \sum_n s'_n \hat{\pi}_n$ if the $s'_n$ values are still well separated. In that case, the associated random variable would be $X'(\omega)$ and assigned this alternate bin values. The eigenvalues $s_n$ essentially amount to what we name the bins and are thus arbitrary. So, if this equation holds at a given set of values $s_n$, then it holds in a small neighborhood around that point. This allows us to differentiate this equation with respect to $s_n$ and so identify that in a perfectly calibrated binning procedure, the binning *must* satisfy:

$$\boxed{p_n = \text{Tr}(\hat{\rho}_S \hat{\pi}_n).} \tag{2.4.46}$$

This is precisely the Born rule! So, we see that in an ideal measurement process where the measurement outcomes are binned so that a state of 0-spread is identified with a consistent value and so that the measurement outcome is an unbiased estimator of the true valued being measured, the Born rule is automatically emergent.

The Born rule as written here is not the form most readily family to beginning quantum mechanics students, who only usually know it through a special case. Consider the special case that the eigenspace projected into by $\hat{\pi}_n$ has dimension 1. Then, we can identify a unit vector $|\phi\rangle$ so that $\hat{\pi}_n = |\phi\rangle\langle\phi|$. Further, consider the special case that $\hat{\rho}_S$ is a pure state. Then, we can identify a state vector $|\psi\rangle$ so that $\hat{\rho}_S = |\psi\rangle\langle\psi|$. Simplifying the Born rule as above we find that in this exceptionally special case, the probability that the measured value lands in bin $n$ becomes $p_n = |\langle\phi|\psi\rangle|^2$, which is the more familiar form to those new to quantum mechanics. We have hopefully made it clear in this section that the magnitude squared inner product $|\langle\phi|\psi\rangle|^2$ for vectors $|\phi\rangle$, $|\psi\rangle$ a priori has no real interpretative meaning. It becomes a meaningful quantity only in the extremely special case of the more general Born rule probability $\text{Tr}(\hat{\rho}_S \hat{\pi}_n)$ which itself is only meaningful as a probability in the context of a perfectly calibrated ideal measurement interaction. On its own outside of this specific dynamical interaction context, $\text{Tr}(\hat{\rho}_S \hat{\pi}_n)$ is just a number.

Let us now look at the final state of the System after it has been measured. The time-dependent state

$\hat{\rho}_S(t)$ is given as usual by the partial trace over everything else, in this case just the Detector:

$$\hat{\rho}_S(t) = \mathrm{Tr}_D(\hat{\rho}(t)) = \mathrm{Tr}_D(\hat{U}(t)\hat{\rho}_0\hat{U}^\dagger(t)). \tag{2.4.47}$$

Using the expression we found for $\hat{U}(t)$ in (2.4.35):

$$\langle m, a_S n'_D|\hat{U} = \sum_n \langle m, a_S n'_D| \left(\hat{\pi}_n \otimes e^{-\frac{i}{\hbar}t\hat{h}_n}\right) = \langle m, a_S| \otimes \left(\langle n'_D|e^{-\frac{i}{\hbar}t\hat{h}_m}\right). \tag{2.4.48}$$

Then:

$$\hat{\rho}_S(t) = \sum_{mnn'ab} |m, a_S\rangle\langle m, a_S n'_D|\hat{U}\hat{\rho}_d\hat{U}^\dagger|n, b_S n'_D\rangle\langle n, b_S| = \sum_{mn} \hat{\pi}_m\hat{\rho}_S\hat{\pi}_n \mathrm{Tr}_D\left(e^{-\frac{i}{\hbar}t\hat{h}_m}\hat{\rho}_D e^{\frac{i}{\hbar}t\hat{h}_n}\right). \tag{2.4.49}$$

Because $\hat{\rho}_D$ appearing on the right hand side is a random variable dependent on the experimental trial, the final state of the system on the left hand side is as well. In the long run and after time averaging the final system state will necessarily relax to some stationary state of the interaction. In individual experimental trials, the initial detector state will be effectively random, but it is possible to assign an averaged detector state $\hat{\bar{\rho}}_D$ according to:

$$\hat{\bar{\rho}}_D = \mathbb{E}\hat{\rho}_D. \tag{2.4.50}$$

Then, the average final state of the system is given by:

$$\mathbb{E}\hat{\rho}_S(t) = \sum_{mn} \hat{\pi}_m\hat{\rho}_S\hat{\pi}_n \mathrm{Tr}_D\left(e^{-\frac{i}{\hbar}t\hat{h}_m}\hat{\bar{\rho}}_D e^{\frac{i}{\hbar}t\hat{h}_n}\right). \tag{2.4.51}$$

In order for $\hat{h}_m$ and $\hat{h}_n$ for $m \neq n$ to drive the Detector toward different equilibria, they must be quite different operators. Consequently, for $m \neq n$, the operator $e^{\frac{i}{\hbar}t\hat{h}_n}e^{-\frac{i}{\hbar}t\hat{h}_m}$ typically oscillates very quickly in time. In individual preparations it is possible for the detector state to be supported on a part of the Hilbert space so that the combination $e^{-\frac{i}{\hbar}t\hat{h}_m}$ acts essentially trivially for some values of $m, n$. However, we assume that our average detector preparation is not biased toward the behavior of any particular such subspace and so under reasonable time averaging for a wide class of systems, this combination time-averages to 0. This is yet again the familiar **decoherence approximation**. So, in the long run (using cyclicity of the trace), when decoherence is valid we have:

$$\mathbb{T}\left(\mathrm{Tr}_D\left(e^{-\frac{i}{\hbar}t\hat{h}_m}\hat{\bar{\rho}}_D e^{\frac{i}{\hbar}t\hat{h}_n}\right)\right) = \delta_{mn}\mathbb{T}\left(\mathrm{Tr}_D\left(\hat{\bar{\rho}}_D e^{\frac{i}{\hbar}t\hat{h}_n}e^{-\frac{i}{\hbar}t\hat{h}_n}\right)\right) = \delta_{mn}. \tag{2.4.52}$$

Thus the ensemble-averaged time-averaged and relaxed final state $\hat{\bar{\rho}}_S^{(f)}$ of the System is given by:

$$\hat{\bar{\rho}}_S^{(f)} = \mathbb{T}\mathbb{E}[\hat{\rho}_S(t)] = \sum_n \hat{\pi}_n\hat{\rho}_S\hat{\pi}_n = \sum_n p_n\hat{\rho}_S^{(n)} \tag{2.4.53}$$

using the Born rule formula (2.4.46). This ensemble-averaged final state is exactly what you would get from assigning final states $\hat{\rho}_S^{(n)}$ each with probability $p_n$ to the system of interest. That is, it is exactly consistent with the collapse rule. However, any non-pure density matrix can be decomposed infinitely many ways as a convex sum, and so this alone is not enough to actually conclude that the system follows the collapse rule trial by trial.

It is also quick to derive (2.4.53) as the averaged final state of the system using the Lindblad equation, since we can treat all of the microscopic degrees of freedom in the detector as an environment to be integrated out precisely as in deriving (2.3.95). However, because the Lindblad equation assumes from the beginning the ensemble behavior of the environmental degrees of freedom rather than their state as being differently realized in every experimental trial, the Lindblad approach has averaged too early and so is unable to derive (2.4.49), which captures precisely how the System's state in each individual trial depends on the initial microscopic state of the detector. This point is often not appreciated. In order to derive the outcomes of individual trials, you are not allowed to do any ensemble averaging. You have to leave the initial conditions as a random variable all the way until the end.

While (2.4.49) formally gives the final state of the system in each trial, deriving the collapse rule from it is much less straightforward than the Born rule. Honestly, I only know of convincing derivations in specific cases. For example, for a spin-1/2 system coupled to a bath at low temperature, it has been thoroughly shown by [7] or for particle tracks in a bubble chamber by [36,54]. It isn't too surprising that the collapse rule is more difficult, because it is also less precisely and less widely observed. There are increasing explorations of "weak measurements" in which the system state is seen to not have been collapsed, only perturbed, during the measurement. In my view, the so-called "measurement problem" in quantum mechanics is basically

just the problem of trying to find a maximally precise and general derivation of when you should expect the collapse rule to hold. It ought to come from appropriate treatment of the initial detector microstate as a random variable and thorough enough understanding of how (2.4.49) should simplify when the detector is thermodynamically large, but you aren't allowed to assume in an individual trial that each part of the detector has the equilibrium density matrix. Solving this problem (and actually convincing people of your solution, which is even harder) would make you very famous.

In the cases where the collapse rule does have a convincing derivation, the style is generally quite similar. The heuristic argument is as follows. The detector meter $X$ is designed to correlate as strongly as possible to the System's quantity $\langle \hat{s} \rangle_S$. Requiring this macroscopic variable to be sensitive to the state of a microscopic variable essentially forces the system to be chaotic (in the sense of a small deviation in initial conditions creating a large effect on the final state). This makes the detector meter also meaningfully sensitive to its internal microscopic variables (since they are also coupled to the system of interest). The necessary binning of the detector meter is accomplished mathematically by detector design providing an effective potential with local minima at the bin values, so that they are the stable equilibria. The initial microstate of the detector chaotically implies which equilibria is reached in each experimental trial, and in an unbiased detector the distribution must follow the Born rule, as we derived. The bins corresponding to this local minima precisely occur when the system sits in the eigenstate that defined the bin, and are precisely stable so that such eigenstates are undisturbed. The strong coupling of the system to the detector meter drives the system into the eigenstate corresponding to the minimum the detector meter is sitting in, giving the collapse rule. Essentially, if you can argue that the system of interest is dynamically driven to an eigenstate of $\hat{s}$, then due to (2.4.53) holding for the average, it must realize each eigenstate with the Born rule probability.

In summary, both of the basic features of quantum measurements typically described as mysterious, the Born rule and the collapse rule, can be understood as emergent from a reasonable interacting model of a macroscopic detector with a microscopic system of interest. The arguments presented here are extremely generic, much like the arguments we gave for how thermodynamic equilibrium is reached by many systems in statistical mechanics. In statistical mechanics there have been many more detailed investigations into the precise way in which specific systems reach thermodynamic equilibrium. The same is true for quantum measurement; there is a meaningful body of literature, with meaningful developments still being made, dedicated to understanding the dynamical emergence of the properties of quantum measurements for particular systems. For more reading on this and closely related subjects, consult [1, 7, 9, 10, 38, 39].

## Exercises

**Problem 2.8: Neumaier Qubit.** This question is based on [38]. Consider the measurement of a quantum system with a two dimensional Hilbert space $\mathbb{H}_S$. Let $|0\rangle$ and $|1\rangle$ form an orthonormal basis of $\mathbb{H}_S$. Let $\hat{A}_S(t)$ be the operator as constructed in (2.4.11) when this system interacts with our detector. The value $\langle \hat{A}_S(t) \rangle$ is intended to map to the physical location of a literal pointer arrow on a piece of equipment with $-\infty$ corresponding to the arrow moving all the way to the left and $\infty$ corresponding to the arrow moving all the way to the right. The detector is calibrated so that if $\hat{\rho}_S = |0\rangle\langle 0|$ then the pointer always moves to the left and so that if $\hat{\rho}_S = |1\rangle\langle 1|$ then the pointer always moves to the right.

(a) Let $a_{mn} = \mathbb{T}(\langle m|\hat{A}_S|n\rangle)$ be the time averaged equilibrium matrix elements of $\hat{A}_S(t)$. Show that the calibration conditions require: $a_{00} = -\infty$ and $a_{11} = \infty$. Explain that $a_{12}$ and $a_{00}/a_{11}$ depend on the experimental preparation and so may be different in each trial.

(b) Assume that the detector and interaction are calibrated so that $a_{12}/a_{00}$ limits to 0 and so that what fraction of the sum $|a_{00}| + |a_{11}|$ is contributed to by each of $|a_{00}|$ and $|a_{11}|$ is uniform. That is, assume the random variable $u$ defined by:

$$u = \frac{a_{00}}{a_{00} - a_{11}}$$

is uniformly distributed. Show that the pointer moves to the left with probability $\langle 0|\hat{\rho}_S|0\rangle$ and that it moves to the right with probability $\langle 1|\hat{\rho}_S|1\rangle$, recovering the Born rule.

**Problem 2.9: Low Rate Double Slit Experiment.** In our previous analysis of the double slit experiment, we have always discussed a continuous interference pattern appearing on the screen. However, if the intensity of the incoming beam is sufficiently low, then the pattern that forms on the screen will form blob by blob and only after many particles in the beam have been sent will the full intensity pattern be visible. In this problem, we will explain these low rate results using our analysis of measurement in quantum mechanics.

Suppose for simplicity in this question that the incoming particle happens to be a pure state and so has some wavefunction. Let vectors like $\vec{x}$ for this question specifically refer to vectors in the plane of the screen.

(a) Our screen is the detector with the bins corresponding to our spatial resolution. In particular, suppose that the screen is divided into regions $\mathbb{M}_n$ labeled by $n$, so that if the incoming particle's wavefunction lies solely within a single $\mathbb{M}_n$ then precisely that pixel will light up upon detection. The property of the incoming wavefunction we wish to measure is its central position, $\langle\hat{\vec{x}}\rangle$. Explain that because of the finite resolution of the bins, the best we can actually measure is the coarse-grained position $\vec{x}_{\text{coarse}}$ defined by:

$$\vec{x}_{\text{coarse}} = \sum_n \langle \vec{x}_n \hat{\pi}_n \rangle$$

where $\vec{x}_n$ is the representative position assigned to $\mathbb{M}_n$ (say, the center of the pixel) and $\hat{\pi}_n$ is the projector of the Hilbert space into the spatial region $\mathbb{M}_n$, or more precisely:

$$\hat{\pi}_n = \int_{\mathbb{M}_n} |\vec{x}\rangle\langle\vec{x}| d^2\vec{x}.$$

A physical screen is a macroscopically large object connected to macroscopic electromagnetic equipment, and so easily satisfies the necessary condition for a detector to not have a precisely preparable microstate described in the section.

(b) Show that in order for which pixel lights up to unbiasedly estimate the value of $\vec{x}_{\text{coarse}}$, we must have the probability $p_n$ that pixel $n$ fires be:

$$p_n = \int_{\mathbb{M}_n} |\psi(\vec{x})|^2 d^2\vec{x}$$

where $\psi(\vec{x})$ is the wavefunction of the incoming particle just before it hits the screen.

(c) Take the limit that the screen becomes arbitrarily precise. Show that $|\psi(\vec{x})|^2$ becomes the probability per unit area on the screen of that point on the screen lighting up for a single incoming particle, thus recovering the standard naive Born rule.

## 2.5 Other Views of Quantum Mechanics

In the century since the beginning of quantum mechanics, there has been enormous confusion about what quantum mechanics even means. This confusion has lead to several viewpoints, which we will discuss in this section. To begin, there are a few empirical facts which are unambiguously tested by experiment that everyone agrees on. Let us begin with the common ground.

**Fact 1: Existence of an ensemble description**

The statistical properties of an ensemble $\mathcal{E}$ of systems of interest prepared macroscopically identically are captured by a density matrix $\hat{\rho}_{\mathcal{E}}$ (which is a self-adjoint, nonnegative, trace-1 operator on the Hilbert space $\mathbb{H}$ of an element of the ensemble). These properties are *statistical* in the sense that $\hat{\rho}_{\mathcal{E}}$ does not in general predict the specific outcome of individual experimental trials or the properties of any particular system within the ensemble; $\hat{\rho}_{\mathcal{E}}$ predicts only distributions of results for possible experiments. If the element systems of the ensemble are isolated, then the density matrix evolves in time according to the von Neumann equation:

$$\frac{d\hat{\rho}_{\mathcal{E}}}{dt} + \frac{1}{i\hbar}[\hat{\rho}_{\mathcal{E}}, \hat{H}] = 0 \tag{2.5.1}$$

where $\hat{H}$ is the self-adjoint Hamiltonian operator on $\mathbb{H}$ for the element systems.

**Fact 2: Operator-Generator correspondence**

A given real-valued generator $s$ of interest of an element of $\mathcal{E}$ which is intended to be "measured" in an appropriate "ideal" measurement is associated with a self-adjoint operator $\hat{s}$. Let $s_n$ be the eigenvalues and $\hat{\pi}_n$ be the eigenspace projectors of $\hat{s}$:

$$\hat{s} = \sum_n s_n \hat{\pi}_n, \qquad \hat{\pi}_n\hat{\pi}_m = \delta_{nm}\hat{\pi}_n, \qquad \sum_n \hat{\pi}_n = \mathbb{1}_S. \tag{2.5.2}$$

$s$ will be measured by coupling the System S to a Detector D through an interaction. In each trial, the detector will register a real value which is able to be observed and recorded by people.

**Fact 3: Certain outcomes**

If $\mathcal{E}$ is such that there exists a value of $n$ so that $\hat{\pi}_n\hat{\rho}_\mathcal{E}\hat{\pi}_n = \hat{\rho}_\mathcal{E}$, then $s$ certainly takes the value $s_n$ in each experimental trial.

**Fact 4: The Born Rule**

For a general ensemble $\mathcal{E}$, in each trial the detector will read one of the $\{s_n\}$ as its value. Within the full ensemble $\mathcal{E}$, the distribution of which $s_n$ is selected is given by the Born rule:

$$\text{Prob}(s = s_n) = p_n = \text{Tr}(\pi_n\hat{\rho}_\mathcal{E}). \tag{2.5.3}$$

**Fact 5: Collapse Rule**

After measuring an ensemble of instances of the system with the Detector, it is possible to form the subensemble $\mathcal{E}_n$ of instances which gave the value $s_n$ under measurement. The statistical properties of that subensemble for future measurements are captured by the density matrix:

$$\hat{\rho}_{\mathcal{E}_n} = \frac{1}{p_n}\hat{\pi}_n\hat{\rho}_\mathcal{E}\hat{\pi}_n. \tag{2.5.4}$$

These facts are empirical in the sense that they are good tools for predicting the correct outcomes of many experiments. Whether these facts are fundamental properties of the universe or results derived from more fundamental physics that happen to be appropriate for the relevant experiments has historically been an interpretative question (with time, advances in measurement technology have made it clear that at least fact 5 is just an approximate fact). All five facts of what everyone agrees on are silent about the properties of an individual experimental trial (an individual system instance within the ensemble $\mathcal{E}$). Once a description of what happens in an individual experimental trial is desired, differences of opinion appear. Any interpretation of quantum mechanics has to be consistent with these five factors. The most conservative possible interpretation of quantum mechanics is the so-called minimal interpretation (also called the statistical interpretation), which is advocated for by Ballentine [2]. In the minimal interpretation, one is only allowed to make statements derivable from Facts 1-5. Therefore, you are not allowed to say anything about individual experimental trials. It is not legal in the minimal interpretation to talk about the energy of a particle; instead one must talk about the average energy of an ensemble of particles. The minimal interpretation has the advantage of protecting one from saying anything plainly wrong by being too bold, but has the conjugate disadvantage of being unable to answer many of the questions we really care about. We live in *one* Universe, and there are many experiments that cannot be repeated by humans (good luck recreating the big bang), and if close adherence to the minimal interpretation is followed, it means we are not legally allowed to say anything about the quantum mechanical aspects of such experiments. The minimal interpretation essentially takes an agnostic approach to the physics of measurement by refusing to attempt to describe what happens in the measurement of an individual quantum system.

In the Thermal Interpretation presented last section: Fact 1 is allowed to be stated even about individual experimental trials; Fact 2 is justified by the form of the Hamiltonian in (2.4.32) implying a set of $\hat{\pi}_n$ and in that interpretation the $s_n$ are essentially arbitrary bin labels; Facts 3 and 4 are determined by the conditions put on detector calibration to be minimally disturbing and unbiased; and Fact 5 can be derived from (2.4.49) for some simple systems, and it is my hope that with time a more general derivation is found. Thus, the Thermal Interpretation is consistent with all of the basic experimental facts of quantum mechanics, though while it is not in any present tension with Fact 5 it also has not yet succeeded in justifying it. In the rest of this section we will discuss some alternate views which are also consistent with these facts.

## 2.5.1 Copenhagen Interpretation

Those unfamiliar with the territory of interpreting quantum mechanics may confuse the minimal interpretation with the Copenhagen interpretation. The Copenhagen interpretation is the one you are most likely to encounter in low-level popular science media or in a beginning quantum mechanics course. Often those who are not especially familiar with quantum mechanics will erroneously think of the Copenhagen interpretation as synonymous with quantum mechanics itself. The Copenhagen view was advocated for historically primarily by Niels Bohr. The minimal interpretation is very conservative and pragmatic; it says that it is only going to be concerned with the statistical properties of many measurements and will agnostically say nothing about the properties of individual experiments. Copenhagen on the other hand is extremely bold because it explicitly asserts that the properties of individual experiments *cannot possibly be answered* even in theory. Since the

Copenhagen interpretation rejects the existence of answers to several questions about quantum mechanics, it has often been called the "shut-up and calculate" approach. There are a few interrelated assumptions defining the Copenhagen interpretation. It asserts Facts 1-5 as exact fundamental physical truths and additionally assumes:

*Individuality*: Fact 1 can be applied to individual experimental runs.

*Indeterminacy*: Quantum mechanics is fundamentally probabilistic with absolutely no possible underlying mechanism for how to explain the Born rule.

*Naive collapse*: When a quantum system is not being measured, it evolves according to the von Neumann (or Schrödinger, in the case of a state vector) equation and when it is measured, that unitary time evolution ceases and the measurement causes the measured system's state to "collapse" *instantly* to the appropriate eigenstate of the operator associated to the measurement described in Fact 5.

*Limited Reality*: A property of a quantum system exists and is physical only when it is being measured. When not measured, it is undefined.

The Copenhagen interpretation agrees with the Thermal interpretation in allowing the discussion of individual quantum systems (Individuality). Indeterminacy is a much bigger leap. Copenhagen treats the state (either state vector or density matrix) of the system being measured as the full set of data relevant in an experiment and then upon observation that the same preparation of such data leads to a distribution of experimental outcomes asserts that it is because of fundamental indeterminacy. In my opinion the Indeterminacy assumption is philosophically and experimentally unwarranted and fundamentally at odds with every other known thing in the domain of science. You could likely guess as much, given that I argued against the scientific sensibility of indeterminism in Orig. section 5.1.4 and provided an in my opinion convincing explicit model for how this distribution of experimental results can be explained with unitary time evolution in the previous section. Because the Copenhagen interpretation cannot predict the results of measurements even in principle, it is often called an incomplete theory (such as by Einstein).

I will no discuss some challenges of the Copenhagen view. This will be a mixture of reasoned argument and my personal opinion based on those arguments. Because the Copenhagen interpretation refuses to describe collapse as a dynamical process given by some equation of motion, it treats collapse as instantaneous at the time of measurement. This assumes an exact moment of measurement can be identified and requires a notion of simultaneity. Relativity is fundamentally at odds with an invariant concept of simultaneity. An equal time slice in one reference frame is not an equal time slice in another boosted reference frame. Therefore, in naive collapse whether the measurement has even happened is not an observer independent statement. Because in the Copenhagen interpretation the state is believed to be instantly changed to an eigenstate of the operator associated to the measurement and whether the measurement has happened is not observer independent, in the Copenhagen interpretation the state of the system is not observer independent. The tension between Copenhagen and relativity worsens the more you think about the nonlocal nature of instantaneous collapse and the spatially extended nature of the density matrix. Since the quantum density matrix describes the whole state of a system, for a spatially extended system it contains data at spatially separated points. If the position operator (or a close approximate) is measured in the Copenhagen interpretation, the state will instantly collapse to an eigenstate of that operator and thus completely localize in space, changing the state instantly at spatially separated points. If the state is physical, this violates relativistic causality. Thus, the Copenhagen interpretation rejects the very notion that the physical state of a system is *physically real*. In my view, a basic comprehension of the scientific method precludes taking this viewpoint seriously.

The denial of reality in Copenhagen is clarified by the final point Limited Reality. In the Copenhagen interpretation, the results of measurements that are actually performed are considered as real. Systems are not considered real when they are not being measured. This thinking is the origin of common popular science descriptions of quantum mechanics. In such a description you might hear the wave-particle duality described as beginning with the wavefunction being a fictitious mathematical tool (obviously the wavefunction can't be real, because it has an imaginary part[4]). Then, when the wavefunction is measured the wave will instantly

[4]This parenthetical is entirely sarcastic. This deranged "proof-by-same-word" logic of incorrectly interpreting what the real in real-numbers means is entirely without merit. Anyone capable of passing an undergraduate math class should be able to understand that.

become a perfect little point particle somewhere (which you should know is not even a mathematically sane concept, based on our work in classical mechanics.). When that wave collapses to a little point particle it suddenly becomes physically real. If it were never measured, it wouldn't have been real. If my grandma had wheels, she would be a bicycle.

A resulting property of the Copenhagen interpretation is *complementarity*. Because the Copenhagen interpretation assumes a physical property of a system is real only at the time that property is measured and because by Fact 2 measurement results in resolving the operator associated to a property of interest into its eigenspaces, the mathematical fact that noncommuting operators cannot be simultaneously diagonalized results in the Copenhagen interpretation to it being impossible to simultaneously assign physically real values to two properties of a quantum system if the operators associated to those to properties do not commute. This is why you will hear those attempting to make quantum mechanics sound mysterious say that in quantum mechanics it is impossible to know the position and momentum of a particle at the same time. This metaphysical weirdness has nothing to do with quantum mechanics; it has solely to do with the Copenhagen interpretation. The central position $\langle\hat{\vec{x}}\rangle$ and total linear momentum $\langle\hat{\vec{p}}\rangle$ of a single particle quantum system are perfectly well-defined functions of time at all times with no weirdness. The same experimental setup one would use to measure $\langle\hat{\vec{x}}\rangle$ by an ideal measurement does not serve as an ideal measurement of $\langle\hat{\vec{p}}\rangle$, which is not at all mysterious, because they require different interaction Hamiltonians. To measure different properties, one needs appropriately differently designed experiments. If you believe in reality, this says nothing about whether the unmeasured values existed. Copenhagen on the other hand, due to its rejection of the density matrix being physically real, must reject the unmeasured values also from being physically real.

Because the Copenhagen interpretation was very early to the historical game, its perspective was heavily used in assigning the standard names given to quantum mechanical objects. This has resulted in countless aspiring physicists baking in many of the concepts of the Copenhagen interpretation deeply into their quantum mechanical thinking, even when completely unnecessary. Many physicists have been taught to think of $|\psi(t,\vec{x})|^2$ as the spatial probability density for finding a single quantum particle (because they always think of $|\langle\psi|\phi\rangle|^2$ as a Born rule probability), to think of $\langle\hat{H}\rangle$ as the (statistical) average energy of a quantum system (because books mostly call it the expectation of $\hat{H}$), to think of a measurement as measuring an operator (even though the operators aside from the density matrix are not even dynamical variables), to think that quantum systems are always in eigenstates of whatever observables you care about (because of interpreting the eigenstate outcomes of the Born rule extremely literally), and so on with many basic unnecessary misconceptions.

Aside from the basic philosophical issues we have already mentioned with the Copenhagen interpretation, it further contains within itself a fundamental flaw: the Copenhagen interpretation does not define what a measurement is. By refusing to give a dynamical interactive picture for what constitutes a measurement, the Copenhagen interpretation admits that it at best provides an incomplete description of physics. Copenhagen simply says that we know when a measurement is performed in a lab because we do the measuring. But, it cannot answer basic scientific questions. Do atoms measure each other? There is a classic joke whose origin I haven't found that goes something like: does an amoeba measure a quantum system by getting too close to it, or does the amoeba have to get a PhD and a spectrometer first? Copenhagen can't say.

The Copenhagen interpretation is based on a familiar error. In the history of classical mechanics, it was commonly said that friction couldn't be explained by Hamiltonian/Action/Lagrangian mechanics. The core of the argument was that if you write down the term friction adds to the equation of motion of a particle, you can show that there is no term you could possibly add to the action which depends only on the degrees of freedom of the particle that results in adding that term to the equation of motion. The critical incorrect assumption here was in thinking you should add friction by only considering the degrees of freedom of the particle. We have long since known that in fact friction and other dissipative effects originate precisely because of ignoring additional degrees of freedom. By adding terms to the action which include the requisite degrees of freedom to dissipate energy into, friction can unambiguously be treated with a variational principle, as treated in the Orig. or in [37,38]. A half century later, in statistical mechanics there were many who would say that a Hamiltonian system cannot approach thermodynamic equilibrium, because the entropy of the phase space density of a Hamiltonian system is conserved. This is again the same error, in which time-averaging and the interaction of the system in question with an environment of external variables is inappropriately 5ignored. Another half century later, the same mistake was made again. The Copenhagen interpretation is born from the observation that the final states in Fact 5 depend nonlinearly on the initial state while the von Neumann equation is purely linear. There is no Hamiltonian operator that could result in the density matrix evolving nonlinearly, therefore (Copenhagen concludes) the fact that the Fact 5 states are empirically good

after measurement must mean that measurement cannot be described by ordinary linear unitary quantum time evolution. Just like the friction example, this is wrong. Measurement is a dynamical interactive process, and so during measurement the System's density matrix does not satisfy the von Neumann equation, the combined System + Detector density matrix satisfies the von Neumann equation. The Copenhagen thinking ignores the Detector degrees of freedom. There is no reason to believe that the System's density matrix evolves linearly when it is a subsystem of an interacting system. In fact, we can explicitly construct physical examples in which it does not (see [7]). The nonlinear relationship between the initial and final states of the System is fully able to be explained by linear unitary time evolution on the full System+Detector system. By neglecting the detector as part of the measurement, Copenhagen misses this.

It is worth noting that the greatest physicist of the quantum era (and quite possibly of all time), Einstein, never believed and in fact actively opposed the Copenhagen interpretation. It is further worth noting that Schrödinger, who actually found the equation of motion of quantum mechanics, also never believed the Copenhagen interpretation. Here is one of many things he had to say about it in his lifetime [55]:

*With very few exceptions (such as Einstein and Laue) all the rest of the theoretical physicists were unadulterated asses and I was the only sane person left. The one great dilemma that ails us day and night is the wave-particle dilemma. In the last decade I have written quite a lot about it and have almost tired of doing so: just in my case the effect is null because most of my friendly nearer colleagues have formed the opinion that I am, naturally enough, in love with my great success in life via wave mechanics reaped at the time I still had all my wits at my command and therefore, so they say, I insist upon the view that "all is waves". Old-age dotage closes my eyes towards the marvelous discovery of "complementarity". So unable is the good average theoretical physicist to believe any sound person could refuse to accept the Copenhagen oracle.*

It is important to understand that typical theory oriented physics graduate students today have a much deeper math education than any of the founders of the Copenhagen interpretation. What Bohr and Heisenberg could or could not imagine a century ago is irrelevant to modern science. Because Facts 1-5, viewed as empirical observations have no interpretative content, the various problematic aspects of Copenhagen historically had no experimental impact. The Copenhagen interpretation does not enable you to explain any observations beyond what Facts 1-5 already explain, as it provides *antianswers* to any further questions. Consequently, Copenhagen has no practical advantages over the minimal interpretation but it does have a mountain of philosophical disadvantages. If you don't want to have an opinion on quantum mechanics, just humbly use the minimal interpretation. We have outgrown Copenhagen.

### 2.5.2 Post-Schrödinger Dynamics

Post-Schrödinger dynamics is not a specific interpretation, but more so a class of similar-in-concept interpretations. The historically most prominent representative of this class is called the de Broglie-Bohm interpretation. Post-Schrödinger interpretations allow Fact 1 to be applied to individual experimental runs and then try to derive, as much as possible, Facts 2-5 as emergent empirical facts rather than fundamental features of reality. So far, this is exactly what the Thermal/Unitary interpretation also wants. The only difference between the Thermal interpretation and post-Schrödinger interpretations is how these facts are intended to be derived. In the Thermal interpretation, these facts are derived by using ordinary unitary von Neumann/Schrödinger time evolution on the combined System+Detector system. Post-Schrödinger interpretations do not typically model the Detector directly; instead, they modify the von Neumann/Schrödinger equation for the System by adding to it some small extra terms that are intended to cause the dynamical derivation of the final states in Fact 5 and/or they add extra degrees of freedom still attached to the System that are supposed to be the variables we actually see in measurement. Post-Schrödinger interpretations thus intend to fundamentally replace/correct quantum mechanics with a new dynamical theory that is approximately quantum mechanics when measurements are not happening. It is very difficult to do this in a way consistent with experiment, because many quantum experiments are extremely precise and these new terms tend to cause at least some small changes in experimental consequences outside of simply Facts 1-5.

In my view, the extra terms and extra degrees of freedom of post-Schrödinger approaches are just a clunky way of trying to derive facts fully explainable without modifications to the dynamical structure of quantum mechanics *if you actually include the full quantum system.* Thus, I view post-Schrödinger approaches as an understandable attempt, from a historical perspective, to figure out what is going on in quantum mechanics though I believe their usefulness is nullified by the Thermal interpretation. Often one

will hear post-Schrödinger approaches described as "hidden variable theories" as they (typically) add extra variables relevant to experiment beyond simply the density matrix or state vector. In the common lore of interpretations of quantum mechanics, it is often claimed that hidden variable theories cannot possibly explain the experimental facts of quantum mechanics. This at-best confused lore is based on a common misunderstanding of Bell's theorem, which we will discuss later in this section.

### 2.5.3 Many Worlds Interpretation

As the troubled nature of the Copenhagen interpretation has approached critical mass of social appreciation among physicists in recent years and a clear view of Bell's theorem has not yet diffused into the general mass of physicsfolk, the Many Worlds interpretation has become the trendy way to think about quantum mechanics. In some ways, the many worlds interpretations is mathematically similar to the Thermal interpretation. The many worlds interpretation does model the full measurement process by System+Detector using the very same Hamiltonian operator (2.4.32) as the Thermal interpretation. Using the same standard decoherence arguments as the Thermal interpretation, the many worlds interpretations shows dynamically that the final density matrix appropriate for a statistical ensemble of the System at the end of measurement is precisely (2.4.53):

$$\hat{\rho}_S^{(f)} = \sum_n p_n \hat{\rho}_S^{(n)}, \qquad p_n = \mathrm{Tr}(\hat{\pi}_n \hat{\rho}_S), \qquad \hat{\rho}_S^{(n)} = \frac{1}{p_n} \hat{\pi}_n \hat{\rho}_S \hat{\pi}_n. \tag{2.5.5}$$

The real leap of the many worlds interpretation comes in how this formula is interpreted. This formula, as the density matrix of an ensemble of final system states, is *consistent* with a statistical distribution over the states $\hat{\rho}_S^{(n)}$ each with respective probability $p_n$. The many worlds interpretation takes this distribution literally and asserts that the ensemble is *always real*. It claims that there are infinitely many copies of the universe and that in each copy, whenever a measurement happens, exactly one of the results $\hat{\rho}_S^{(n)}$ is achieved, and that among all of the copies the fraction in which that specific result is achieved is $p_n$. The many worlds interpretation is thus essentially hyper-frequentist. In some sense it claims that there is no true randomness because all of the possibilities happen, we just simply don't know which of the infinite number of universes in the ensemble we are in for a particular measurement until we read out the measurement results.

The calculation leading to (2.5.5) (the steps leading to (2.4.53)) as the appropriate ensemble density matrix for the System after measurement is entirely uncontroversial in modern quantum mechanics. There is a second more "out there" calculation that is used to motivate the many worlds interpretation specifically. As presented here, that calculation is based on the calculation done in section 3.7 of [63], which in turn is based on several older papers cited in that section. Assume that the final ensemble density matrix above is taken by hypothesis to be the appropriate density matrix to assign in the final state immediately before reading the experimental result of a single experimental trial. Then, consider $N$ copies of the experiment running in parallel. The overall density matrix of these $N$ copies then would be:

$$\hat{\rho}^{(N)} = \hat{\rho}_S^{(f)} \otimes ... \otimes \hat{\rho}_S^{(f)} = \sum_{n_1 ... n_N} p_{n_1} ... p_{n_N} \hat{\rho}_S^{(n_1)} \otimes ... \otimes \hat{\rho}_S^{(n_N)}. \tag{2.5.6}$$

Define the frequency operators $\hat{P}_n^{(N)}$ as operators which count precisely what fraction of times each outcome occurred when acting on states in definitive outcome eigenspaces. That is, if $|\psi_n\rangle$ is a vector in the Hilbert space of a single experiment which is purely in the eigenspace projected into by $\hat{\pi}_n$, then:

$$\hat{P}_n^{(N)}(|\psi_{n_1}\rangle \otimes ... \otimes |\psi_{n_N}\rangle) = \frac{C(n_1, ..., n_N; n)}{N} |\psi_{n_1}\rangle \otimes ... \otimes |\psi_{n_N}\rangle \tag{2.5.7}$$

where $C(n_1, ..., n_N; n)$ is the number of times $n$ occurs among the list $n_1, ..., n_N$. Then, we can compute:

$$\langle \hat{P}_n^{(N)} \rangle = \mathrm{Tr}(\hat{\rho}_n^{(N)} \hat{\rho}^{(N)}) = \sum_{n_1, ..., n_N} \frac{C(n_1, ..., n_N; n)}{N} p_{n_1} ... p_{n_N}. \tag{2.5.8}$$

For a fixed number $N_n = C(n_1, ..., n_N; n)$ of occurrences of $n$, there are $N$ choose $N_n$ equivalent reorderings of the labels $n_1, ..., n_N$. So, with this factor, we can unambiguously fix the first $N_n$ to be $n$ and guarantee that the remaining $N - N_n$ are not $n$:

$$\langle \hat{P}_n^{(N)} \rangle = \sum_{N_n=0}^{N} \sum_{\substack{n_{N_n+1}, ..., n_N \\ \text{None } = n}} \frac{N_n}{N} \binom{N}{N_n} p_n^{N_n} p_{N_n+1} ... p_{n_N} = \sum_{N_n=0}^{N} \frac{N_n}{N} \binom{N}{N_n} p_n^{N_n} (1 - p_n)^{N - N_n} \tag{2.5.9}$$

$$= \frac{p_n}{N}\frac{\partial}{\partial p_n}\sum_{N_n=0}^{N} p_n^{N_n} q^{N-N_n}\bigg|_{q=1-p_n} = \frac{p_n}{N}\frac{\partial}{\partial p_n}(p_n+q)^N\bigg|_{q=1-p_n} = p_n. \tag{2.5.10}$$

Similarly:

$$\langle\left(\hat{P}_n^{(N)}\right)^2\rangle = \sum_{N_n=0}^{N}\sum_{\substack{n_{N_n+1},\ldots,n_N \\ \text{None }=n}} \frac{N_n^2}{N^2}\binom{N}{N_n} p_n^{N_n} p_{N_n+1}\ldots p_{n_N} = \sum_{N_n=0}^{N}\frac{N_n^2}{N^2}\binom{N}{N_n} p_n^{N_n}(1-p_n)^{N-N_n} \tag{2.5.11}$$

$$= \frac{p_n}{N^2}\frac{\partial}{\partial p_n} p_n \frac{\partial}{\partial p_n}\sum_{N_n=0}^{N} p_n^{N_n} q^{N-N_n}\bigg|_{q=1-p_n} = \frac{p_n}{N^2}\frac{\partial}{\partial p_n} p_n \frac{\partial}{\partial p_n}(p_n+q)^N\bigg|_{q=1-p_n} \tag{2.5.12}$$

$$= \frac{p_n(1-p_n)}{N} + p_n^2. \tag{2.5.13}$$

Thus:

$$\langle \hat{P}_n^{(N)}\rangle = p_n, \qquad \text{Spr}(\hat{P}_n^{(N)}) = \frac{p_n(1-p_n)}{N}. \tag{2.5.14}$$

Imagining the $N\to\infty$ limit, we see that in that limit the state $\hat{\rho}^{(N)}$ tends toward having zero spread with respect to $\hat{P}_n^{(N)}$ and assigns it a summary value $p_n$. The many worlds thinking interprets this calculation as showing that the fraction of states in the ensemble of systems which land in the state $\hat{\rho}_S^{(n)}$ is precisely $p_n$, because that is (so it is claimed) what $\langle\hat{P}_n^{(N)}\rangle$ represents, since it counts occurrences in each individual eigenspace.

In my opinion, there are a two unjustified steps of reasoning in coming up with the many worlds interpretation. The first is that the density matrix $\hat{\rho}_S^{(f)}$ is assigned to *each* individual final state right before measurement when it is in fact the correct density matrix for the *ensemble* of final states. The decoherence argument necessary to derive $\hat{\rho}_S^{(f)}$ relies on a time-averaging, relaxation to equilibrium, and ensemble averaging. The time evolution that brings the System's state in this form is reliant on its coupling to the Detector (whose existence is suddenly ignored in the act of constructing the "many worlds" in $\hat{\rho}^{(N)}$). Thus, inadvertently the many worlds interpretation assumes the requisite ensemble on the first step of calculation. The second error in interpretation is the meaning assigned to $\hat{P}_n^{(N)}$. It is true (by construction) that when $\hat{P}_n^{(N)}$ acts on a state of the form in (2.5.7), it has the effect of simply multiplying by the count fraction $N_n/N$. However, it is a separate leap to conclude that in some linear combination of such states, generally forming a state which is not an eigenstate of $\hat{P}_n^{(N)}$, the operator summary of $\hat{P}_n^{(N)}$ still represents anything like the fraction of counts of the experiment which will be assigned the final state $\hat{\rho}_S^{(n)}$. In fact, by hypothesis, the many worlds interpretation already assumes each final state is precisely assigned to be $\hat{\rho}_S^{(f)}$, not $\hat{\rho}_S^{(n)}$.

I believe that the many worlds interpretation's errors are born from a misunderstanding of the physical properties of a linear combination. It is not unique to this interpretation for quantum mechanists to misunderstand linearity, so remember the following points when thinking about quantum mechanics in general. The second error in the previous paragraph may (to someone not being careful) seem logically okay – if every element in a linear combination has a property then certainly the linear combination itself should have that property right? No. Absolutely not. That is not how linearity works. A linear combination of Dirac deltas at different points is a not a Dirac delta, so even though every state in such a linear combination has a position spread of 0, the linear combination does not. A vector pointing east in the plane is a linear combination of vectors pointing northeast and southeast, each of which have the property of changing the coordinate along the north-south axis, but their sum does not. In quantum electrodynamics, the electric field operator (defined in (1.10.30) and (1.10.32)) is a linear combination of creation and annihilation operators. Therefore, the electric field operator when acting on an $N$ photon state results in a linear combination of states of $N-1$ and $N+1$ photons. Consequently, the classical electric field $\langle\hat{\vec{E}}\rangle$ of every single $N$ photon state is 0. However, a linear combination of states of different photon numbers gives a nonzero classical electric field. The literal existence of classical electromagnetism is proof that the logic of this second error is wrong.

At its core, the many worlds interpretation looks at the density matrix $\hat{\rho}_S^{(f)}$ and asserts that it is a genuine statistical mixture of the $\hat{\rho}_S^{(n)}$ with probabilities $p_n$. This is, in my opinion, a misunderstanding of how linearity works. The same final state ensemble density matrix $\hat{\rho}_S^{(f)}$ can be decomposed into a sum of the form (2.5.5) for multiple different sets of $\hat{\rho}_S^{(n)}$ and $p_n$; that decomposition is ambiguous. Therefore

any argument that attempts to derive the distribution $p_n$ of outcomes $\hat{\rho}^{(n)}$ from the density matrix $\hat{\rho}_S^{(f)}$ is doomed to fail; it is wrong on principal. See the example given in the text of subsection 2.1.2. In order to actually justify the distribution of the ensemble density matrix, you must start from individual instances of the system and derive probabilities. The math is not reversible.

I find the many worlds interpretation not only logically unsatisfying but also philosophically vacuous. The silver bullet that the many worlds view uses to explain the empirical facts of quantum mechanics is the introduction of an infinite number of inaccessible other copies of the universe. But, the interpretation offers no even formal structure through which these other copies of the universe could ever even in principle be measured or interacted with, and therefore the claim of their existence is **NOT** a scientific hypothesis, as it is entirely unfalsifiable. The central claim of the many worlds interpretation is literally not science.

### 2.5.4 The Damned Cat

Ugh. Is it time to talk about Schrödinger's cat? I guess it is time to talk about Schrödinger's cat. If somehow you haven't heard about this poor cat, let me explain. Consider a cat in a sealed box which also contains some poison in a container with a release mechanism, a radioactive particle, and a detector capable of reading the radioactivity. The release mechanism is such that if the detector goes off, the poison is released, killing the cat (boooo). Initially the cat, viewed as a quantum system, is definitely in a subregion of its Hilbert space where it has the property of "alive". After the decay, the cat is definitely in a subregion of its Hilbert space where it has the property of "dead". While the box is closed and the particle is evolving, the particle's quantum state will generally be a linear combination of a state of particle number 1 and a state of particle number 0. It is supposed then that before the decay has happened, the cat is similarly in a linear combination of its before and after decay states. Thus, the cat is in a linear combination of states, the first term in the linear combination being a state which taken on its own is one in which the cat is definitely alive and the second being a state which taken on its own is one in which the cat is definitely dead. The thought experiment has the effect that a microscopic linear combination (of the radioactive particle) induces a macroscopic linear combination (of the cat's dead and alive states). The question is, what are the properties of this macroscopic linear combination?

Schrödinger originally considered this thought experiment to make fun of the Copenhagen interpretation. Within Copenhagen, the thought experiment reveals a basic question about what counts as measurement. Does the cat cause measurements? Does the detector cause measurements? Or, is it we who open the box after a while who cause the measurement? If the cat, by being able to hear the detector click, is constantly measuring whether or not the decay has happened, then for some amount of time the cat is in the alive state until it hears a click and shortly after dies. If the cat does not constitute a measurer but the detector does, then the macroscopic linear combination of the cat between an alive state and dead state progressively forms and one has to wonder what it means for a cat to be in such a linear combination. At some point, the detector notices the decay and the cat dies. The final conceptualization of what counts as measurement is what Schrödinger was really trying to make fun of originally, which is to think that neither the cat nor the detector count as taking measurements. It is humans who count as performing the measurement when we open the box. Then, before we have opened the cat is in its macroscopic linear combination. When we open the box, either the detector will have gone off or not, and correspondingly the cat will suddenly collapse to either the alive state or the dead state. This position is so bizarre because it places humans in a special role in nature, which is antithetical to basic science (the Copernican principle). In a popular science magazine, you are likely to read the cat as being described as "both dead and alive" before the box is opened. The Copenhagen interpretation is not capable of resolving which of these three participants (the cat, the detector, or the humans) are doing the measuring, and so is again incomplete as a theory. In any of the three cases, when the box is opened the cat is seen to either be dead or alive and the probabilities computed from Facts 1-5 are the same. Copenhagen rejects reality except at the time of measurement, but in this case the time of measurement is not determined by Copenhagen, and so it is unclear what the interpretation counts as real.

The many worlds and thermal interpretations essentially match on this topic, differing only in the mechanism behind how the decay itself occurs, and that isn't really the heart of this thought experiment. In either of these two interpretations, the cat and the detector are both macroscopic. Whether or not the detector has gone off always has a physically real answer. Once the detector has gone off the poison is released and the cat dies. We learn the answer to whether the cat is dead or alive once the box is opened but this does not mean anything like "before the box is opened the cat is dead and alive". The cat goes from a state of definitely alive to definitely dead at an unpredictable time. That time is unpredictable because the time

of the decay is unpredictable, and that time is unpredictable (within the thermal interpretation) because of the coupling of the radioactive particle to infinitely many uncontrolled background degrees of freedom (the various fields that exist in our universe which weakly couple to basically everything at any nonzero temperature) or (within the many worlds interpretation) because we don't know which of the infinitely many possible universes we are in.

The objection one is likely to hear to the thermal and many world interpretations' accounting of events is about what does the macroscopic linear combination of dead and alive look like for the cat before the decay has happened? This question of course was also present in the Copenhagen interpretation, but because that interpretation rejects reality before the measurement, it doesn't really have to answer it. Interpretations that believe in realism do. The answer is simple: the whole concept that the cat being in a linear combination of a state in the alive part of the Hilbert and a state in the dead part of the Hilbert space means the cat is "both dead and alive" is a complete misunderstanding of how linear combination works. There is absolutely no reason to believe that macroscopic properties (deadness/aliveness) are linear over the microscopic state of the system. For an easy physical example you can see with your eyes, consider the visible region of the electromagnetic spectrum. There is a certain band of the spectrum corresponding to red light; there is another band that corresponds to purple light. A state has the property "is red" if it is approximately purely in the spectrum band that is red (approximately within the resolution of the eye or other detector to see the components of colors outside of this band). A state has the property "is not red" if it has dominating components outside this band. Similarly, one can define the properties "is purple" and "is not purple" from its corresponding band. The color pink is a linear combination of both red light and purple light. Does this mean that pink is both red and also purple? Being a linear combination of two things, each with a property, does not mean that the combination has both or either property. Pink is not red and it is not purple, it is just a different thing. Back to the cat, for every state in the cat's Hilbert space, it is possible to assign the property or alive, dead, or unclear to that state. The unclear condition exists because biologically what counts as alive is a little bit fuzzy (heart stopping versus brain death, for example) and we can basically ignore it for this question. Every state in the Hilbert space takes one of these three properties definitely. The properties are not linear over the linear structure of the Hilbert space, and so if one takes an arbitrary state from the alive region and combines it with an arbitrary state from the dead region, it will end up in some region, but completely unclear which. This breaking of linearity when assigning properties that do not respect the lienar structure happens for even the simplest nontrivial vector space. Consider a two dimensional plane, divided into four quadrants. We can define a vector as type NE, NW, SE, SW, N, E, S, W, or 0 (with letters corresponding to the cardinal directions) based on in the first four cases which quadrant it points from the origin into, in the fifth through eighth cases which pure cardinal direction it points along, and in the last case if it is the 0 vector. Thus every vector in the plane has the property of definitely fitting into exactly one of these nine categories. The property of which one it fits into is not linear over the vectors. Consider the vector $(3, 2)$ which is NE and the vector $(3, -3)$ which is SE; their linear combination is $(6, -1)$ which is SE. Alternatively, consider the vector $(3, 3)$ which is NE and the vector $(3, -1)$ which is SE; their linear combination is $(6, 2)$ which is NE. Therefore, from the properties assigned to the vectors (SE'lyness and NE'lyness) alone it is impossible to know what property to assign to their linear combination. Even worse, it would not even be an invariant property to assign $(6, 2)$ the property of "both NE and SE" as would be suggested by dead-and-alive logic; you can also write it as a linear combination of $(4, 1)$ and $(2, 1)$ which are both NE, no SE part to be found. This refutation is already possible in a 2 dimensional vector space. In the unimaginably high dimension Hilbert space of a cat, the complicated properties of aliveness and deadness have absolutely no reason to be asserted to be anything even related to linear over the microscopic state of the cat. That is neither how math nor how physics works[5].

In short, the entire Schrödinger's cat discussion is basically irrelevant for actual physics. It is a funny jab at the ridiculousness of the Copenhagen interpretation but all of the weirdness described in the thought experiment is either an artifact of the absurdity of the interpretation used (in the case of Copenhagen) or of a misunderstanding of the nature of linear combination (in the case of claiming the cat is "both dead and alive"). In a sane world, we could have stopped telling this joke 90 years ago.

[5]But wait, I thought that all properties were linear over the microscopic state because the summary map is linear? No. The values of extensive generators of canonical transformations are linear in the density matrix. Other quantities generally aren't. The spread of a single operator is a property of the state of a system but is not linear over the density matrix. The entire density matrix is determined by the summary map, and therefore any property of a quantum system can be written as a function of the operator summaries, each of which is linear over the system state. But, one can consider nonlinear functions of these operator summaries, and the resulting property from that function is then nonlinear over the system state. Aliveness and NE'lyness are nonlinear functions of operator summaries.

### 2.5.5 Bell-type Theorems and Their Consequences

This section is based on [39]. Bell's theorem is one of the most frequently misunderstood results in all of physics. If you have only engaged with it at the Wikipedia-level, you are essentially certain to completely misunderstand it, but I promise it isn't actually that complicated. Bell's theorem and various similar theorems at their core are mathematical statements. Bell-type theorems all consist of two parts. In the first, they give examples of quantum systems satisfying interesting equalities or inequalities. These relations are interesting because in the second part they show that no classical probability space could also satisfy them, where the operator summary is interpreted as a statistical expectation and the operators are interpreted as random variables in that probability space. The effect of this is to show that quantum mechanics is not equivalent to classical probability theory, which as a mathematical statement is unambiguously true. There is a layer of interpretation that is very often added to Bell-type theorems after this which claims that Bell's theorem disproves the possibility that the apparent randomness in quantum mechanics could be explained by a real local hidden variable theory. This is certainly what Bell believed. I will first explain the mathematics of Bell's theorem, then how this interpretive layer is added, then finally why that interpretive layer is generally mistaken. The fundamental flaw in the interpretive layer is in its understanding of the word "real". Essentially, the way it constructs the classical probability space is not how one would construct the probability space in the relevant hidden variable theories. Let's explain the details.

We will treat only the most classic example of a Bell inequality, but the importance and logic of the result is the same regardless of the example. Consider a quantum system consisting of two interacting subsystems, each with a two dimensional Hilbert space (for example, each subsystem could be a spin-1/2 particle). Call the first system A (or often Alice, for cuteness) and the second system B (often Bob, similarly). Consider two different operators of interest in each system, $\hat{A}_0$ and $\hat{A}_1$ for Alice and $\hat{B}_0$ and $\hat{B}_1$ for Bob, defined by:

$$\hat{A}_0 = \sigma_z^{(A)} \otimes \mathbb{1}^{(B)}, \qquad \hat{A}_1 = \sigma_x^{(A)} \otimes \mathbb{1}^{(B)}, \qquad \hat{B}_0 = -\mathbb{1}^{(A)} \otimes \frac{\sigma_x^{(B)} + \sigma_z^{(B)}}{\sqrt{2}}, \qquad \hat{B}_1 = \mathbb{1}^{(A)} \otimes \frac{\sigma_x^{(B)} - \sigma_z^{(B)}}{\sqrt{2}}. \tag{2.5.15}$$

Let $|+\rangle$ and $|-\rangle$ denote the eigenstates of $\sigma_z$ with eigenvalues $+1$ and $-1$, and consider the state vector:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|+_A -_B\rangle - |-_A +_B\rangle). \tag{2.5.16}$$

In this state, it follows that:

$$\langle \hat{A}_0 \rangle = 0, \qquad \langle \hat{A}_1 \rangle = 0, \qquad \langle \hat{B}_0 \rangle = 0, \qquad \langle \hat{B}_1 \rangle = 0. \tag{2.5.17}$$

As well, it follows that:

$$\langle \hat{A}_0 \hat{B}_0 \rangle = \langle \hat{A}_0 \hat{B}_1 \rangle = \langle \hat{A}_1 \hat{B}_0 \rangle = -\langle \hat{A}_1 \hat{B}_1 \rangle = \frac{1}{\sqrt{2}} \tag{2.5.18}$$

from which we have:

$$\langle \hat{A}_0 \hat{B}_0 \rangle + \langle \hat{A}_0 \hat{B}_1 \rangle + \langle \hat{A}_1 \hat{B}_0 \rangle - \langle \hat{A}_1 \hat{B}_1 \rangle = 2\sqrt{2}. \tag{2.5.19}$$

This identity, purely within the domain of our quantum system, is not very interesting. The above result is given import by considering a naively related system in probability theory.

Consider a probability space in which there are four random variables, $a_0, a_1, b_0, b_1$, so that in each outcome of the probability space, each of these variables takes the value $+1$ or $-1$. Plainly:

$$a_0 b_0 + a_0 b_1 + a_1 b_0 - a_1 b_1 = (a_0 + a_1) b_0 + (a_0 - a_1) b_1. \tag{2.5.20}$$

Because both $a_0$ and $a_1$ are either $\pm 1$, it must be that $a_0 = a_1$ or $a_0 = -a_1$. In the first case, the $b_1$ term on the right hand side above does not contribute and the expression has magnitude 2. In the second case, the $b_0$ term on the right hand side above does not contribute and the expression again has magnitude 2. Therefore, it is identically true that:

$$|a_0 b_0 + a_0 b_1 + a_1 b_0 - a_1 b_1| = 2. \tag{2.5.21}$$

There will be some probability that the expression inside the absolute value takes the value 2 and some probability it takes the value -2. But, because these are the only values it may take, its average cannot exceed 2 in magnitude. Therefore, for any distribution on these four variables:

$$|\mathbb{E}[a_0 b_0 + a_0 b_1 + a_1 b_0 - a_1 b_1]| \leq 2. \tag{2.5.22}$$

This identity considered purely in the domain of probability theory is also not very interesting. Let us now

see how (2.5.19) and (2.5.22) are often interpreted together.

Return to our quantum system and think about how it behaves under measurement without applying any particular interpretation (so we can only use Fact 1-5). We take the operators $\hat{A}_0, \hat{A}_1, \hat{B}_0, \hat{B}_1$ to all correspond to properties of experimental interest. Each of these operators has eigenvalues $+1$ and $-1$. We are going to prepare a large number of identical copies of our system and perform many measurements. Because $\hat{A}_0$ and $\hat{A}_1$ do not commute, in each trial we will only measure the property associated to one of them. Similarly so, in each trial we will only measure the property associated to either $\hat{B}_0$ or $\hat{B}_1$. That is, in our different trials we will have different settings on our measurement equipment appropriate for measuring different properties of the system. The $\hat{A}$ commute with the $\hat{B}$ and so in each trial we will select an $\hat{A}$ operator and a $\hat{B}$ operator to measure the properties associated to each of them. Let the measurement results associated to our operators be called $a_0, a_1, b_0, b_1$. By the 5 basic facts of measurement, in each experimental trial no matter which subset of $a_0, a_1, b_0, b_1$ are measured, each of the ones which are measured will take the value $+1$ or $-1$. Different trials will give different values for the measured variables following the Born rule distribution, and for now we don't care why that it is. We now add two interpretative elements that are crucial in Bell's argument. The first is we assume **Bell-realism**: that in each experimental trial the properties $a_0, a_1, b_0, b_1$ are all physically real properties of our quantum system, regardless of if they are measured, and the act of measuring that system reveals the true value of some of these variables (whichever pair of properties has been selected). The second is we assume **Bell-locality**: that because $\hat{A}_0$ and $\hat{A}_1$ act only on system A and $\hat{B}_0$ and $\hat{B}_1$ act only on system B, they each only affect data in their respective system, and so the property associated to the product $\hat{A}_\mu \hat{B}_\nu$ is simply $a_\mu b_\nu$. Under the assumption of Bell-realism, in each experiment trial (which we now view as an outcome in probability space), there is a true value of all of $a_0, a_1, b_0, b_1$ and so we can assign a probability of occurrence in our experiment to each of the 16 possible combinations of values of the properties, with those probabilities the same regardless of the measurement setting. Automatically then, (2.5.22) must hold. However, by the Born rule and Bell-locality, $\langle \hat{A}_0 \hat{B}_0 \rangle = \mathbb{E}[a_0 b_0]$ and similarly so for the other combinations of operators, and so we have by (2.5.19) that the left hand side of (2.5.22) must take the value $2\sqrt{2}$, violating the right hand side of (2.5.22). We have arrived at a contradiction. Therefore, either Bell-realism or Bell-locality must be incorrect. That is Bell's theorem.

You will notice that the above argument had absolutely nothing to do with hidden variables. Hidden variables are typically invoked in Bell-type discussions to explain why when the system of interest is prepared identically different measurement results are achieved. This argument goes by saying perhaps this is some fully deterministic theory in which $a_0, a_1, b_0, b_1$ are 4 among its variables (the others being "hidden" i.e. not measured in this experiment). Due to the imperfect knowledge of the other variables in each experimental trial, the properties $a_0, a_1, b_0, b_1$ have some probability distribution. But, because no probability distribution can possibly be consistent with (2.5.19) (when interpreted with Bell-realism and Bell-locality) one is forced to conclude that local real hidden variables cannot correctly account for the randomness of quantum mechanics. This argument is entirely correct, if you assume Bell-realism and Bell-locality are true. But, the apparent conceptual tension was already evident last paragraph before we mentioned anything about hidden variables. The true content of Bell's theorem is that **quantum mechanics is not equivalent to the combination of probability theory, Bell-realism, Bell-locality**. Some have interpreted Bell's theorem to prove that quantum mechanics is nonlocal, although you will notice that (1) it is not immediately clear that Bell-locality and locality in the sense you would mean in say relativity are at all the same thing and (2) there is another alternative.

The simplest resolution to the problem of Bell's theorem is just to accept that **Bell-realism is not true**. Discarding Bell-realism does not make you give up ordinary philosophical realism. Bell-realism assumes that the possible measurement outcomes $a_0, b_0, a_1, b_1$ are real properties of the quantum system being measured, regardless of which ones actually are measured. However, as argued in section 2.4, those outcomes of individual experimental trials are actually properties of the experimental apparatus which in a well-designed experiment become correlated with part of the incoming state of the system being measured through their interaction. Because $\hat{A}_0$ and $\hat{A}_1$ do not commute there is no way to prepare any interaction Hamiltonian operator $\hat{H}_{\text{int}}$ capable of ideally measuring $\hat{A}_0$ and $\hat{A}_1$ in the same trial. Because $a_0$ and $a_1$ are defined through this measurement interaction, there is thus no way to define both $a_0$ and $a_1$ in a single trial. This is incompatible with Bell-realism. Bell-realism assumes that $a_0$ and $a_1$ each have an unknown true value, whether measured or not, and that the experiment measures one of them based on its detector setting. In the thermal interpretation, the quantum system being measured has a real physical state (represented by its single trial density matrix) and setting the experiment (choosing the detector setup) to measure $a_0$ or $a_1$ determines which one will be defined, while the other *does not even exist* in that setting (because there is no

detector variable corresponding to it); the defined variable then reveals some partial information about the actual state of the system of interest (the incoming density matrix). So, within the thermal interpretation it is no surprise it is impossible to design a joint probability distribution for $a_0, a_1, b_0, b_1$ able to describe the outcome of the experiment. The interacting mechanics of measurement never has more than two of those four variables even being well-defined in the same trial! The correct mathematical description of probability in the thermal interpretation is that there is a probability space appropriate for each experimental setting (choice of things to measure, called a **context**).

That concept of **contextuality** is essential to understanding the meaning of Bell's theorem. Contextuality is the property that the set of possible outcomes of experimental results is determined by the experimental setting (i.e. which variables you can define values for is determined by your interaction Hamiltonian). Contextuality is the conceptual opposite of Bell-realism. Within the thermal interpretation one still has ordinary realism, i.e. there is a a physical state of the system which is real (specifically, the single trial density matrix of the full interacting system and detector), and one is completely able to produce the results of Bell's theorem without contradiction because Bell-realism is not a property of the thermal interpretation. Bell-realism holds that the unknown hypothetical outcomes of experiments not performed are real which is experimentally vacuous but theoretically a disaster. The simple answer is to just understand that the state (the density matrix) is real and the results of measurement are consequently contextual. Looking back at Bell's inequalities with contextuality understood, the random variable $a_0$ only exists when your experiment is set to measure the property associated to $\hat{A}_0$, and similarly so for the other properties. In each of the four possible experimental settings, one can then fully define a probability distribution using the Born rule appropriate for the two variables which are defined in that context and then one sees by a brief calculation that indeed:

$$\mathbb{E}_{00}[a_0 b_0] + \mathbb{E}_{01}[a_0 b_1] + \mathbb{E}_{10}[a_1 b_0] - \mathbb{E}_{11}[a_1 b_1] = 2\sqrt{2} \tag{2.5.23}$$

exactly as is required for consistency. In the above the subscripts on the expectation indicate which experimental context is used (and thus the appropriate Born rule distribution and probability space). There is no joint distribution or probability space for all four variables because there is no experimental context in which all four variables exist. It is amazing just how confused the common discussion around Bell's theorem is. Please again refer to [39] for more analysis and for references to other works that hath understanding.

### Exercises

**Problem 2.10: ⋆ The most important homework question.**

(a) Go take a talk through nature with someone you like, even if that someone is just yourself.

(b) Go pet a cat, if you can.

## 2.6 Quantum Statistical Mechanics

Fundamentally, we got started in the business of quantum mechanics because of statistical mechanics. In the case of particles, the macroscopic properties of matter are inconsistent with the existence of point particles. We have since replaced point particles with Schrödinger wave packets, which are able to explain the macroscopic properties of matter. As well, the inability for classical field theories to reach thermodynamic equilibrium (illustrated by the ultraviolet catastrophe in black body radiation) has shown us that we need to alter classical field theory as well. Its replacement is quantum field theory, which we will delve fully into next chapter. However, the very little bit of quantum field theory we developed in section 1.10 will be enough for us to properly solve the ultraviolet catastrophe, thus showing the power of quantum mechanics to solve a fundamental problem of classical mechanics. It is time to learn quantum statistical mechanics.

In quantum statistical mechanics we are interested in the statistical expected values of the physical properties of interest of an ensemble of "equivalent" systems, just as in classical statistical mechanics. We again can have this ensemble originating from many trials of an experiment where we only have imperfect control over the initial conditions (and thus use a distribution to model our ignorance of them) or from us uncontrolledly selecting a subsystem from a single system comprised of many mechanically identical subsystems (one particle in a box of gas). As we have established in quantum mechanics, all symmetry generators from Noether's theorem are extensive and take the form of the operator summary of a self-adjoint operator. Further, we have already shown that the statistical expectation of an operator summary is still of the form of an operator summary but now with that summary taken by using the statistical expectation of

the system's density matrix. Thus, in quantum statistical mechanics all of the physical properties of interest are characterized by the statistical expectation of the system's density matrix, which we will just call $\hat{\rho}$ as it acts as the effective density matrix of the full ensemble. Of course, the full density matrix of our system satisfies the von Neumann equation (2.1.63).

### 2.6.1 First Law of Thermodynamics

The fundamental results of thermodynamics carry over to quantum mechanics with hardly any change. Let us begin with the first law. Consider our system in the active picture so that now the density matrix is constant $\hat{\rho}_0$ and the coordinate and momentum operators evolve with time. Suppose our system has free Hamiltonian operator $\hat{H}$ and is subject to some external forces which cause modification to the evolution of the momentum operators as usual:

$$\frac{d\hat{Q}^a}{dt} = \frac{1}{i\hbar}[\hat{Q}^a, \hat{H}], \qquad \frac{d\hat{P}_a}{dt} = \frac{1}{i\hbar}[\hat{P}_a, \hat{H}] + \hat{F}_a^{\text{ext}}. \tag{2.6.1}$$

We now want to compute the time evolution of the expected energy of the system. We write our Hamiltonian operator in the form of (2.2.39):

$$\hat{H} = \sum_{\substack{m=0\\n=0}}^{\infty,\infty} H^{(m,n)}{}_{a_1\dots a_m}{}^{b_1\dots b_n} \text{Sym}(\hat{Q}^{a_1}, ..., \hat{Q}^{a_m}, \hat{P}_{b_1}, ..., \hat{P}_{b_n}) \tag{2.6.2}$$

where we have taken the point $(q, p) = (0, 0)$ and given an abbreviated name to the coefficients of the expansion. Those coefficients are constant because the underlying system has time translational symmetry. Now we can directly compute:

$$\frac{d\hat{H}}{dt} = \sum_{\substack{m=0\\n=0}}^{\infty,\infty} H^{(m,n)}{}_{a_1\dots a_m}{}^{b_1\dots b_n} \left( m\text{Sym}\left(\frac{d\hat{Q}^{a_1}}{dt}, ..., \hat{Q}^{a_m}, \hat{P}_{b_1}, ..., \hat{P}_{b_n}\right) + n\text{Sym}\left(\hat{Q}^{a_1}, ..., \hat{Q}^{a_m}, \frac{d\hat{P}_{b_1}}{dt}, ..., \hat{P}_{b_n}\right)\right) \tag{2.6.3}$$

$$= \frac{1}{i\hbar}[\hat{H}, \hat{H}] + \sum_{\substack{m=0\\n=0}}^{\infty,\infty} H^{(m,n)}{}_{a_1\dots a_m}{}^{b_1\dots b_n} n\text{Sym}\left(\hat{Q}^{a_1}, ..., \hat{Q}^{a_m}, \hat{F}_{b_1}^{\text{ext}}, ..., \hat{P}_{b_n}\right). \tag{2.6.4}$$

This last series can be understood as effectively differentiating $\hat{H}$ as a function with respect to $\hat{P}_a$ and inserting $\hat{F}_a^{\text{ext}}$ at each place $\hat{P}_a$ was removed. Therefore, we define as a piece of notation:

$$\left(\hat{F}_a^{\text{ext}} \frac{\partial}{\partial \hat{P}_a}\right) \hat{H} = \sum_{\substack{m=0\\n=0}}^{\infty,\infty} H^{(m,n)}{}_{a_1\dots a_m}{}^{b_1\dots b_n} n\text{Sym}\left(\hat{Q}^{a_1}, ..., \hat{Q}^{a_m}, \hat{F}_{b_1}^{\text{ext}}, ..., \hat{P}_{b_n}\right) \tag{2.6.5}$$

$$\implies \frac{d\hat{H}}{dt} = \left(\hat{F}_a^{\text{ext}} \frac{\partial}{\partial \hat{P}_a}\right) \hat{H}. \tag{2.6.6}$$

Immediately then, because $\text{E} = \langle \hat{H} \rangle$:

$$\frac{d\text{E}}{dt} = \langle \left(\hat{F}_a^{\text{ext}} \frac{\partial}{\partial \hat{P}_a}\right) \hat{H} \rangle. \tag{2.6.7}$$

If we now divide our variables as usual in statistical mechanics into microscopic $m$ and macroscopic $M$ degrees of freedom we have (simply by expanding the sum):

$$\frac{d\text{E}}{dt} = \langle \left(\hat{F}_{ma}^{\text{ext}} \frac{\partial}{\partial \hat{P}_{ma}}\right) \hat{H} \rangle + \langle \left(\hat{F}_{Ma}^{\text{ext}} \frac{\partial}{\partial \hat{P}_{Ma}}\right) \hat{H} \rangle. \tag{2.6.8}$$

So, just as in classical mechanics the contributions to the change in energy decompose nicely into effects due to macroscopic external forces and effects due to microscopic external forces. Just as in classical mechanics then, these contributions provide the fundamental definitions of work and heat:

$$\frac{đ\text{W}}{dt} = \langle \left(\hat{F}_{Ma}^{\text{ext}} \frac{\partial}{\partial \hat{P}_{Ma}}\right) \hat{H} \rangle, \qquad \frac{đ\mathcal{Q}}{dt} = \langle \left(\hat{F}_{ma}^{\text{ext}} \frac{\partial}{\partial \hat{P}_{ma}}\right) \hat{H} \rangle \tag{2.6.9}$$

and the first law of thermodynamics is entirely unchanged:

$$\frac{d\text{E}}{dt} = \frac{đ\text{W}}{dt} + \frac{đ\mathcal{Q}}{dt}. \tag{2.6.10}$$

### 2.6.2 Thermodynamic Equilibrium

The notion of thermodynamic equilibrium and the generalized Boltzmann distribution also carry over very nicely into quantum mechanics. Let us return to the passive picture and consider the evolution of the density matrix under the von Neumann equation. Let $\hat{\bar{\rho}}$ indicate the time-averaged density matrix:

$$\hat{\bar{\rho}} = \lim_{\tau\to\infty} \frac{1}{\tau}\int_0^\tau \hat{\rho}(t)dt. \tag{2.6.11}$$

$\hat{\bar{\rho}}$ captures the long term behavior of our system. Using the von Neumann equation, we have that:

$$[\hat{\bar{\rho}}, \hat{H}] = \lim_{\tau\to\infty} \frac{1}{\tau}\int_0^\tau i\hbar\frac{d\hat{\rho}(t)}{dt}dt = i\hbar \lim_{\tau\to\infty} \frac{\hat{\rho}(\tau)-\hat{\rho}(0)}{\tau}. \tag{2.6.12}$$

The numerator will always remain a bounded operator while the denominator goes to 0 and so in fact:

$$[\hat{\bar{\rho}}, \hat{H}] = 0. \tag{2.6.13}$$

The steady state density matrix must commute with the Hamiltonian operator.

Suppose our system has $K$ independent extensive conserved quantities $G_1, ...G_K$. In quantum mechanics, these can be written as the operator summaries of $K$ self-adjoint operators $\hat{G}_1, ..., \hat{G}_K$ and each of these operators must commute with the Hamiltonian operator in order to have the conservation law:

$$[\hat{G}_k, \hat{H}] = 0. \tag{2.6.14}$$

By independence we require that these quantities cannot be written in terms of each other and that their operators commute:

$$[\hat{G}_k, \hat{G}_l] = 0. \tag{2.6.15}$$

Because we assume these are a complete list of our system's conserved quantities, in order for $\hat{\bar{\rho}}$ to commute with $\hat{H}$ it must be a function of the operators $\hat{G}_1, ...\hat{G}_K$. That is, for some functional form $F$:

$$\hat{\bar{\rho}} = F(\hat{G}_1, ..., \hat{G}_K). \tag{2.6.16}$$

The particulars of this functional form are fixed by the initial conditions of the system. However, if our experimental design is unbiased in certain ways we can determine a specific function $F$ just like in classical statistical mechanics. For example, if we always prepare our system to identically have energy E and to be totally unbiased otherwise on average, then necessarily the only one of the $\hat{G}_k$ that $F$ depends on is $\hat{H}$ itself and $\hat{\bar{\rho}}$ will take the form:

$$\hat{\bar{\rho}} = N\delta(\hat{H} - \mathrm{E}) \tag{2.6.17}$$

where $N$ is a normalization factor. Because $\hat{\bar{\rho}}$ must have trace 1 we can determine this normalization factor in terms of the density of states $g(\mathrm{E})$:

$$g(\mathrm{E}) = \mathrm{Tr}(\delta(\hat{H} - \mathrm{E})) \tag{2.6.18}$$

and in doing so we achieve the quantum microcanonical ensemble:

$$\boxed{\hat{\rho}_{\text{micro}} = \frac{1}{g(\mathrm{E})}\delta(\hat{H} - \mathrm{E}).} \tag{2.6.19}$$

Alternatively, we can say that if we bring together two subsystems which are noninteracting but under the same thermodynamic conditions then they should be statistically independent of each other and their extensive conserved quantities should add. Independence requires that their density matrices multiply by (2.1.110):

$$\hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2. \tag{2.6.20}$$

Imposing this on (2.6.16), the additivity of extensive conserved quantities then means:

$$F(\hat{\vec{G}}_1 \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\vec{G}}_2) = F(\hat{\vec{G}}_1) \otimes F(\hat{\vec{G}}_2) \tag{2.6.21}$$

where we have formed a vector $\hat{\vec{G}} = (\hat{G}_1, ..., \hat{G}_K)$. By the identical calculation leading to Orig. (5.5.44), we have that for some real numbers $\beta_a$ and $Z$:

$$F(\hat{\vec{G}}) = \frac{e^{-\beta_a \hat{G}^a}}{Z}. \tag{2.6.22}$$

Thus we recover an analog of the generalized Boltzmann distribution as the generalized canonical ensemble in

quantum mechanics. The most common special case of course is when all of our conserved quantities other than the energy are 0 on average and so happen to have vanishing multipliers $\beta_a$, leaving us with the usual canonical ensemble:

$$\boxed{\hat{\rho}_{\text{canon}} = \frac{e^{-\beta \hat{H}}}{Z}.} \tag{2.6.23}$$

The temperature is (in direct analogy to classical statistical mechanics) defined as the reciprocal of the $\beta$ factor corresponding to the energy:

$$\boxed{T = \frac{1}{\beta}} \tag{2.6.24}$$

and the partition function $Z$ is fixed as always by requiring appropriate normalization of the distribution:

$$\text{Tr}(\hat{\rho}_{\text{canon}}) \stackrel{!}{=} 1 \implies \boxed{Z = \text{Tr}(e^{-\beta \hat{H}}).} \tag{2.6.25}$$

It is clear how (2.6.18) and (2.6.25) reduce in the classical limit to the classical definitions of the density of states and the partition function by using (2.2.123) and that in the classical limit the Wigner function associated to $H$ will become the classical Hamiltonian $\mathcal{H}$. Again this explains the inclusion of factors $2\pi\hbar$ all the way back in the phase space measure in Orig. (5.2.4); those factors were put there to make this classical limit as simple as possible due to the factors of $2\pi\hbar$ which appear automatically in the Weyl transform.

The familiar formula Orig. (5.5.20) for the expected energy in the canonical ensemble even holds in quantum mechanics! By an essentially identical calculation:

$$\frac{\partial Z}{\partial \beta} = \frac{\partial}{\partial \beta}\text{Tr}(e^{-\beta \hat{H}}) = -\text{Tr}(\hat{H} e^{-\beta H}) = -Z\text{Tr}(\hat{H}\hat{\rho}) = -\text{E}Z \implies \boxed{\text{E} = -\frac{\partial}{\partial \beta} \ln Z.} \tag{2.6.26}$$

Unsurprisingly given how much of the rest of thermodynamic structure is shared between classical and quantum mechanics, the free energy is defined in exactly the same way between them:

$$\boxed{\mathcal{F} = -T \ln Z.} \tag{2.6.27}$$

The $T$ derivative of $\mathcal{F}$ immediately then gives (using (2.6.24) and (2.6.26)):

$$\text{E} = \mathcal{F} - T\frac{\partial \mathcal{F}}{\partial T}. \tag{2.6.28}$$

### 2.6.3 Entropy and the Second Law of Thermodynamics

The classical formula for the entropy of a phase space density Orig. (5.6.64) can't directly apply to the density matrix because the density matrix is not a function on the phase space. An obvious suggestion would be to take Orig. (5.6.64) and replace the phase space density by the Wigner function associated to the density matrix, however this is not a valid definition of entropy because the Wigner function associated to the density matrix can be negative, rendering the logarithm meaningless (on real numbers). Also, we don't always want to commit ourselves to using configuration Hilbert spaces. Quantum mechanics is formulated on an entirely generic Hilbert space, and we should try to keep it that way. Because the density matrix is a nonnegative self-adjoint operator of unit trace, we can diagonalize it into the form:

$$\hat{\rho} = \sum_n p_n \hat{\pi}_n \tag{2.6.29}$$

where the projectors $\hat{\pi}_n$ and eigenvalues $p_n$ satisfy:

$$\pi_n \pi_k = \delta_{nk}\pi_n, \qquad \sum_n \hat{\pi}_n = \mathbb{1}, \qquad p_n \geq 0, \qquad \sum_n p_n = 1. \tag{2.6.30}$$

By (2.1.80), the density matrix (2.6.29) has the same statistical expectations as you would get from treating the $\hat{\pi}_n$ as a list of density matrices and randomly sampling from them with associated probabilities $p_n$. So, a possible natural definition of the entropy of the density matrix is then:

$$S = -\sum_n p_n \ln p_n \tag{2.6.31}$$

as this is the information entropy you would get by using that probability distribution. This entropy is known as the von Neumann entropy and we will now show that it does in fact have the desired properties we would

like the entropy to have. Recall that a function of a diagonalizable operator is defined as the action of that function on the operator's eigenvalues in any of its diagonal bases. With this in mind, we can immediately rewrite the von Neumann entropy as:

$$\boxed{S = -\mathrm{Tr}(\hat{\rho}\ln\hat{\rho}).} \tag{2.6.32}$$

Because the von Neumann entropy is an information entropy of a probability distribution, it is automatically nonnegative:

$$S \geq 0. \tag{2.6.33}$$

This is superior to the situation in classical mechanics where the entropy can become arbitrarily negative (because differential entropy is cursed). The explanation for how the possible negativity arises in classical mechanics is that the classical limit of (2.6.32) often gives an infinite result (as happened originally in the derivation of the entropy of a continuous random variable) and the differential entropy of classical mechanics comes from setting an arbitrary 0 point (dropping an overall infinite constant a la $-\ln\hbar$ in the $\hbar \to 0^+$ limit).

In classical statistical mechanics, the various ensembles of interest have the maximum possible entropy consistent with their relevant constraints. The very same is true in quantum mechanics. Let us show this for the canonical ensemble, as for the other ensembles the calculations follow essentially the same steps. We wish to maximize the $S$ for a $\hat{\rho}$ with a given expected energy E and so consider the variations of:

$$S(\hat{\rho}) = -\mathrm{Tr}(\hat{\rho}\ln\hat{\rho}) + \alpha(1 - \mathrm{Tr}(\hat{\rho})) + \beta(\mathrm{E} - \mathrm{Tr}(\hat{H}\hat{\rho})) \tag{2.6.34}$$

with the Lagrange multipliers $\alpha$ and $\beta$. Varying with respect to $\hat{\rho}$ yields:

$$\frac{\partial S}{\partial \hat{\rho}} = -\ln\hat{\rho} - \mathbb{1} - \alpha\mathbb{1} - \beta\hat{H} \stackrel{!}{=} 0 \tag{2.6.35}$$

where the derivative of a function with respect to an operator is defined in an arbitrary orthonormal basis as:

$$\frac{\partial S}{\partial \hat{\rho}} = \sum_{mn} \frac{\partial S}{\partial \rho_{nm}} |m\rangle\langle n| \tag{2.6.36}$$

and where the components $\rho_{nm}$ are defined as usual by:

$$\rho_{nm} = \langle n|\hat{\rho}|m\rangle. \tag{2.6.37}$$

Continuing from (2.6.35) we have:

$$\hat{\rho} = e^{-\alpha}e^{-\beta\hat{H}} \tag{2.6.38}$$

which upon normalization just recovers the canonical ensemble. The second derivative of the entropy again reveals that this is the maximum value attained.

Another foundational argument in favor of the suitability of von Neumann's definition of the entropy of a density matrix is that it precisely recovers the usual relationship between the entropy and the free energy. In particular, for the canonical ensemble:

$$S = -\mathrm{Tr}(\hat{\rho}\ln\hat{\rho}) = -\mathrm{Tr}(\hat{\rho}(-\beta\hat{H} - \mathbb{1}\ln Z)) = \beta\mathrm{E} + \ln Z = \frac{\mathrm{E} - \mathcal{F}}{T} = -\frac{\partial\mathcal{F}}{\partial T} \tag{2.6.39}$$

totally identically to classical statistical mechanics.

Like classical entropy, the von Neumann entropy is extensive. If two systems are independent so that their combined density matrix factorizes as $\hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2$ then:

$$\hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2 \implies \ln\hat{\rho} = \ln\hat{\rho}_1 \otimes \mathbb{1} + \mathbb{1} \otimes \ln\hat{\rho}_2 \implies S(\hat{\rho}) = S(\hat{\rho}_1) + S(\hat{\rho}_2). \tag{2.6.40}$$

Von Neumann's entropy is also subadditive, just like classical information entropy. Unfortunately, the proof is a bit more complicated. First we need a few smaller results.

**Lemma: Convexity of the trace.** For a concave up function $f(x)$, arbitrary self-adjoint $\hat{A}$ and $\hat{B}$, and $t \in (0, 1)$:

$$\mathrm{Tr}(f((1-t)\hat{A} + t\hat{B})) \leq (1-t)\mathrm{Tr}(f(\hat{A})) + t\mathrm{Tr}(f(\hat{B})) \tag{2.6.41}$$

with equality holding if and only if $\hat{A} = \hat{B}$. Proof:

Let $\hat{A}$ and $\hat{B}$ be self-adjoint operators and let $f(x)$ be a differentiable real valued function of a real variable. Define the operator:

$$\hat{C}(t) = (1-t)\hat{A} + t\hat{B} \implies \hat{C}(0) = \hat{A}, \qquad \hat{C}(1) = \hat{B}. \tag{2.6.42}$$

Let $|n(t)\rangle$ be a set of orthonormal eigenvectors of $\hat{C}(t)$ with eigenvalues $\lambda_n(t)$. Then, by perturbation theory around $t$ (from $t$ to $t+\delta t$, say) it follows that:

$$\lambda_n'(t) = \langle n(t)|(\hat{B}-\hat{A})|n(t)\rangle \tag{2.6.43}$$

where the prime indicates the derivative with respect to $t$. So:

$$\mathrm{Tr}(f(\hat{C}(t+\delta t))) = \sum_n f(\lambda_n(t+\delta t)) = \sum_n f(\lambda_n(t)) + \delta t \sum_n \lambda_n'(t) f'(\lambda_n(t)) + \mathcal{O}(\delta t^2) \tag{2.6.44}$$

$$= \sum_n \langle n(t)|f(\hat{C}(t))|n(t)\rangle + \delta t \sum_n \langle n(t)|(\hat{B}-\hat{A})f'(\hat{C}(t))|n(t)\rangle + \mathcal{O}(\delta t^2) \tag{2.6.45}$$

$$\implies \frac{d}{dt}\mathrm{Tr}(f((1-t)\hat{A}+t\hat{B})) = \mathrm{Tr}((\hat{B}-\hat{A})f'((1-t)\hat{A}+t\hat{B})). \tag{2.6.46}$$

Now consider $f(x)$ to be a concave up function. A concave up function is a function for which for $x \neq y$ and $t \in (0,1)$:

$$f((1-t)x+ty) < (1-t)f(x) + tf(y). \tag{2.6.47}$$

Let $|m'\rangle$ denote any orthonormal basis of our Hilbert space which is not an eigenbasis of $\hat{A}$. Using the concave up-ness of $f(x)$ and the diagonal representation of $\hat{A}$:

$$\mathrm{Tr}(f(\hat{A})) = \sum_m \langle m'|f(\hat{A})|m'\rangle = \sum_{mn} f(\lambda_n(0))\langle m'|n(0)\rangle\langle n(0)|m'\rangle \tag{2.6.48}$$

$$> \sum_m f\left(\sum_n a_n |\langle m'|n(0)\rangle|^2\right) = \sum_m f\left(\langle m'|\hat{A}|m'\rangle\right). \tag{2.6.49}$$

If we instead have $|m'\rangle = |n_A\rangle$ we get exact equality immediately and so in general:

$$\mathrm{Tr}(f(\hat{A})) \geq \sum_m f(\langle m'|\hat{A}|m'\rangle) \tag{2.6.50}$$

for any concave up $f(x)$, self-adjoint $\hat{A}$, and generic orthonormal basis $|m'\rangle$ with equality holding if and only if $|m'\rangle$ is an eigenbasis of $\hat{A}$. Let us now apply this result with $\hat{A}$ replaced by $\hat{C}(t)$. Let us first assume $\hat{A}$ and $\hat{B}$ commute but $\hat{A} \neq \hat{B}$. Then, we can select $|n(t)\rangle$ to be a simultaneous eigenbasis of them and because $\hat{A} \neq \hat{B}$ there is at least one $n$ so that $\langle n(t)|\hat{A}|n(t)\rangle \neq \langle n(t)|\hat{B}|n(t)\rangle$. With this, (2.6.47), (2.6.50) we have:

$$\mathrm{Tr}(f((1-t)\hat{A}+t\hat{B})) = \sum_n f((1-t)\langle n(t)|\hat{A}|n(t)\rangle + t\langle n(t)|\hat{B}|n(t)\rangle) \tag{2.6.51}$$

$$< \sum_n \Big((1-t)f(\langle n(t)|\hat{A}|n(t)\rangle) + tf(\langle n(t)|\hat{B}|n(t)\rangle)\Big) = (1-t)\mathrm{Tr}(f(\hat{A})) + t\mathrm{Tr}(f(\hat{B})). \tag{2.6.52}$$

Now let us assume $\hat{A}$ and $\hat{B}$ do not commute. Then, the eigenbasis $|n(t)\rangle$ will be an eigenbasis of neither $\hat{A}$ nor $\hat{B}$ and so again using (2.6.47) and (2.6.50):

$$\mathrm{Tr}(f((1-t)\hat{A}+t\hat{B})) = \sum_n f((1-t)\langle n(t)|\hat{A}|n(t)\rangle + t\langle n(t)|\hat{B}|n(t)\rangle) \tag{2.6.53}$$

$$\leq \sum_n \Big((1-t)f(\langle n(t)|\hat{A}|n(t)\rangle) + tf(\langle n(t)|\hat{B}|n(t)\rangle)\Big) < (1-t)\mathrm{Tr}(f(\hat{A})) + t\mathrm{Tr}(f(\hat{B})). \tag{2.6.54}$$

If we had instead chosen $\hat{B} = \hat{A}$ then plainly $\mathrm{Tr}(f((1-t)\hat{A}+t\hat{B})) = (1-t)\mathrm{Tr}(f(\hat{A})) + t\mathrm{Tr}(f(\hat{B}))$. So, in all cases the lemma holds. ■

**Lemma: Klein's Inequality.** For any differentiable concave up function $f(x)$ and self-adjoint operators $\hat{A}$ and $\hat{B}$:

$$\mathrm{Tr}(f(\hat{B})) - \mathrm{Tr}(f(\hat{A})) \geq \mathrm{Tr}((\hat{B}-\hat{A})f'(\hat{A})) \tag{2.6.55}$$

with equality if and only if $\hat{A} = \hat{B}$. Proof:

Consider $\hat{B} \neq \hat{A}$ and from our function $f(x)$ define the function $F(t)$:

$$F(t) = \mathrm{Tr}(f((1-t)\hat{A} + t\hat{B})). \tag{2.6.56}$$

It immediately follows for $t \in (0,1)$ and $x, y \in [0,1]$ with $y \neq x$ from the convexity of the trace that $F(t)$ is also concave up:

$$(1-t)F(x) + tF(y) = (1-t)\mathrm{Tr}(f((1-x)\hat{A} + x\hat{B})) + t\mathrm{Tr}(f((1-y)\hat{A} + y\hat{B})) \tag{2.6.57}$$

$$> \mathrm{Tr} f((1-t)(1-x)\hat{A} + (1-t)x\hat{B} + t(1-y)\hat{A} + ty\hat{B}) = F((1-t)x + ty). \tag{2.6.58}$$

As well, define the function $g(t)$:

$$g(t) = \frac{F(t) - F(0)}{t}. \tag{2.6.59}$$

Because $F(t)$ is concave up, it follows that $g(t)$ is increasing. We can see this by evaluating the concave up condition on $F(t)$ in terms of $x = 0$, $y = b$, $t = \frac{a}{b}$ (note $a < b$ because $t < 1$):

$$b > a \implies \frac{F(b) - F(0)}{b} > \frac{F(a) - F(0)}{a} \implies g(b) > g(a). \tag{2.6.60}$$

By applying the concave up condition at $x = 0$, $y = 1$:

$$F(1) - F(0) > g(t). \tag{2.6.61}$$

Because $g(t)$ is an increasing function of $t$ and using its explicit definition, we also have:

$$F(1) - F(0) > g(t) > \lim_{t \to 0^+} g(t) = F'(0). \tag{2.6.62}$$

Let us now apply (2.6.46) to simplify this inequality into:

$$\mathrm{Tr}(f(\hat{B})) - \mathrm{Tr}(f(\hat{A})) > \mathrm{Tr}((\hat{B} - \hat{A})f'(\hat{A}). \tag{2.6.63}$$

Plainly, if we had used $\hat{B} = \hat{A}$ then both sides of (2.6.63) become 0 and we get equality. ■

**Lemma: Nonnegativity of quantum relative information.** The quantum relative information $\mathcal{I}(\hat{\mathfrak{p}}, \hat{\mathfrak{q}})$ between the density matrices $\hat{\mathfrak{p}}$ and $\hat{\mathfrak{q}}$ defined by:

$$\mathcal{I}(\hat{\mathfrak{p}}|\hat{\mathfrak{q}}) = \mathrm{Tr}(\hat{\mathfrak{p}} \ln \hat{\mathfrak{p}}) - \mathrm{Tr}(\hat{\mathfrak{p}} \ln \hat{\mathfrak{q}}) \tag{2.6.64}$$

is nonnegative and is 0 if and only if $\hat{\mathfrak{p}} = \hat{\mathfrak{q}}$. (Note that $\mathcal{I}$ is the natural operator generalization of Orig. (5.6.9) for the relative information between two probability distributions.) Proof:

Consider the function:

$$f(x) = x \ln x. \tag{2.6.65}$$

Plainly for $x > 0$, $f(x)$ is differentiable. As well, its second derivative is positive (simply $1/x$), which we can then integrate from $a > 0$ to $b > a$ to show that $f'(x)$ is an increasing function:

$$f''(x) = \frac{1}{x} > 0 \implies f'(b) - f'(a) > 0. \tag{2.6.66}$$

We can now take this condition and replace $b$ by $x$ and then integrate from $a$ to $b$ or replace $a$ by $x$ and then integrate from $a$ to $b$. Doing each of these and combining the results gives:

$$f'(a) < \frac{f(b) - f(a)}{b - a} < f'(b). \tag{2.6.67}$$

If we now consider $c > b$, we have:

$$\frac{f(c) - f(a)}{c - a} = \frac{\frac{f(c)-f(b)}{c-b}(c-b) + \frac{f(b)-f(a)}{b-a}(b-a)}{c-a} > \frac{f'(b)(c-b) + \frac{f(b)-f(a)}{b-a}(b-a)}{c-a} \tag{2.6.68}$$

$$> \frac{\frac{f(b)-f(a)}{b-a}(c-b) + \frac{f(b)-f(a)}{b-a}(b-a)}{c-a} = \frac{f(b) - f(a)}{b-a}. \tag{2.6.69}$$

Taking $a = x$, $b = (1-t)x + ty$, and $c = y$ with $y > x$ and $t \in (0,1)$ we find with a small amount of algebra that indeed $f(x)$ is concave up:

$$f((1-t)x + ty) < (1-t)f(x) + tf(y). \tag{2.6.70}$$

Now consider two different valid density matrices $\hat{\mathfrak{q}}$ and $\hat{\mathfrak{p}}$. Because $f(x)$ is concave up, we can apply Klein's

inequality to find:

$$\mathrm{Tr}(\hat{\mathfrak{p}}\ln\hat{\mathfrak{p}}) - \mathrm{Tr}(\hat{\mathfrak{q}}\ln\hat{\mathfrak{q}}) \geq \mathrm{Tr}((\hat{\mathfrak{p}}-\hat{\mathfrak{q}})(\ln\hat{\mathfrak{q}}+\mathbb{1})) \implies \mathrm{Tr}(\hat{\mathfrak{p}}\ln\hat{\mathfrak{p}}) - \mathrm{Tr}(\hat{\mathfrak{p}}\ln\hat{\mathfrak{q}}) \geq 0 \tag{2.6.71}$$

with equality holding if and only if $\hat{\mathfrak{q}} = \hat{\mathfrak{p}}$. ■

**Proof of the subadditivity of entropy:**

Consider a quantum system which decomposes into two subsystems. Let $\hat{\rho}$ be the density matrix of the full system and let $\hat{\rho}_1$ and $\hat{\rho}_2$ be the marginal density matrices of the two subsystems. As well, let $\hat{\rho}_{\text{indep}}$ be the tensor product of $\hat{\rho}_1$ and $\hat{\rho}_2$:

$$\hat{\rho}_1 = \mathrm{Tr}_2(\hat{\rho}), \qquad \hat{\rho}_2 = \mathrm{Tr}_1(\hat{\rho}), \qquad \hat{\rho}_{\text{indep}} = \hat{\rho}_1 \otimes \hat{\rho}_2. \tag{2.6.72}$$

Of course, $\hat{\rho}_{\text{indep}}$ is the density matrix the system would have if the two subsystems were completely independent and noninteracting. Then:

$$S(\hat{\rho}_1) + S(\hat{\rho}_2) - S(\hat{\rho}) = -\mathrm{Tr}_1(\hat{\rho}_1\ln\hat{\rho}_1) - \mathrm{Tr}_2(\hat{\rho}_2\ln\hat{\rho}_2) + \mathrm{Tr}(\hat{\rho}\ln\hat{\rho}) \tag{2.6.73}$$

$$= -\mathrm{Tr}(\hat{\rho}\ln(\hat{\rho}_1\otimes\mathbb{1})) - \mathrm{Tr}(\hat{\rho}\ln(\mathbb{1}\otimes\hat{\rho})) + \mathrm{Tr}(\hat{\rho}\ln\hat{\rho}) \tag{2.6.74}$$

$$= \mathrm{Tr}(\hat{\rho}\ln\hat{\rho}) - \mathrm{Tr}(\hat{\rho}\ln\hat{\rho}_{\text{indep}}) = \mathcal{I}(\hat{\rho}|\hat{\rho}_{\text{indep}}) \geq 0 \tag{2.6.75}$$

with equality holding if and only if $\hat{\rho} = \hat{\rho}_{\text{indep}}$. After one final step we have, as promised:

$$\boxed{S(\hat{\rho}) \leq S(\hat{\rho}_1) + S(\hat{\rho}_2)} \tag{2.6.76}$$

with equality if and only if $\hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2$. ■

Just like in classical statistical mechanics, the entropy is conserved by Hamiltonian time evolution. Under Hamiltonian time evolution, the density matrix has the simple time dependence:

$$\hat{\rho}(t) = \hat{U}(t,t_0)\hat{\rho}(t_0)\hat{U}^\dagger(t,t_0). \tag{2.6.77}$$

Any continuous function of $\hat{\rho}(t)$ then satisfies:

$$f(\hat{\rho}(t)) = \hat{U}(t,t_0)f(\hat{\rho}(t_0))\hat{U}^\dagger(t,t_0) \tag{2.6.78}$$

(to see this, first use that any continuous function can be arbitrarily well approximated by an analytic function and that for an analytic function this follows automatically by using its power series expression). So, the entropy is constant by the cyclicity of the trace:

$$S(t) = -\mathrm{Tr}(\hat{\rho}(t)\ln\hat{\rho}(t)) = -\mathrm{Tr}(\hat{U}(t,t_0)\hat{\rho}(t_0)\ln\hat{\rho}(t_0)\hat{U}^\dagger(t,t_0)) = -\mathrm{Tr}(\hat{\rho}(t_0)\ln\hat{\rho}(t_0)) = S(t_0). \tag{2.6.79}$$

Exactly as in classical mechanics, the entropy of an isolated system only changes due to unaccounted for forces (which we could add by time evolving using (2.6.1)). The combination of the subadditivity of entropy together with its conservation under Hamiltonian time evolution give that the justification of the second law of thermodynamics established in Orig. (5.7.25) is completely unchanged in quantum mechanics. The second law of thermodynamics carries over exactly.

### 2.6.4 Planck's Law

Now that we have shown how thermodynamics plays nicely with quantum mechanics, let us show how quantum field theory explains Planck's law. Because the electromagnetic Hamiltonian in Orig. (4.4.23) depends only quadratically on the fields and because all of the symmetries of electromagnetism act linearly on the potentials, we know that we can write down a Hamiltonian operator for quantum electromagnetism with all of the correct symmetries by taking that classical Hamiltonian as the Wigner function of the theory. Because of this, all of the coordinate transformations leading to Orig. (5.12.70) remain appropriate in quantum mechanics, with that $\mathcal{H}$ now interpreted as the Wigner function for $\hat{H}$. So, the Hamiltonian operator for quantum mechanical electromagnetic waves in a conducting box of volume side length $\ell$ is then:

$$\hat{H} = \sum_{n_x=1}^{\infty}\sum_{n_y=1}^{\infty}\sum_{n_z=1}^{\infty}\left(\frac{\hat{P}^2_{\vec{n}-1}}{2} + \frac{\hat{P}^2_{\vec{n}+1}}{2} + \frac{\pi^2|\vec{n}|^2}{2\ell^2}(\hat{A}^2_{\vec{n}-1} + \hat{A}^2_{\vec{n}+1})\right) + \sum_{n_y=1}^{\infty}\sum_{n_z=1}^{\infty}\left(\frac{\hat{P}^2_{x0n_yn_z}}{2} + \frac{\pi^2|\vec{n}|^2}{2\ell^2}\hat{A}^2_{x0n_yn_z}\right)$$
$$+ \sum_{n_x=1}^{\infty}\sum_{n_z=1}^{\infty}\left(\frac{\hat{P}^2_{yn_x0n_z}}{2} + \frac{\pi^2|\vec{n}|^2}{2\ell^2}\hat{A}^2_{yn_x0n_z}\right) + \sum_{n_x=1}^{\infty}\sum_{n_y=1}^{\infty}\left(\frac{\hat{P}^2_{zn_xn_y0}}{2} + \frac{\pi^2|\vec{n}|^2}{2\ell^2}\hat{A}^2_{zn_xn_y0}\right). \tag{2.6.80}$$

This is now the form of an infinite number of quantum harmonic oscillators! Because the structure of thermodynamics is totally unchanged by quantum mechanics, the Helmholtz free energy remains extensive and so all of the arguments leading to Orig. (5.12.79) are unchanged. Therefore, we still have the the free energy of quantum electromagnetism is given by:

$$\mathcal{F} = \frac{\mathbb{V}}{\pi^2}\int_0^\infty \mathcal{F}_1(T,\omega)\omega^2 d\omega \tag{2.6.81}$$

where $\mathcal{F}_1(T,\omega)$ is the free energy of a single harmonic oscillator at temperature $T$ of angular frequency $\omega$. However, now it is the free energy of a quantum harmonic oscillator because each mode within (2.6.80) is a qho. The ground state energy of any system is entirely arbitrary and unphysical as we can shift the energy by an overall constant. By convention, we set the ground state energy of a quantum field theory to be 0 by definition. This convention is well-motivated in a Lorentz invariant theory by wanting the vacuum state to be Lorentz invariant, which only happens if its 4-momentum is 0. The other common ground state convention for a single qho, in which the lowest energy is $\frac{\hbar\omega}{2}$ is an obnoxious choice for field theory, because when this is integrated over all modes the result is infinite. There is no reason to add an overall infinite, non-Lorentz invariant constant to your energy, so don't. So, we have that for a single quantum harmonic oscillator we take the Hamiltonian operator:

$$\hat{H} = \hbar\omega a^\dagger \hat{a} \tag{2.6.82}$$

in terms of the creation and annihilation operators of the oscillator. Now using (2.6.25) we have:

$$Z_1 = \mathrm{Tr}(e^{-\beta\hat{H}}) = \sum_{n=0}^\infty e^{-\beta E_n} = \sum_{n=0}^\infty e^{-n\beta\hbar\omega} = \frac{1}{1-e^{-\beta\hbar\omega}} \tag{2.6.83}$$

$$\implies \mathcal{F}_1 = -T\ln Z_1 = T\ln\left(1-e^{-\frac{\hbar\omega}{T}}\right). \tag{2.6.84}$$

This is *exactly* the same free energy that we guessed in Orig. (5.12.102). Therefore, quantum mechanics genuinely explains that formula. From this free energy, we can legitimately follow the rest of the steps of Orig. section 5.12 to derive Planck's law and the Stefan-Boltzmann law. We can also explain Orig. (5.12.112). Using the density of states definition in (2.6.18):

$$g(\mathrm{E}) = \mathrm{Tr}(\delta(\hat{H}-\mathrm{E})) = \sum_n \delta(E_n - \mathrm{E}) = \sum_n (\mathrm{E} - n\hbar\omega) \tag{2.6.85}$$

exactly as desired. Amazingly, quantum mechanics fully resolves the tension between classical field theory and thermodynamics.

### 2.6.5 Bose-Einstein and Fermi-Dirac Statistics

In the next chapter we will show that a Lorentz invariant quantum field theory for a field of integer spin has creation and annihilation operators that satisfy the usual commutation relations but for a field of half-integer spin the creation and annihilation operators satisfy a slightly different *anticommutation* relation. In particular, if $\hat{a}$, $\hat{a}'$ are the annihilation operators for two different modes of angular frequency $\omega, \omega'$ for an integer spin field, we will show:

$$[\hat{a},\hat{a}'] = 0, \qquad [\hat{a},\hat{a}'^\dagger] = 0, \qquad [\hat{a},\hat{a}^\dagger] = [\hat{a}',\hat{a}'^\dagger] = 1. \tag{2.6.86}$$

However, if $\hat{a}$ and $\hat{a}'$ are the annihilation operators for two different modes of angular frequency $\omega, \omega'$ for a half-integer spin field, we will show:

$$[\hat{a},\hat{a}']_+ = 0, \qquad [\hat{a},\hat{a}'^\dagger]_+ = 0, \qquad [\hat{a},\hat{a}^\dagger]_+ = [\hat{a}',\hat{a}'^\dagger]_+ = 1 \tag{2.6.87}$$

where we have defined the anticommutator of two operators:

$$[\hat{A},\hat{B}]_+ = \hat{A}\hat{B} + \hat{B}\hat{A}. \tag{2.6.88}$$

Plainly, the anticommutator is twice the symmetrized product of two operators, but for our thinking it will tend to be nicer to represent it in this $[\cdot,\cdot]_+$ notation so that it appears more similar to the usual commutator. This connection between the type of commutator used and the spin of the field is called the **spin-statistics theorem**. While we will have to wait until next chapter to justify it, it has important implications for quantum statistical mechanics, and so we will study some of those consequences here. Fields which have creation and annihilation operators satisfying the usual commutation relations (2.6.86) are called **bosons** and fields which have creation and annihilation operators satisfying the anticommutation relations (2.6.87)

are called **fermions**. The existence of fermions surprises sane people at first, but you will learn soon that they make perfectly good sense.

As we have seen, the electromagnetic field is bosonic. Let us see the basics of the statistical mechanics of bosonic fields in general. Suppose we have a quadratic field theory which we have transformed into creation and annihilation operators $\hat{a}_k, \hat{a}_k^\dagger$ where $k$ here is a label identifying which field mode is under consideration. Let $E_k$ be the energy of a single particle state corresponding to mode $k$. Then, the Hamiltonian operator for the theory is simply:

$$\hat{H} = \sum_k E_k \hat{a}_k^\dagger \hat{a}_k. \tag{2.6.89}$$

From here we can easily compute the partition function of the theory because the modes for different $k$ commute with each other. Define the number operator $\hat{N}$ which counts the total number of particles:

$$\hat{N} = \sum_k \hat{a}_k^\dagger \hat{a}_k. \tag{2.6.90}$$

Then, in exact analogy of Orig. (5.10.42), the grand canonical ensemble density matrix is given by:

$$\hat{\rho} = \frac{e^{-\beta(\hat{H}-\mu\hat{N})}}{\mathcal{Z}} \tag{2.6.91}$$

where the grand partition function is determined by the normalization of the density matrix:

$$\mathrm{Tr}(\hat{\rho}) = 1 \implies \mathcal{Z} = \mathrm{Tr}(e^{-\beta(\hat{H}-\mu\hat{N})}). \tag{2.6.92}$$

In the classical statistical mechanics of identical particles we found that the occupation number density was very useful. We can define an exactly analogous and equally useful quantity in the quantum statistical mechanics of identical particles. Define the occupation number $n_k$ of mode $k$ as the expected number of particles in mode $k$:

$$n_k = \mathbb{E}[\langle \hat{a}_k^\dagger \hat{a}_k \rangle] = \mathrm{Tr}(\hat{a}_k^\dagger \hat{a}_k \hat{\rho}). \tag{2.6.93}$$

Because the Hamiltonian and number operators in the grand canonical ensemble are sums of commuting terms for each different $k$, we can decompose the density matrix has:

$$\hat{\rho} = \prod_k \left( \frac{e^{-\beta(E_k-\mu)\hat{a}_k^\dagger \hat{a}_k}}{\mathcal{Z}_k} \right) \tag{2.6.94}$$

where each mode's grand partition function is defined by:

$$\mathcal{Z}_k = \mathrm{Tr}_k(e^{-\beta(E_k-\mu)\hat{a}_k^\dagger \hat{a}_k}) \tag{2.6.95}$$

and the trace is only taken over the Hilbert space of mode $k$. With our decomposition of the density matrix, the computation of $n_k$ becomes quite simple:

$$n_k = \frac{\mathrm{Tr}\left(\hat{a}_k^\dagger \hat{a}_k e^{-\beta(E_k-\mu)\hat{a}_k^\dagger \hat{a}_k} \prod_{k' \neq k} e^{-\beta(E_{k'}-\mu)\hat{a}_{k'}^\dagger \hat{a}_{k'}}\right)}{\prod_{k'} \mathcal{Z}_{k'}} = \frac{\mathrm{Tr}(\hat{a}_k^\dagger \hat{a}_k e^{-\beta(E_k-\mu)\hat{a}_k^\dagger \hat{a}_k})}{\mathrm{Tr}(e^{-\beta(E_k-\mu)\hat{a}^\dagger \hat{a}_k})} = \frac{1}{\beta}\frac{\partial}{\partial \mu} \ln \mathcal{Z}_k. \tag{2.6.96}$$

We can easily compute $\mathcal{Z}_k$ because we already know the eigenvalues of the (bosonic) quantum harmonic oscillator:

$$\mathcal{Z}_k = \sum_{n=0}^{\infty} e^{-\beta(E_k-\mu)n} = \frac{1}{1-e^{-\beta(E_k-\mu)}}. \tag{2.6.97}$$

Using this, we immediately have:

$$n_k = \frac{1}{e^{\beta(E_k-\mu)}-1}. \tag{2.6.98}$$

This is the expected number of particles in mode $k$ at a given temperature and chemical potential. Equation (2.6.98) is called the **Bose-Einstein** occupation formula and the particles of bosonic fields are said to satisfy Bose-Einstein statistics.

It is useful to define the occupation number operator $\hat{n}$ as an operator on the single particle Hilbert space of the field theory by:

$$\hat{n} = \sum_k n_k |k\rangle\langle k| \tag{2.6.99}$$

where $|k\rangle$ indicates a single particle state in mode $k$. Automatically then, if $\hat{H}_1$ is the effective single particle Hamiltonian operator of the field theory, we have that the total expected number of particles and total expected energy are given by:

$$\mathrm{N} = \mathrm{Tr}(\hat{n}), \qquad \mathrm{E} = \mathrm{Tr}(\hat{H}_1\hat{n}). \tag{2.6.100}$$

Let $n(Q,P)$ be the Wigner function associated to $\hat{n}$ and $Q,P$ be the phase space coordinates of the phase space of a single particle. Similarly, let $H_1$ be the Wigner function for $\hat{H}_1$. Then, using (2.2.123) we have:

$$\mathrm{N} = \int n(Q,P)d\mathbb{P}_1, \qquad \mathrm{E} = \int H_1(Q,P)n(Q,P)d\mathbb{P}_1 \tag{2.6.101}$$

in exact analogy with Orig. (5.10.56). Very often in applications (but not always), the single particle states can be considered as basically classical because we may have inadequate resolution to notice single particle quantum effects. However, even in such a case, quantum mechanics affects the statistics of the many particle system. In this pseudo-classical case, we approximate $H_1$ simply by the classical Hamiltonian $\mathcal{H}_1$ of a single particle of the field theory in the classical limit. As well, then $E_k$ limits toward being $\mathcal{H}$ evaluated at the relevant point in the single particle phase space. Thus, in this pseudo-classical limit we approximate:

$$\boxed{n_{\mathrm{BE}}(Q,P) \approx \frac{1}{e^{\beta(\mathcal{H}_1-\mu)}-1}} \tag{2.6.102}$$

where the subscript BE indicates that this is appropriate for Bose-Einstein particles.

We can perform all of the above analysis for fermionic fields as well, and surprisingly little changes about the calculations. Before we do so, we need to introduce a few basic facts about how fermions work. For the moment, suppose we have a single fermionic quantum harmonic oscillator. That is, we have an operator $\hat{a}$ that we call an annihilation operator and its adjoint $\hat{a}^\dagger$ which together satisfy:

$$[\hat{a},\hat{a}]_+ = 0, \qquad [\hat{a},\hat{a}^\dagger]_+ = 1 \tag{2.6.103}$$

and so that the Hamiltonian operator for the system is:

$$\hat{H} = \hbar\omega\hat{a}^\dagger\hat{a}. \tag{2.6.104}$$

We want to now understand the Hilbert space of this system and to find the eigenstates and eigenvalues of $\hat{H}$. Because of the anticommutation relations, we immediately have that $\hat{a}$ and $\hat{a}^\dagger$ as operators square to the zero operator:

$$\hat{a}^2 = \left(\hat{a}^\dagger\right)^2 = 0. \tag{2.6.105}$$

Let $|\psi\rangle$ be a normalized eigenstate of $\hat{H}$ of eigenvalue $E$:

$$\langle\psi|\psi\rangle = 1, \qquad \hat{H}|\psi\rangle = E|\psi\rangle. \tag{2.6.106}$$

Now let us act $\hat{a}^\dagger$ on the eigenvalue equation:

$$\hbar\omega\hat{a}^\dagger\hat{a}^\dagger\hat{a}|\psi\rangle = E\hat{a}^\dagger|\psi\rangle. \tag{2.6.107}$$

The left-hand side is automatically zero because $\hat{a}^\dagger\hat{a}^\dagger = 0$. Therefore:

$$\hat{H}|\psi\rangle = E|\psi\rangle \implies E\hat{a}^\dagger|\psi\rangle = 0. \tag{2.6.108}$$

So, each eigenstate of $\hat{H}$ either has eigenvalue 0 or satisfies $\hat{a}^\dagger|\psi\rangle = 0$. Let us first consider an eigenstate of eigenvalue 0 and let us call it $|\psi_0\rangle$. Then, acting $\hat{a}$ on the eigenvalue equation:

$$\hat{H}|\psi_0\rangle = 0 \implies \hat{a}\hat{a}^\dagger\hat{a}|\psi_0\rangle = 0 \implies (\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a})\hat{a}|\psi_0\rangle - \hat{a}^\dagger\hat{a}\hat{a}|\psi_0\rangle = 0 \implies \hat{a}|\psi_0\rangle = 0. \tag{2.6.109}$$

So, if an eigenstate of $\hat{H}$ has eigenvalue 0 then it is annihilated by the annihilation operator. Now let is see how $|\psi_0\rangle$ is affected by the creation operator. Let $|\psi_1\rangle$ be given by:

$$|\psi_1\rangle = \hat{a}^\dagger|\psi_0\rangle \implies \langle\psi_1|\psi_1\rangle = \langle\psi_0|\hat{a}\hat{a}^\dagger|\psi_0\rangle = \langle\psi|(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a})|\psi_0\rangle - \langle\psi_0|\hat{a}^\dagger\hat{a}|\psi_0\rangle = 1 \tag{2.6.110}$$

and so $|\psi_1\rangle$ is properly normalized. As well:

$$\hat{H}|\psi_1\rangle = \hbar\omega\hat{a}^\dagger\hat{a}\hat{a}^\dagger|\psi_0\rangle = \hbar\omega(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger)|\psi_1\rangle - \hbar\omega\hat{a}\hat{a}^\dagger\hat{a}^\dagger|\psi_0\rangle = \hbar\omega|\psi_1\rangle \tag{2.6.111}$$

and so $|\psi_1\rangle$ is also an eigenstate of $\hat{H}$, but with eigenvalue $\hbar\omega$. Because $\omega \neq 0$ and all of $\hat{H}$'s eigenstates satisfy $E\hat{a}^\dagger|\psi\rangle = 0$, we immediately have that $|\psi_1\rangle$ cannot be raised any further:

$$\hat{a}^\dagger|\psi_1\rangle = 0. \tag{2.6.112}$$

We can also recover $|\psi_0\rangle$ from $|\psi_1$ by lowering in the usual way:

$$\hat{a}|\psi_1\rangle = \hat{a}\hat{a}^\dagger|\psi_0\rangle = (\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a})|\psi_0\rangle - \hat{a}^\dagger\hat{a}|\psi_0\rangle = |\psi_0\rangle. \tag{2.6.113}$$

So, given an eigenstate of $\hat{H}$ with eigenvalue 0 we can produce a state of eigenvalue $\hbar\omega$ and neither can be raised or lowered to achieve any new independent states beyond these two.

Next let us consider an eigenstate of $\hat{H}$ which does not have eigenvalue 0. Such a state must have $a^\dagger|\psi\rangle = 0$. Let us now act $\hat{a}$ on this equation:

$$\hat{a}\hat{a}^\dagger|\psi\rangle = 0 \implies (\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a})|\psi\rangle - \hat{a}^\dagger\hat{a}|\psi\rangle = 0 \implies \hat{H}|\psi\rangle = \hbar\omega|\psi\rangle. \tag{2.6.114}$$

Thus, if an eigenstate of $\hat{H}$ does not have eigenvalue 0 then it must have eigenvalue $\hbar\omega$. By our previous calculation, we know from this state we can produce a state of eigenvalue 0 by simply considering $\hat{a}|\psi\rangle$. So, the eigenstates of $\hat{H}$ comes precisely in disjoint pairs with one member of each pair having eigenvalue 0 and the other having eigenvalue $\hbar\omega$ and the two elements of the pair related by raising/lowering. We can refine our definition of $\hat{H}$ slightly to guarantee that it posses precisely one such pair. If $\hat{H}$ had multiple pairs of eigenvalues, we could define a creation operator for each pair by projecting the full creation operator onto that pair's subspace of the Hilbert space. Then, the full Hamiltonian operator is given by summing over Hamiltonian operators each of which now manifestly acting only on its two dimensional Hilbert space orthogonal to all the others. Therefore, as part of the definition of a fermionic quantum harmonic oscillator we include the condition that its Hilbert space be purely two dimensional. This is a very stark difference to the bosonic oscillator which had an infinite dimensional Hilbert space! We now call the two eigenstates of $\hat{H}$ $|0\rangle$ and $|1\rangle$ and the number simply refers to the mode occupation number, or in other words, the number of particles in this mode. Surprisingly, **a fermionic field can never have more than one particle in any specific state**. That fact is called the **Pauli exclusion principle**. This enormously affects the statistical behavior of fermions. We can summarize our conclusions on fermions by saying that the Hamiltonian is simply diagonalized by states $|0\rangle$ and $|1\rangle$ satisfying:

$$\langle 0|0\rangle = \langle 1|1\rangle = 1, \quad \langle 0|1\rangle = 0, \quad |1\rangle = \hat{a}^\dagger|0\rangle, \quad |0\rangle = \hat{a}|1\rangle, \quad \hat{H}|0\rangle = 0, \quad \hat{H}|1\rangle = \hbar\omega|1\rangle. \tag{2.6.115}$$

Because the Hilbert space of a single mode of a fermion is just two dimensional, it is way simpler than that of a boson (which as an ordinary qho has infinite dimensional Hilbert space). We can give the fermionic $\hat{a}$, $\hat{a}^\dagger$, and $\hat{H}$ very simple matrix representations following the above. In particular:

$$|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \qquad |1\rangle = \begin{pmatrix}0\\1\end{pmatrix} \implies \hat{a} = \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}, \qquad \hat{a}^\dagger = \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}, \qquad \hat{H} = \hbar\omega\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}. \tag{2.6.116}$$

In (1.10.56) we found that the two photon wavefunction is symmetric under exchanging the parameters of the two photons, which comes from the fact that the creation operators for each of those particles commute with each other and they are of the same species. This generalizes in the obvious way so that the $N$ particle wavefunction of any single species bosonic field is symmetric under exchanging any pair of particles. For fermions, because the creation operators anticommute with each other, the $N$ particle wavefunctions are instead antisymmetric under particle exchange. For example, if $\hat{a}_h(\vec{k})$ is the annihilation operator for a half-integer spin field with $h$ some index covering the internal field degrees of freedom and $\vec{k}$ the wavenumber then precisely like (1.10.54) we can represent a generic two particle state as:

$$|\Psi\rangle_2 = \frac{1}{\sqrt{2}}\int_{\mathbb{K}'}\sum_{hh'}\psi_{hh'}(\vec{k},\vec{k}')\hat{a}_h^\dagger(\vec{k})\hat{a}_{h'}^\dagger(\vec{k}')|0\rangle d^3\vec{k}d^3\vec{k}' \tag{2.6.117}$$

where $|0\rangle$ is the vacuum state (in there there are 0 particles of any mode). We can rewrite the product $\hat{a}_h^\dagger(\vec{k})\hat{a}_{h'}^\dagger(\vec{k}')$ in terms of their commutator and anticommutator:

$$\hat{a}_h^\dagger(\vec{k})\hat{a}_{h'}^\dagger(\vec{k}') = \frac{1}{2}[\hat{a}_h^\dagger(\vec{k}),\hat{a}_{h'}^\dagger(\vec{k}')] + \frac{1}{2}[\hat{a}_h^\dagger(\vec{k}),\hat{a}_{h'}^\dagger(\vec{k}')]_+ = \frac{1}{2}[\hat{a}_h^\dagger(\vec{k}),\hat{a}_{h'}^\dagger(\vec{k}')] \tag{2.6.118}$$

because, due to the spin-statistics theorem, the different mode creation and annihilation operators anticommute. Returning this to $|\Psi\rangle_2$ and relabeling indices:

$$|\Psi\rangle_2 = \frac{1}{\sqrt{2}}\int_{\mathbb{K}'}\sum_{hh'}\frac{\psi_{hh'}(\vec{k},\vec{k}') - \psi_{h'h}(\vec{k}',\vec{k})}{2}\frac{1}{2}[\hat{a}_h^\dagger(\vec{k}),\hat{a}_{h'}^\dagger(\vec{k}')]|0\rangle d^3\vec{k}d^3\vec{k}' \tag{2.6.119}$$

$$= \frac{1}{\sqrt{2}}\int_{\mathbb{K}'}\sum_{hh'}\frac{\psi_{hh'}(\vec{k},\vec{k}') - \psi_{h'h}(\vec{k}',\vec{k})}{2}\hat{a}_h^\dagger(\vec{k})\hat{a}_{h'}^\dagger(\vec{k}')|0\rangle d^3\vec{k}d^3\vec{k}'. \tag{2.6.120}$$

This is just (2.6.117) with the wavefunction antisymmetrized, and so because the wavefunction is only defined through its representation of the two particle state in that equation, we have:

$$\psi_{hh'}(\vec{k},\vec{k}') = -\psi_{h'h}(\vec{k}',\vec{k}). \tag{2.6.121}$$

In the $N$ particle case, the fermionic wavefunction is antisymmetric under the swap of any pair of particles.

Now let us determine the occupation number distribution for fermionic fields. All of the analysis of (2.6.89) through (2.6.96) and (2.6.99) through (2.6.101) is completely unchanged from the bosonic case. The only thing we need to treat differently is the calculation of $\mathcal{Z}_k$ and the resulting expression for $n_k$. Now the eigenvalues of $\hat{a}_k^\dagger \hat{a}_k$ are simply 0 and 1 and so:

$$\mathcal{Z}_k = \sum_{n=0}^{1} e^{-\beta(E_k-\mu)n} = 1 + e^{-\beta(E_k-\mu)} \implies n_k = \frac{1}{e^{\beta(E_k-\mu)}+1}. \tag{2.6.122}$$

What a small change! However, that $-1$ versus $+1$ is qualitatively hugely important. If the exponential is small in the bosonic case, the denominator approaches 0 and the occupation number becomes arbitrarily large. Because $n_k$ is necessarily nonnegative from its definition, $E_k > \mu$ always in the bosonic case. If the exponential is small in the fermionic case, the denominator approaches 1 and the occupation number never exceeds 1, consistent with Pauli exclusion. In the pseudo-classical limit, we naturally have the Fermi-Dirac (FD) occupation number distribution:

$$\boxed{n_{\text{FD}}(Q,P) \approx \frac{1}{e^{\beta(\mathcal{H}_1-\mu)}+1}.} \tag{2.6.123}$$

At very low temperatures $\beta$ is large and so the eigenvalues of the operator $e^{-\beta\hat{H}}$ fall of very quickly. Because of this, the density matrix at low temperatures is very small except near the low energy states of the system. In the bosonic case, this causes all of the particles to cram together in low energy states (known as a Bose-Einstein condensate) since arbitrarily many particles can occupy a single state. This is reflected in (2.6.102) going to 0 if $\mathcal{H}_1 > \mu$ and $\beta$ is very large but in the denominator going to 0 (and so the result exploding) when $\mathcal{H}_1 \to \mu$. In the fermionic case, low temperature causes all of the lowest energy states to have essentially filled occupation number 1 and for there to be a cutoff energy (called the Fermi energy) above which the occupation number drops to 0. This is reflected in (2.6.123) going to 1 if $\mathcal{H}_1 < \mu$ or to 0 if $\mathcal{H}_1 > \mu$ whenever $\beta$ is very large. Thus at very low temperatures the different statistics of the particles matters a lot. Away from the extremely low temperature limit, in regions of the phase space where single particles have high energy compared to the chemical potential, $e^{\beta(\mathcal{H}_1-\mu)}$ becomes much greater than 1 and the exponential dominates the denominator in the both the Bose-Einstein and Fermi-Dirac cases. The leading order result simply reduces then to the classical result, the Maxwell-Boltzmann distribution in Orig. (5.10.62), for both cases.

## 2.7 Summary and Additional Reading

Important homework questions from this chapter:

- **Problem 2.6: Affinely Realized Symmetries** showed that for a classical system whose symmetry generators other than the Hamiltonian operator are polynomials of at most quadratic degree in the phase space variables there is a natural notion of quantization by taking those symmetry generators to be the Wigner functions for the corresponding quantum theory. This will be heavily used in quantum field theory.
- **Problem 2.10: The most important homework question** should help you get a grip on reality after spending time thinking about philosophical heady nonsense.

In this chapter we developed the general theory of quantum mechanics, which is the most fundamental formalism we present know for doing physics. In section 2.1 we showed how quantum mechanics arises naturally as the faithful representation of the algebra of operators we introduced all the way back in Orig. section 1.3. We found that under very simple assumptions on how operators can map to number in a way compatible with experiment, the physical state of a quantum system is its density matrix, with evolution given by the von Neumann equation, for which the wavefunction and Schrödinger equation are special cases. In section 2.2 we understood in general how to take the classical limit of a quantum system, and we found the Wigner-Weyl representation of quantum mechanics in phase space with the corresponding

Moyal bracket, giving a way to do quantum mechanics without operators and to view it as a deformation of classical mechanics. We also showed there is no coordinate covariant way to map a classical theory to a quantum one in general, thus canonical quantization does not really exist. But, in the case that you are only interested in affinely realized symmetries, there is a canonical quantization map compatible with specifically those symmetries, wherein the classical symmetry generators are identified with the Wigner functions of the quantum theory. In section 2.3 we derived the Lindblad equation, which is the most general possible closed linear equation of motion that can be satisfied by an open quantum system. In general an open quantum system does not satisfy a closed equation of motion, and even if it does, it may not be linear. In the case of weakly coupling to a large bath, it is possible to maintain closure and linearity, so the Lindblad equation is useful to studying the average statistical effects of dissipation and thermodynamics in quantum mechanics. In section 2.4 we introduced the Thermal/Unitary interpretation of quantum mechanics, in which the Born rule and the collapse rule are macroscopic dynamically emergent effects of underlying deterministic unitary dynamics of coupling a system to a thermodynamically large detector and the inherent need to treat the state of such a detector with statistical mechanics. The Born rule emerges as the only possible rule for describing the outcomes of a macroscopic unbiased binning procedure of measurement. The collapse rule can be derived circumstantially depending on the kind of measurement, with the lack of a well understand more general derivation essentially being the extent to which the measurement problem remains unsolved. In section 2.5 we described several other interpretations of quantum mechanics and discussed their pros and cons. We also proved Bell's theorem and tried to clarify the actual philosophical implications of it. In section 2.6 we outlined the basics of quantum statistical mechanics, which quite miraculously carry almost everything from classical statistical mechanics with almost nothing changed. We also introduced the concepts of bosonic and fermionic particles and hinted at the spin-statistics theorem, which will be a fundamental result relativistic quantum field theory. We now have all of the background necessary for us to dive head-first into proper quantum field theory in the next chapter.

This chapter is not a complete treatment of essential topics in quantum mechanics. Refer to [2, 33, 48, 51, 57, 63] for general treatments. For more on the theory of measurement and on the interpretative elements of quantum mechanics, see [1, 7, 9, 10, 16, 17, 25, 36, 38, 39, 54].

# Chapter 3

# Quantum Field Theory

*I insist upon the view that all is waves.*
– Erwin Schrödinger [55]

*The career of a young theoretical physicist consists of treating the harmonic oscillator in ever-increasing levels of abstraction.*
– Sidney Coleman [12]

*If my calculations are correct, when this baby hits 88 miles per hour... you're gonna see some serious shit.*
– Emmett Brown [64]

Quantum field theory is a notorious subject. The notation is often obtuse, the derivations are often of very low quality, mathematical rigor or even a coherent use of definitions is typically nowhere to be found. Conceptually, it is for the most part not actually so difficult. Rather, calculations in quantum field theory are typically quite tedious and some of the essential foundations of the theory remain poorly understood. In some ways quantum field theory is the opposite of general relativity. General relativity is perfectly rigorously mathematically well defined and most of the time calculations are just too hard. We have enormously developed tools for systematic calculations in quantum field theory, but we don't know how to define it nearly as rigorously as our other theories in physics. There are too many limits to take, and they don't commute with each other. We will try to do our best deriving things as always, but sadly many of our derivations going forward will not be at the same standards as our previous ones. All quantum field theory really is is ordinary quantum mechanics on an ordinary Hilbert space, where that specific Hilbert space is that formed from the multiparticle states from the creation and annihilation operators that are the mode operators of some fields of interest. The issues all come from the fact that fields have infinitely many modes and any sane calculation must start with a finite number of modes and take limits. It is difficult to do so in a way that properly maintains the desired symmetries of the field theory and gives finite results. While it is a bit uphill at the beginning, once you crest the peak, a wonderful view comes into focus. Let's get started and see what happens.

Having now found through the resolution of the ultraviolet catastrophe and the various statistical properties of multiparticle quantum mechanics that quantum field theory is necessary, we can state the basic assumptions of our new worldview. We have:

**Assumption :: LQFTI - Quantum Field State of Nature**

The state of the Universe is described by a density matrix $\hat{\rho}$ which evolves according to the von Neumann equation on a Hilbert space $\mathbb{H}$ of field functionals.

We are not ready to attempt dealing with general relativity, so we will look at simply special relativistic

quantum field theory at this time. Thus, we assume:

**Assumption :: LQFTII - Lorentz Boost Symmetry**

The Universe has symmetry under applying a Lorentz boost.

and:

**Assumption :: LQFTIII - Spacetime is static and Minkowskian**

Spacetime is not itself dynamical and is instead fixed without any curvature or punctures, extending infinitely in all directions.

When doing quantum field theory no sane person nor anyone who values their time uses units except those in which $\hbar = c = 1$, so it is time we finally switch to high energy units as promised in section 1.5. These are units in which the energy of a normalized wavefunction of definite angular frequency is exactly that angular frequency and in which the distance traveled by light in a given time duration is precisely that duration.

## 3.1 Representations of the Lorentz Group

Viewed in the passive picture, we will have some field operators $\hat{\phi}^a(\vec{x})$ which act on our Hilbert space $\mathbb{H}$ of states and the physical state of our qft will be some time dependent density matrix $\hat{\rho}(t)$. The field operators are functions of spatial position for the same reason that classical fields are: a field theory is really just an ordinary theory in which spatial position is a "continuous index" for designating which variable is being described. Because we are working with a relativistic theory, this separated treatment of time $t$ as the evolution parameter and position $\vec{x}$ as a field label is inconvenient, as Lorentz boosts will mix the two. So, typically in relativistic quantum field theory we will use either the active picture or the interaction picture, so that $(t, \vec{x})$ appears more covariantly. We will often forgo explicitly putting a subscript on the field operators to indicate which picture they are taken in just to keep equations from being extremely cluttered. Due to this, be careful to read the text around equations which will always make it clear what picture is used. Following (2.1.60) and (2.1.62), we now go to the active picture so that the field operators are $\hat{\phi}^a(t, \vec{x})$ and evolve according to (2.1.47) and the density matrix $\hat{\rho}$ is constant and simply defines the summary map.

What can our field operators be? In general it is not such an easy question, but with some reasonable restrictions we can enumerate the possibilities. Whatever operators we have, $\langle \hat{\phi}^a(t, \vec{x}) \rangle$ will be just be classical fields (for exactly the same reasons that in a generic quantum system $\langle \hat{Q}^a \rangle$ and $\langle \hat{P}_a \rangle$ are classical variables. As always, we assume that there exists a stable equilibrium state in the phase space of the universe so that if we start at that state we stay there and if we start nearby we stay nearby. We choose to define our fields so that $\langle \hat{\phi}^a(t, \vec{x}) \rangle = 0$ corresponds to such an equilibrium (shifting your canonical variables by a constant is always an allowed canonical transformation). Then, by the logic of the theorem of normal modes from Orig. section 1.15, nearby that equilibrium our Hamiltonian will be approximately some quadratic polynomial in our canonical variables. Because our Hamiltonian is of the form $\langle \hat{H} \rangle$, this requires that nearby equilibrium our Hamiltonian operator is a quadratic polynomial in our field operators.

Whatever fields $\phi^a(X)$ we have in our theory, they will transform in some way under a Lorentz transformation. As well, the coordinates of the evaluation point $X$ becomes transformed under a Lorentz transformation to new coordinates $X'$. Therefore, under Lorentz transformation, our transformed fields at the transformed coordinates $\phi'^a(X')$ will be related in some way to our untransformed fields at the untransformed coordinates:

$$\phi'^a(X') = R^a(\phi(X)) \tag{3.1.1}$$

for some functions $R^a$ dependent on the specific Lorentz transformation applied. The functions $R$ are not taken to depend on the coordinates of $X$ due to the spacetime translational invariance of our theory (and so the fields at each point should transform under the same transformation law). Under such a transformation,

we define $\mathscr{D}_{\mu\nu}\phi^a$ as:

$$\mathscr{D}_{\mu\nu}\phi^a \stackrel{\text{def}}{=} \left.\frac{dR^a(\phi(X))}{d\theta^{\mu\nu}}\right|_{\theta=0} \tag{3.1.2}$$

where the $\theta_{\mu\nu}$ are our transformation parameters. These $\mathscr{D}_{\mu\nu}\phi^a$ are consequently functions of all of the fields. Suppose we are considering a theory near some stable equilibrium (as we usually are) and define our fields so that $\phi^a = 0$ is the stable equilibrium (by the theorem of normal modes this is always possible if such an equilibrium exists). Then, we keep our Lagrangian density as a Taylor expansion about 0 up to some order $N$, giving a polynomial of degree $N$ in the fields (and their derivatives). If we want our Lagrangian density to have Lorentz symmetry then it must be the case that replacing $\phi^a(X)$ by $\phi'^a(X')$ results in a Lagrangian density which differs only by a 4-divergence. Thus, the transformed Lagrangian density must also be a polynomial of degree $N$ in the fields up to a 4-divergence. Because we are performing an expansion to order $N$ in powers of $\phi$, we only need to consider transformation functions $R^a(\phi)$ which are themselves polynomials of up to degree $N$. It is possible to define new fields $\phi_{\text{new}}$ as a polynomial in $\phi$ so that at least up to the necessary order $N$ the $\phi_{\text{new}}$ transform affinely under Lorentz transformation. (To do this, write $\phi_{\text{new}}$ as a general polynomial in $\phi$. $\phi_{\text{new}}$ then transforms under Lorentz transformation by applying $R$ to each factor of $\phi$. Do this explicitly by inserting the polynomial expression for $R$ then fix the coefficients of how $\phi_{\text{new}}$ is expressed in terms of $\phi$ so that the transformed $\phi_{\text{new}}$ is an affine function of the untransformed $\phi_{\text{new}}$. This gives a system of equations for the coefficients at each order which can be solved in principle. We do not need the explicit solution, just to know that it can be done.)

Assume now that we have defined our fields as described so that by definition they transform affinely under Lorentz transformations. Automatically then under Lorentz transformation:

$$\phi'^a(X') = R^a(\phi(X)) = U^a{}_b\phi^b(X) + \xi^a \tag{3.1.3}$$

for some $U^a{}_b$ and $\xi^a$ which depend on the Lorentz transformation. However, consider the stable equilibrium point $\phi^a = 0$. Under this transformation the equilibrium becomes $\phi'^a(X') = \xi^a$. Our Lagrangian density is a series expansion about $\phi^a = 0$ and so for our expansion to also be valid for our transformed fields the expansion point must be the same (otherwise, previously neglected terms in the series expansion could be made large), meaning we must have $\phi'^a(X') = 0$ whenever $\phi^a = 0$ for our polynomial expansion to be valid. Thus, $\xi^a = 0$. So, for a Taylor expanded Lagrangian up to degree $N$ in the fields, the form of the Lorentz transformation must be linear:

$$\phi'^a(X') = R^a(\phi(X)) = U^a{}_b\phi^b(X). \tag{3.1.4}$$

Differentiating we have:

$$\mathscr{D}_{\mu\nu}\phi^a = (\mathscr{D}_{\mu\nu}U^a{}_b)\phi^b. \tag{3.1.5}$$

For convenience, let us define:

$$(\mathcal{S}_{\mu\nu})^a{}_b \stackrel{\text{def}}{=} \mathscr{D}_{\mu\nu}U^a{}_b \implies \mathscr{D}_{\mu\nu}\phi^a = (\mathcal{S}_{\mu\nu})^a{}_b\phi^b. \tag{3.1.6}$$

Because the Lorentz transformations form a group, the composition of such transformations is itself such a transformation. In particular, consider a Lorentz transformation $L_1$ with parameters $\varepsilon_1\theta^{\mu\nu}$ and another $L_2$ with parameters $\varepsilon_2\theta^{\mu\nu}$. Then, the composition of such transformations gives:

$$L_1L_2 = e^{\varepsilon_1\theta^{\mu\nu}J_{\mu\nu}}e^{\varepsilon_2\theta^{\mu\nu}J_{\mu\nu}} = e^{(\varepsilon_1+\varepsilon_2)\theta^{\mu\nu}J_{\mu\nu}} \tag{3.1.7}$$

which is then a Lorentz transformation with parameters $(\varepsilon_1+\varepsilon_2)\theta^{\mu\nu}J_{\mu\nu}$. This manipulation of the exponentials is valid because the matrix $\theta^{\mu\nu}J_{\mu\nu}$ of course commutes with itself. If we now consider a Lorentz transformation by $\varepsilon\theta^{\mu\nu}$ letting $\phi_0$ be our untransformed fields and $\phi_\varepsilon$ be our transformed fields, we have:

$$\phi^a_\varepsilon = U^a_{\varepsilon\ b}\phi^b_0 \tag{3.1.8}$$

Applying our composition rule to the fields, we see that:

$$\phi^a_{\varepsilon_1+\varepsilon_2} = U^a_{(\varepsilon_1+\varepsilon_2)_b}\phi^b_0 = U^a_{\varepsilon_1\,c}U^c_{\varepsilon_2\,b}\phi^b_0 = U^a_{\varepsilon_1\,b}\phi^b_{\varepsilon_2}. \tag{3.1.9}$$

If we now differentiate with respect to $\varepsilon_1$ at $\varepsilon_1 = 0$ and use the chain rule:

$$\left.\frac{d\phi^a_{\varepsilon_1+\varepsilon_2}}{d\varepsilon_1}\right|_{\varepsilon_1=0} = \left.\frac{dU^a_{\varepsilon_1\,b}}{d\varepsilon_1}\right|_{\varepsilon_1=0}\phi^b_{\varepsilon_2} \tag{3.1.10}$$

$$\implies \left.\frac{d\phi^a_{\varepsilon_1+\varepsilon_2}}{d\varepsilon_2}\right|_{\varepsilon_1=0} = \left.\frac{dU^a{}_b}{d\theta^{\mu\nu}}\right|_{\theta=0}\left.\frac{d(\varepsilon_1\theta^{\mu\nu})}{d\varepsilon_1}\right|_{\varepsilon_1=0}\phi^b_{\varepsilon_2} \tag{3.1.11}$$

$$\implies \frac{d\phi^a_{\varepsilon_2}}{d\varepsilon_2} = (\mathscr{D}_{\mu\nu}U^a{}_b)\theta^{\mu\nu}\phi^b_{\varepsilon_2} = \theta^{\mu\nu}(S_{\mu\nu})^a{}_b\phi^b_{\varepsilon_2}. \tag{3.1.12}$$

As a matrix equation, this is a familiar linear differential equation with solution given by the matrix exponential:

$$\phi^a_\varepsilon = \left(e^{\varepsilon\theta^{\mu\nu}S_{\mu\nu}}\right)^a{}_b\phi^b_0. \tag{3.1.13}$$

So, we can see by matching (3.1.8) and (3.1.13) and setting $\varepsilon = 1$ that:

$$U^a{}_b = \left(e^{\theta^{\mu\nu}S_{\mu\nu}}\right)^a{}_b. \tag{3.1.14}$$

Now consider two distinct Lorentz transformations which on coordinates have matrices with components $L^\mu_{1\nu}$ and $L^\mu_{2\nu}$. Then, by the group property there exists a Lorentz transformation with matrix components $L^\mu_{3\nu}$ so that:

$$L_1L_2 = L_3. \tag{3.1.15}$$

Suppose that the Lorentz transformation parameters of $L_1$ are $\varepsilon\theta_1^{\mu\nu}$ and that the Lorentz transformation parameters of $L_2$ are $\varepsilon\theta_2^{\mu\nu}$. Then:

$$L_1L_2 = \left(\mathbb{1} + \varepsilon\theta_1^{\mu\nu}J_{\mu\nu} + \frac{\varepsilon^2}{2}\left(\theta_1^{\mu\nu}J_{\mu\nu}\right)^2 + \mathcal{O}(\varepsilon^3)\right)\left(\mathbb{1} + \varepsilon\theta_2^{\rho\sigma}J_{\rho\sigma} + \frac{\varepsilon^2}{2}\left(\theta_2^{\rho\sigma}J_{\rho\sigma}\right)^2 + \mathcal{O}(\varepsilon^3)\right) \tag{3.1.16}$$

$$= \mathbb{1} + \varepsilon(\theta_1^{\mu\nu} + \theta_2^{\mu\nu})J_{\mu\nu} + \frac{\varepsilon^2}{2}\left(\theta_1^{\mu\nu}J_{\mu\nu}\right)^2 + \varepsilon^2\theta_1^{\mu\nu}\theta_2^{\rho\sigma}J_{\mu\nu}J_{\rho\sigma} + \frac{\varepsilon^2}{2}\left(\theta_2^{\rho\sigma}J_{\rho\sigma}\right)^2 + \mathcal{O}(\varepsilon^3). \tag{3.1.17}$$

As well, $L_3$ must have some Lorentz transformation parameters $\theta_3^{\mu\nu}(\varepsilon)$ which will depend on $\varepsilon$ as some series:

$$\theta_3^{\mu\nu}(\varepsilon) = \theta_{30}^{\mu\nu} + \varepsilon\theta_{31}^{\mu\nu} + \varepsilon^2\theta_{32}^{\mu\nu} + \mathcal{O}(\varepsilon^3). \tag{3.1.18}$$

When $\varepsilon = 0$, $L_1$ and $L_2$ are the identity and so $L_3$ is as well. Then:

$$L_3|_{\varepsilon=0} = e^{\theta_{30}^{\mu\nu}J_{\mu\nu}} = \mathbb{1} \tag{3.1.19}$$

and so we take $\theta_{30}^{\mu\nu} = 0$ by definition. Then, expanding the exponential form for $L_3$:

$$L_3 = e^{\theta_3^{\mu\nu}J_{\mu\nu}} = e^{\varepsilon\theta_{31}^{\mu\nu}J_{\mu\nu} + \varepsilon^2\theta_{32}^{\mu\nu}J_{\mu\nu} + \mathcal{O}(\varepsilon^3)} \tag{3.1.20}$$

$$= \mathbb{1} + \varepsilon\theta_{31}^{\mu\nu}J_{\mu\nu} + \frac{1}{2}\varepsilon^2(\theta_{31}^{\mu\nu}J_{\mu\nu})^2 + \varepsilon^2\theta_{32}^{\mu\nu}J_{\mu\nu} + \mathcal{O}(\varepsilon^3). \tag{3.1.21}$$

Matching this against (3.1.17) at order $\varepsilon$ we see:

$$\theta_1^{\mu\nu} + \theta_2^{\mu\nu} = \theta_{31}^{\mu\nu}. \tag{3.1.22}$$

At order $\varepsilon^2$ then we see:

$$\frac{1}{2}\left(\theta_1^{\mu\nu}J_{\mu\nu}\right)^2 + \theta_1^{\mu\nu}\theta_2^{\rho\sigma}J_{\mu\nu}J_{\rho\sigma} + \frac{1}{2}(\theta_2^{\rho\sigma}J_{\rho\sigma})^2 = \frac{1}{2}(\theta_1^{\mu\nu}J_{\mu\nu} + \theta_2^{\mu\nu}J_{\mu\nu})^2 + \theta_{32}^{\mu\nu}J_{\mu\nu} \tag{3.1.23}$$

$$\implies \frac{1}{2}\theta_1^{\mu\nu}\theta_2^{\rho\sigma}(J_{\mu\nu}J_{\rho\sigma} - J_{\rho\sigma}J_{\mu\nu}) = \theta_{32}^{\mu\nu}J_{\mu\nu}. \tag{3.1.24}$$

Using then equation the Lorentz algebra:

$$\frac{1}{2}\theta_1^{\mu\nu}\theta_2^{\rho\sigma}\left(\eta_{\mu\sigma}J_{\rho\nu} - \eta_{\mu\rho}J_{\sigma\nu} - \eta_{\nu\sigma}J_{\rho\mu} + \eta_{\nu\rho}J_{\sigma\mu}\right) = \theta_{32}^{\mu\nu}J_{\mu\nu} \tag{3.1.25}$$

$$\implies (\theta_{1\rho}^\nu\theta_2^{\rho\mu} - \theta_{1\rho}^\mu\theta_2^{\rho\nu})J_{\mu\nu} = \theta_{32}^{\mu\nu}J_{\mu\nu} \tag{3.1.26}$$

$$\implies \theta_{32}^{\mu\nu} = \theta_{1\rho}^\nu\theta_2^{\rho\mu} - \theta_{1\rho}^\mu\theta_2^{\rho\nu}. \tag{3.1.27}$$

Thus, we have the Lorentz transformation composition rule:

$$e^{\varepsilon\theta_1^{\mu\nu}J_{\mu\nu}}e^{\varepsilon\theta_2^{\mu\nu}J_{\mu\nu}} = e^{\varepsilon(\theta_1^{\mu\nu} + \theta_2^{\mu\nu})J_{\mu\nu} - 2\varepsilon^2\theta_{1\rho}^\mu\theta_2^{\rho\nu}J_{\mu\nu} + \mathcal{O}(\varepsilon^3)}. \tag{3.1.28}$$

Applying this composition rule to the fields then demands:

$$U_1U_2 = U_3 \tag{3.1.29}$$

Because $U$ is given as the matrix exponential of $\mathcal{S}$ in the same way that $L$ is given as the matrix exponential of $J$, we immediately have:

$$U_1U_2 = \mathbb{1} + \varepsilon(\theta_1^{\mu\nu} + \theta_2^{\mu\nu})\mathcal{S}_{\mu\nu} + \frac{\varepsilon^2}{2}\left(\theta_1^{\mu\nu}\mathcal{S}_{\mu\nu}\right)^2 + \varepsilon^2\theta_1^{\mu\nu}\theta_2^{\rho\sigma}\mathcal{S}_{\mu\nu}\mathcal{S}_{\rho\sigma} + \frac{\varepsilon^2}{2}\left(\theta_2^{\rho\sigma}\mathcal{S}_{\rho\sigma}\right)^2 + \mathcal{O}(\varepsilon^3). \tag{3.1.30}$$

and:

$$U_3 = \mathbb{1} + \varepsilon\theta_{31}^{\mu\nu}\mathcal{S}_{\mu\nu} + \frac{1}{2}\varepsilon^2(\theta_{31}^{\mu\nu}\mathcal{S}_{\mu\nu})^2 + \varepsilon^2\theta_{32}^{\mu\nu}\mathcal{S}_{\mu\nu} + \mathcal{O}(\varepsilon^3). \tag{3.1.31}$$

At order $\varepsilon$, these two match automatically by using that $\theta_{31}^{\mu\nu}$ is given in (3.1.22). At order $\varepsilon^2$ we get:

$$\frac{1}{2}\theta_1^{\mu\nu}\theta_2^{\rho\sigma}(\mathcal{S}_{\mu\nu}\mathcal{S}_{\rho\sigma} - \mathcal{S}_{\rho\sigma}\mathcal{S}_{\mu\nu}) = -2\theta_{1\rho}^{\mu}\theta_2^{\rho\nu}J_{\mu\nu} \tag{3.1.32}$$

$$\implies \theta_1^{\mu\nu}\theta_2^{\rho\sigma}\left([\mathcal{S}_{\mu\nu},\mathcal{S}_{\rho\sigma}] - \eta_{\nu\sigma}\hat{\mathcal{S}}_{\mu\rho} + \eta_{\mu\sigma}\hat{\mathcal{S}}_{\nu\rho} + \eta_{\nu\rho}\hat{\mathcal{S}}_{\mu\sigma} - \eta_{\mu\rho}\hat{\mathcal{S}}_{\nu\sigma}\right) = 0. \tag{3.1.33}$$

Because this holds for all $\theta_1^{\mu\nu}$ and $\theta_2^{\rho\sigma}$, we have:

$$[\mathcal{S}_{\mu\nu},\mathcal{S}_{\rho\sigma}] = \eta_{\nu\sigma}\hat{\mathcal{S}}_{\mu\rho} - \eta_{\mu\sigma}\hat{\mathcal{S}}_{\nu\rho} - \eta_{\nu\rho}\hat{\mathcal{S}}_{\mu\sigma} + \eta_{\mu\rho}\hat{\mathcal{S}}_{\nu\sigma}. \tag{3.1.34}$$

Thus, these $\mathcal{S}$ matrices must satisfy the exact same algebraic relations as the original Lorentz generators and thus are said to be a representation of the Lorentz algebra (the $U$ matrices then form a representation of the Lorentz group). Conventionally it is common in physics to instead use the matrices $\hat{S}_{\mu\nu}$ defined by:

$$\hat{S}_{\mu\nu} = i\hat{\mathcal{S}}_{\mu\nu} \tag{3.1.35}$$

$$\implies [\hat{S}_{\mu\nu},\hat{S}_{\rho\sigma}] = i\left(\eta_{\nu\sigma}\hat{S}_{\mu\rho} - \eta_{\mu\sigma}\hat{S}_{\nu\rho} - \eta_{\nu\rho}\hat{S}_{\mu\sigma} + \eta_{\mu\rho}\hat{S}_{\nu\sigma}\right). \tag{3.1.36}$$

These $\hat{S}_{\mu\nu}$ are called the spin matrices for our fields. Define as well:

$$\hat{J}_a = \frac{1}{2}\epsilon_a{}^{bc}\hat{S}_{bc}, \qquad \hat{K}_a = \hat{S}_{a0}. \tag{3.1.37}$$

The commutation relation for the $\hat{S}_{\mu\nu}$ then becomes:

$$[\hat{J}_a,\hat{J}_b] = i\epsilon_{ab}{}^c\hat{J}_c, \qquad [\hat{J}_a,\hat{K}_b] = i\epsilon_{ab}{}^c\hat{K}_c, \qquad [\hat{K}_a,\hat{K}_b] = -i\epsilon_{ab}{}^c\hat{J}_c. \tag{3.1.38}$$

These commutators can be decomposed nicely by defining the operators:

$$\hat{L}_a = \frac{\hat{J}_a + i\hat{K}_a}{2}, \qquad \hat{R}_a = \frac{\hat{J}_a - i\hat{K}_a}{2}. \tag{3.1.39}$$

Then, $\hat{L}$ and $\hat{R}$ satisfy a much simpler algebra:

$$[\hat{L}_a,\hat{L}_b] = [\frac{\hat{J}_a + i\hat{K}_a}{2}, \frac{\hat{J}_b + i\hat{K}_b}{2}] = \frac{1}{4}[\hat{J}_a,\hat{J}_b] + \frac{i}{4}[\hat{K}_a,\hat{J}_b] + \frac{i}{4}[\hat{J}_a,\hat{K}_b] - \frac{1}{4}[\hat{K}_a,\hat{K}_b] \tag{3.1.40}$$

$$= \frac{1}{4}i\epsilon_{ab}{}^c\hat{J}_c - \frac{1}{4}\epsilon_{ab}{}^c\hat{K}_c - \frac{1}{4}\epsilon_{ab}{}^c\hat{K}_c + \frac{1}{4}i\epsilon_{ab}{}^c\hat{J}_c = i\epsilon_{ab}{}^c\frac{\hat{J}_c + i\hat{K}_c}{2} = i\epsilon_{ab}{}^c\hat{L}_c \tag{3.1.41}$$

$$[\hat{L}_a,\hat{R}_b] = [\frac{\hat{J}_a + i\hat{K}_a}{2}, \frac{\hat{J}_b - i\hat{K}_b}{2}] = \frac{1}{4}[\hat{J}_a,\hat{J}_b] + \frac{i}{4}[\hat{K}_a,\hat{J}_b] - \frac{i}{4}[\hat{J}_a,\hat{K}_b] + \frac{1}{4}[\hat{K}_a,\hat{K}_b] \tag{3.1.42}$$

$$= \frac{1}{4}i\epsilon_{ab}{}^c\hat{J}_c - \frac{1}{4}\epsilon_{ab}{}^c\hat{K}_c + \frac{1}{4}\epsilon_{ab}{}^c\hat{K}_c - \frac{1}{4}i\epsilon_{ab}{}^c\hat{J}_c = 0. \tag{3.1.43}$$

$$[\hat{R}_a,\hat{R}_b] = [\frac{\hat{J}_a - i\hat{K}_a}{2}, \frac{\hat{J}_b - i\hat{K}_b}{2}] = \frac{1}{4}[\hat{J}_a,\hat{J}_b] - \frac{i}{4}[\hat{K}_a,\hat{J}_b] - \frac{i}{4}[\hat{J}_a,\hat{K}_b] - \frac{1}{4}[\hat{K}_a,\hat{K}_b] \tag{3.1.44}$$

$$= \frac{1}{4}i\epsilon_{ab}{}^c\hat{J}_c + \frac{1}{4}\epsilon_{ab}{}^c\hat{K}_c + \frac{1}{4}\epsilon_{ab}{}^c\hat{K}_c + \frac{1}{4}i\epsilon_{ab}{}^c\hat{J}_c = i\epsilon_{ab}{}^c\frac{\hat{J}_c - i\hat{K}_c}{2} = i\epsilon_{ab}{}^c\hat{R}_c. \tag{3.1.45}$$

That is, the $\hat{L}_a$ satisfy the SU(2) commutation relations among themselves, the $\hat{R}_a$ also satisfy the SU(2) commutation relations among themselves, and the $\hat{L}_a$ and $\hat{R}_a$ commute with each other and therefore are simultaneously diagonalizable. So, we can decompose our fields $\phi^a$ into different sets of fields which transform together under these $\hat{L}_a$ and $\hat{R}_a$ but where there is no mixing across sets. Each individual set is then called an irreducible representation of the Lorentz group. Suppose then we are studying such an irreducible set of fields. We found in Theorem 1.11 that the irreducible representations of the SU(2) algebra are specified by a half-integer number and then have given matrix forms. Because $\hat{L}_a$ and $\hat{R}_a$ all commute with each other, each one is allowed an independent value for that half integer. That is, an irreducible representation of the Lorentz group is specified by $(s_l, s_r)$ where $s_l$ and $s_r$ are half-integers corresponding to the $s$ in Theorem 1.11 for the $\hat{L}_a$ and $\hat{R}_a$. Representations of $\hat{L}_a$ are called "left-handed" and representations of $\hat{R}_a$ are called "right-handed" for reasons we will explore soon. The transformation of an irreducible representation of fields under $(s_l, s_r)$ is then given by $s_l \otimes s_r$ meaning the $s_l$ representation of the $\hat{L}_a$ and $s_r$ representation of the $\hat{R}_a$ act independently on their respective components of the fields. Any reducible representation of the fields then

can be written as a direct sum of irreducible representations of this form. Because the different sets in the sum don't mix into each other under Lorentz transformation, we tend to study each irreducible representation independently and then put them all together at the end. Once we have specified $(s_l, s_r)$, $\hat{L}_a$ and $\hat{R}_a$ are determined by Theorem 1.11. Then, we can invert to find $\hat{J}_a$ and $\hat{K}_a$:

$$\hat{J}_a = \hat{L}_a + \hat{R}_a, \qquad \hat{K}_a = i\hat{R}_a - i\hat{L}_a. \tag{3.1.46}$$

In turn these give us $\hat{S}_{\mu\nu}$ from (3.1.37). Then, from (3.1.14):

$$\hat{U} = e^{-i\theta^{\mu\nu}\hat{S}_{\mu\nu}}. \tag{3.1.47}$$

Thus, we then know fully how the fields transform under Lorentz transformation.

## 3.2 Spin-Statistics Theorem

As we found before, any stable Hamiltonian near its linearly stable equilibrium can be written as a sum of harmonic oscillators.

discuss C, P, and T classically discuss C, P, and T quantum mechanically

## 3.3 Path Integral

### 3.3.1 Path Integral for Time Evolution

Feynman developed a convenient formula, known as the path integral, for computing the time evolution operator $\hat{U}(t, t_0)$. For ordinary quantum mechanics his formula is actually obnoxiously inconvenient. But, while ordinary operator methods are generally much more painful for calculations in quantum field theory compared to single particle quantum mechanics, the path integral isn't really any worse when we make the leap to field theory. So, it becomes convenient by comparison. Strictly speaking, every calculation in quantum mechanics (including in quantum field theory) can be done without ever using the path integral. However, it is often the most convenient route to take for very complicated calculations. Recall that the unitary time evolution operator $\hat{U}(t, t_0)$ defined by:

$$i\hbar\frac{d\hat{U}(t,t_0)}{dt} = \hat{H}(t)\hat{U}(t,t_0), \qquad \hat{U}(t,t) = \mathbb{1}, \tag{3.3.1}$$

can be formally expressed in terms of the Hamiltonian operator using the time ordering transformation as:

$$\hat{U}(t,t_0) = \mathscr{T} : e^{-\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t')dt'} \tag{3.3.2}$$

and satisfies the composition property:

$$\hat{U}(t,t_0) = \hat{U}(t,t_1)\hat{U}(t_1,t_0). \tag{3.3.3}$$

The path integral is used to explicitly compute the configuration space propagator $\hat{U}(t, Q, t_0, Q_0)$, which is the matrix element function of $\hat{U}$ in coordinate space:

$$U(t,Q,t_0,Q_0) = \langle Q|\hat{U}(t,t_0)|Q_0\rangle \tag{3.3.4}$$

Consider a list of $N+1$ times $t_0, t_1, ..., t_{N-1}, t_N = t$ that divide up the time interval $[t_0, t]$ with $t_n < t_{n+1}$. We will eventually seek to take $N$ large so that we can also take the duration between adjacent times to become arbitrarily small. Using the composition property, we have immediately:

$$\hat{U}(t,t_0) = \hat{U}(t,t_{N-1})\hat{U}(t_{N-1},t_{N-2})...\hat{U}(t_2,t_1)\hat{U}(t_1,t_0) = \mathscr{T} : \prod_{n=0}^{N-1}\hat{U}(t_{n+1},t_n). \tag{3.3.5}$$

To compute $U(t, Q, t_0, Q_0)$, we insert a complete set of $|Q\rangle$ states between each factor in this product (we use the labels $Q_0, Q_1, ..., Q_{N-1}, Q_N = Q$:

$$\begin{aligned} U(t,Q,t_0,Q_0) &= \int \langle Q|\hat{U}(t,t_{N-1})|Q_{N-1}\rangle\langle Q_{N-1}|\hat{U}(t_{N-1},t_{N-2})|Q_{N-2}\rangle... \\ &\qquad \cdot \langle Q_2|\hat{U}(t_2,t_1)|Q_1\rangle\langle Q_1|\hat{U}(t_1,t_0)|Q_0\rangle d^dQ_1...d^dQ_{N-1} \\ &= \int U(t,Q,t_{N-1},Q_{N-1})U(t_{N-1},Q_{N-1},t_{N-2},Q_{N-2})... \end{aligned} \tag{3.3.6}$$

$$\cdot U(t_2,Q_2,t_1,Q_1)U(t_1,Q_1,t_0,Q_0)d^dQ_1...d^dQ_{N-1} \tag{3.3.7}$$

$$=\int\left(\prod_{n=0}^{N-1}U(t_{n+1},Q_{n+1},t_n,Q_n)\right)d^dQ_1...d^dQ_{N-1}. \tag{3.3.8}$$

This is true for each $N$, and so in the $N\to\infty$ limit:

$$U(t,Q,t_0,Q_0)=\lim_{N\to\infty}\int\left(\prod_{n=0}^{N-1}U(t_{n+1},Q_{n+1},t_n,Q_n)\right)d^dQ_1...d^dQ_{N-1}. \tag{3.3.9}$$

We will now compute each of these factors which in the large $N$ limit causes the time interval $t_{n+1}-t_n$ to become short. Note that because each $Q_n$ is integrated over its all range, the displacement $Q_{n+1}-Q_n$ does not become small.

Consider $U(t+\delta t,Q+\delta Q,t,Q)$ for $\delta t$ small but $\delta Q$ arbitrary. Using (3.3.2):

$$U(t+\delta t,Q+\delta Q,t,Q)=\langle Q+\delta Q|\hat{U}(t+\delta t,t)|Q\rangle=\langle Q+\delta Q|\mathscr{T}:e^{-\frac{i}{\hbar}\int_t^{t+\delta t}\hat{H}(t')dt'}|Q\rangle \tag{3.3.10}$$

$$=\langle Q+\delta Q|\left(\mathbb{1}-\frac{i}{\hbar}\int_t^{t+\delta t}\hat{H}(t')dt'+\mathcal{O}(\delta t^2)\right)|Q\rangle \tag{3.3.11}$$

$$=\langle Q+\delta Q|\left(\mathbb{1}-\frac{i}{\hbar}\hat{H}(t)\delta t+\mathcal{O}(\delta t^2)\right)|Q\rangle \tag{3.3.12}$$

$$=\langle Q+\delta Q|Q\rangle-\frac{i}{\hbar}\delta t\langle Q+\delta Q|\hat{H}(t)|Q\rangle+\mathcal{O}(\delta t^2). \tag{3.3.13}$$

Now use (2.2.17) with $H(t,Q,P)$ the Wigner function associated to $\hat{H}$:

$$U(t+\delta t,Q+\delta Q,t,Q)=\delta(\delta Q)-\frac{i}{\hbar}\delta t\int e^{\frac{i}{\hbar}P_a\delta Q^a}H\left(t,Q+\frac{\delta Q}{2},P\right)\frac{d^dP}{(2\pi\hbar)^d}+\mathcal{O}(\delta t^2) \tag{3.3.14}$$

$$=\int\frac{e^{\frac{i}{\hbar}P_a\delta Q^a}}{(2\pi\hbar)^d}d^dP-\frac{i}{\hbar}\delta t\int e^{\frac{i}{\hbar}P_a\delta Q^a}H\left(t,Q+\frac{\delta Q}{2},P\right)\frac{d^dP}{(2\pi\hbar)^d}+\mathcal{O}(\delta t^2) \tag{3.3.15}$$

$$=\int e^{\frac{i}{\hbar}P_a\delta Q^a}\left(1-\frac{i}{\hbar}\delta tH\left(t,Q+\frac{\delta Q}{2},P\right)+\mathcal{O}(\delta t^2)\right)\frac{d^dP}{(2\pi\hbar)^d} \tag{3.3.16}$$

$$=\int e^{\frac{i}{\hbar}\left(P_a\delta Q^a-H(t,Q+\delta Q/2,P)\delta t+\mathcal{O}(\delta t^2)\right)}\frac{d^dP}{(2\pi\hbar)^d}. \tag{3.3.17}$$

Define the phase-action change $\delta S$ by:

$$\delta S=P_a\delta Q^a-H(t,Q+\delta Q/2,P)\delta t \tag{3.3.18}$$

$$\implies U(t+\delta t,Q+\delta Q,t,Q)=\int e^{\frac{i}{\hbar}\delta S+\mathcal{O}(\delta t^2)}\frac{d^dP}{(2\pi\hbar)^d}. \tag{3.3.19}$$

Returning this to (3.3.9) and noting that if we divide up $[t_0,t_f]$ evenly, $t_{n+1}-t_n=\frac{t_f-t_0}{N}$:

$$U(t_f,Q_f,t_0,Q_0)=\lim_{N\to\infty}\int\left(\prod_{n=0}^{N-1}\int e^{\frac{i}{\hbar}\delta S_n+\mathcal{O}(1/N^2)}\frac{d^dP_n}{(2\pi\hbar)^d}\right)d^dQ_1...d^dQ_{N-1} \tag{3.3.20}$$

$$=\lim_{N\to\infty}\int e^{\frac{i}{\hbar}\sum_{n=0}^{N-1}\delta S_n+\mathcal{O}(1/N)}\frac{d^dP_0}{(2\pi\hbar)^d}\frac{d^dQ_1d^dP_1}{(2\pi\hbar)^d}...\frac{d^dQ_{N-1}d^dP_{N-1}}{(2\pi\hbar)^d}. \tag{3.3.21}$$

Here we used that the correction terms $\mathcal{O}(\delta t^2)$ are $\mathcal{O}(1/N^2)$ when written in terms of $N$ and that a sum of $N$ terms which are $\mathcal{O}(1/N^2)$ is $\mathcal{O}(1/N)$. If we now assume that it is safe to "complete" this limit, then in the $N\to\infty$ limit these $1/N$ terms vanish:

$$\boxed{U(t_f,Q_f,t_0,Q_0)=\lim_{N\to\infty}\int e^{\frac{i}{\hbar}\sum_{n=0}^{N-1}\delta S_n}\frac{d^dP_0}{(2\pi\hbar)^d}\frac{d^dQ_1d^dP_1}{(2\pi\hbar)^d}...\frac{d^dQ_{N-1}d^dP_{N-1}}{(2\pi\hbar)^d}} \tag{3.3.22}$$

where each $\delta S_n$ is given by:

$$\delta S_n=P_{na}(Q^a_{n+1}-Q^a_n)-H\left(t_n,\frac{Q_n+Q_{n+1}}{2},P_n\right)\frac{t_f-t_0}{N} \tag{3.3.23}$$

and $Q_N=Q_f$.

Equation (3.3.22) is the relatively safe formula to use for the path integral, but it is almost never what you will see written. It is called the path integral because we could view $Q_0, Q_1, ..., Q_{N-1}, Q_f$ and $P_0, ..., P_{N-1}$ as forming a path through classical phase space associated with the times $t_0, t_1, ..., t_{N-1}, t_f$ beginning at $Q_0$ at time $t_0$ and ending at $Q_f$ at time $t_f$. We integrate over all the intermediate points. In the $N \to \infty$ limit, this becomes (at least formally) "an integration over all paths" in phase space connecting $(t_0, Q_0)$ to $(t_f, Q_f)$. This is somewhat like the classical variational problem which considers variations of all such paths to find the distinguished classical path (by the one that makes the action stationary). Now we are integrating over all such paths. The integral measure over all such paths is quite a mess to write out, and you will note that we perform one more $P$ integration than $Q$ integration. Call the path that we are integrating over $(Q(t), P(t))$. This path then has boundary conditions $Q(t_0) = Q_0$ and $Q(t_f) = Q_f$. For compactness of notation, it is common to make the replacement:

$$\frac{d^d P_0}{(2\pi\hbar)^d}\frac{d^d Q_1 d^d P_1}{(2\pi\hbar)^d}...\frac{d^d Q_{N-1} d^d P_{N-1}}{(2\pi\hbar)^d} \to \mathcal{D}Q\mathcal{D}P = \mathcal{D}\mathbb{P}. \tag{3.3.24}$$

This replacement is just lazy notation and the expression on the far left side should be used when being careful. The expression in the exponent in the $N \to \infty$ limit actually goes to the form of a Riemann sum:

$$\lim_{N\to\infty} \sum_{n=0}^{N-1} \delta S_n = \lim_{N\to\infty} \sum_{n=0}^{N-1} \left( P_{na} \frac{Q^a_{n+1} - Q^a_n}{t_{n+1} - t_n} - H\left(t_n, \frac{Q_n + Q_{n+1}}{2}, P_n\right)\right)(t_{n+1} - t_n) \tag{3.3.25}$$

$$= \int_{t_0}^{t_f} \Big( P_a(t)\dot{Q}^a(t) - H(t, Q(t), P(t))\Big)\, dt \tag{3.3.26}$$

This integral is precisely the action integral one would get for the path $(Q, P)$ using the Wigner function $H(t, Q, P)$ associated to the Hamiltonian operator $\hat{H}$ as if it were the Hamiltonian for the quantities $(Q, P)$. Thus, define the Wigner action functional:

$$S[Q, P] = \int_{t_0}^{t_f} \Big( P_a\dot{Q}^a - H(t, Q, P)\Big)\, dt. \tag{3.3.27}$$

In terms of this functional, the condensed expression for the path integral becomes:

$$\boxed{U(t_f, Q_f, t, Q) = \int_{Q(t_0)=Q_0}^{Q(t_f)=Q_f} e^{\frac{i}{\hbar}S[Q,P]}\mathcal{D}Q\mathcal{D}P.} \tag{3.3.28}$$

Equation (3.3.28) has a very nice interpretation and that interpretation significantly predates Feynman – it was identified by Huygens nearly 300 years prior – the path integral gives the propagation amplitude for an initial wave disturbance (localized at $Q_0$ at time $t_0$) to a final wave disturbance (localized at $Q_f$ at time $t_f$) by summing (integrating) over all paths that connect the initial and final points (paths taken in phase space) and weighted by the phase accumulated along that particularly trajectory. Feynman's identification was to appreciate that the phase is itself given by (up to a multiplier) a functional of a very similar form to the classical action.

### 3.3.2 Important Subtleties

There are a few common errors made in the use of the path integral. These errors are common because they are safe for some common special cases but they fail in general. You should view them as useful shortcuts that are valid under appropriate circumstances, and you should remember to actually check the circumstances before you use them. First, the $H(t, Q, P)$ function in the path integral is the Wigner function associated to the Hamiltonian operator $\hat{H}$, it is **NOT** the Hamiltonian function of the classical system gotten by taking the classical limit of the quantum theory. It is a common error to think of the path integral as a "quantization scheme" (which you know by now doesn't exist) where you take the Hamiltonian function of a classical system and put it into (3.3.27). This is **WRONG**. Because generators in classical mechanics transform quantities by the Poisson bracket while functions on phase space in quantum mechanics transform according to the Moyal bracket, if you perform "path integral quantization" in one coordinate system and your friend performs the same process on the same classical system expressed in different coordinates, the resulting quantum theories will not generally be equivalent. You used the wrong transformation properties! In one perspective, $\mathcal{H}$ (the classical Hamiltonian) is the $\hbar \to 0^+$ limit of $H$ and $\hbar$ corrections are mixed around by quantum coordinate transformations. Throwing them away in one coordinate system (unphysically) designates a

preferred coordinate system. This distinction is very important for symmetry considerations. In order for the quantum system to have a desired symmetry, the Wigner function $H$ must have appropriate symmetry under the Moyal bracket with the generator of that symmetry. If you look solely for a function symmetric under the Poisson bracket with the generator, your theory will not in general actually have the right symmetry (people may call this an anomaly, but it isn't very anomalous to get the wrong answer by doing wrong math). The reason this error is sometimes safe is that sometimes all of our symmetry transformations of interest act linearly on our coordinates and momenta. If $Q$ and $P$ transform linearly then the generator of that transformation is necessarily a quadratic polynomial in $Q$ and $P$ and as we showed when proving Groenewold's theorem, there is a unique faithful quantization scheme of polynomials of quadratic degree of less and it is equivalent to identifying Wigner functions with classical generators. In essence, if at least one of the two functions in a bracket is a quadratic polynomial in $Q$ and $P$, then the Moyal and Poisson brackets of those functions agree. So, if *all* of our symmetry generators of interest other than the Hamiltonian are classically (at most) quadratic polynomials of $Q$ and $P$ and we develop a classical Hamiltonian with all of those symmetries, then using that Hamiltonian as the Wigner $H$ of a quantum theory will give a quantum theory with the correct symmetries.

Those who commit the second error have almost always committed the first. So, suppose for now you have already mistakenly replaced $H$ with the classical Hamiltonian $\mathcal{H}$. In that case, the functional $S$ in (3.3.27) becomes the classical Hamiltonian action. The second oft repeated error with the path integral is to claim that the momentum integrals can generally be completed so that afterward the integrand of the $Q$ integrals is of the form $e^{\frac{i}{\hbar}S_{\mathcal{L}}}$ where $S_{\mathcal{L}}$ is now the Lagrangian action of the classical theory. Let me explain where this incorrect claim comes from. Suppose that $H$ happens to depend on $P$ only as a quadratic polynomial and that the coefficients of the term quadratic in $P$ do not depend on $Q$. Such a simple $H$ takes the form:

$$H(t,Q,P)=\frac{1}{2}W^{ab}(t)P_aP_b+\Gamma^a(t,Q)P_a+V(t,Q) \tag{3.3.29}$$

for some functions $W^{ab}$, $\Gamma^a$, $V$. For shorthand, define:

$$\bar{Q}_n=\frac{Q_n+Q_{n+1}}{2},\qquad \dot{Q}_n=\frac{Q_{n+1}-Q_n}{t_{n+1}-t_n},\qquad \delta t_n=t_{n+1}-t_n. \tag{3.3.30}$$

Now using our simple form of $H$ and returning to (3.3.22), we can explicitly complete the $P$ integrals using the results of the multivariable Gaussian integral. In particular, if $\vec{x}$ is a vector in $\mathbb{R}^n$ and $A_{ab}=A_{ba}$, $B_a$, and $C$ are some real coefficients independent of $x$ and the matrix $A_{ab}$ is nondegenerate, then:

$$\int_{\mathbb{R}^n}e^{-\frac{i}{2}A_{ab}x^ax^b+iB_ax^a+iC}d^n\vec{x}=\frac{(2\pi)^{\frac{n}{2}}}{i^{\frac{n}{2}}\sqrt{\det A}}e^{\frac{i}{2}A^{-ab}B_aB_b+iC} \tag{3.3.31}$$

where $A^{-ab}$ is the inverse matrix of $A_{ab}$ (this result is proven in the homework). With these, in this special case we can simplify (3.3.22) to:

$$U(t_f,Q_f,t_0,Q_0)=\lim_{N\to\infty}\int e^{\frac{i}{\hbar}\sum_{n=0}^{N-1}\delta S_{\mathcal{L}n}}\prod_{n=0}^{N-1}\left(\frac{1}{(2\pi i\hbar\delta t_n)^{\frac{d}{2}}\sqrt{\det W(t_n)}}\right)d^dQ_1...d^dQ_{N-1} \tag{3.3.32}$$

where we have defined:

$$\delta S_{\mathcal{L}n}=\frac{1}{2}W^{-1}_{ab}(t_n)(\dot{Q}^a_n-\Gamma^a_n(t_n,\bar{Q}_n))(\dot{Q}^b_n-\Gamma^b_n(t_n,\bar{Q}_n))\delta t_n-V(t_n,\bar{Q}_n)\delta t_n. \tag{3.3.33}$$

Notice that if $H$ happens to be the classical Hamiltonian $\mathcal{H}$ and it takes the form (3.3.29) then by the Legendre transformation, the classical Lagrangian is:

$$\mathcal{L}=\frac{1}{2}W^{-1}_{ab}(\dot{Q}^a-\Gamma^a)(\dot{Q}^b-\Gamma^b)-V \tag{3.3.34}$$

and so in this very special case:

$$\lim_{N\to\infty}\sum_{n=0}^{N-1}\delta S_{\mathcal{L}n}=S_{\mathcal{L}} \tag{3.3.35}$$

where $S_{\mathcal{L}}$ is the classical Lagrangian action. As well, factors in the extra product appearing in (3.3.32) are all independent of $Q$. Because of this, in our super special case we can rewrite $U(t_f,Q_f,t_0,Q_0)$ as:

$$U(t_f,Q_f,t_0,Q_0)=\mathcal{N}(t_f,f_0)\int_{Q(t_0)=Q_0}^{Q(t_f)=Q_f}e^{\frac{i}{\hbar}S_{\mathcal{L}}}\mathcal{D}Q \tag{3.3.36}$$

where $\mathcal{N}$ is some complex function independent of $Q$ and we have made the replacement:

$$d^dQ_1...d^dQ_{N-1} \to \mathcal{D}Q. \tag{3.3.37}$$

Because of the unitarity of $\hat{U}$ and the fact that $\mathcal{N}$ does not depend on $Q$, it is possible to cancel $\mathcal{N}$ from any computations of operator summaries:

$$\langle \hat{G} \rangle(t) = \mathrm{Tr}(\hat{\rho}(t)\hat{G}) = \frac{\mathrm{Tr}(\hat{U}(t,0)\hat{\rho}_0\hat{U}^\dagger(t,0)\hat{G})}{\mathrm{Tr}(\hat{U}(t,0)\hat{\rho}_0\hat{U}^\dagger(t,0))} \tag{3.3.38}$$

$$= \frac{\int \langle Q_0|\hat{\rho}_0|Q_0'\rangle\langle Q_0'|\hat{U}^\dagger|Q_f'\rangle\langle Q_f'|\hat{G}|Q_f\rangle\langle Q_f|\hat{U}|Q_0\rangle d^dQ_0 d^dQ_0' d^dQ_f d^dQ_f'}{\int \langle Q_0|\hat{\rho}_0|Q_0'\rangle\langle Q_0'|\hat{U}^\dagger|Q_f\rangle\langle Q_f|\hat{U}|Q_0\rangle d^dQ_0 d^dQ_0' d^dQ_f} \tag{3.3.39}$$

$$= \frac{\int \langle Q_0|\hat{\rho}_0|Q_0'\rangle \int_{Q(0)=Q_0'}^{Q(t)=Q_f'} e^{-\frac{i}{\hbar}S_{\mathcal{L}}} \mathscr{D}Q \langle Q_f'|\hat{G}|Q_f\rangle \int_{Q(0)=Q_0}^{Q(t)=Q_f} e^{\frac{i}{\hbar}S_{\mathcal{L}}} \mathscr{D}Q d^dQ_0 d^dQ_0' d^dQ_f d^dQ_f'}{\int \langle Q_0|\hat{\rho}_0|Q_0'\rangle \int_{Q(0)=Q_0'}^{Q(t)=Q_f} e^{-\frac{i}{\hbar}S_{\mathcal{L}}} \mathscr{D}Q \int_{Q(0)=Q_0}^{Q(t)=Q_f} e^{\frac{i}{\hbar}S_{\mathcal{L}}} \mathscr{D}Q d^dQ_0 d^dQ_0' d^dQ_f}. \tag{3.3.40}$$

Equation (3.3.36) is a very useful special case of the path integral, because half of the integrals have already been performed, but very importantly if $H$ is not of the form of (3.3.29) then it is just a wrong formula. Regardless of the form taken by $H$, one can in principle complete the $P$ integrals in (3.3.22) and leave behind something that now is purely a function of all of the time steps and $Q_n$. In the $N \to \infty$ limit this leaves a function of $t_0, t_f, Q_0, Q_f$ and a functional of the path connecting $Q_0$ and $Q_f$. Thus, formally for some such functional $S'$ and some arbitrary function $\mathcal{N}'(t_f, t_0)$:

$$\mathcal{N}'(t_f,t_0)e^{\frac{i}{\hbar}S'[Q]} = \lim_{N\to\infty} \int e^{\frac{i}{\hbar}\sum_{n=0}^{N-1}\delta S_n} \frac{d^dP_0}{(2\pi\hbar)^d}\frac{d^dP_1}{(2\pi\hbar)^d}...\frac{d^dP_{N-1}}{(2\pi\hbar)^d}. \tag{3.3.41}$$

From this it automatically follows that:

$$U(t_f,Q_f,t_0,Q_0) = \mathcal{N}'(t_f,t_0)\int_{Q(t_0)=Q_0}^{Q(t_f)=Q_f} e^{\frac{i}{\hbar}S'[Q]}\mathcal{D}Q. \tag{3.3.42}$$

This of course bears a striking resemblance to (3.3.36). However, there is absolutely no reason to believe that $S'[Q]$ has any relationship whatsoever to the classical Lagrangian action in general. In fact, in general you cannot even guarantee that it is possible to select $\mathcal{N}'$ so that the functional $S'[Q]$ is real-valued (which of course $S_{\mathcal{L}}[Q]$ always is) nor do you have any reasonable expectation of $S'[Q]$ generically manifesting the symmetries of your theory in an obvious way.

### 3.3.3 Path Integral for Vacuum Expectation Values

The path integral is especially useful for computing so-called "vacuum expectation values" or vevs for short. Many quantities of interest in quantum mechanics, in particular scattering amplitudes, can be computed from vevs. We will see precisely how over the coming chapters. A vev specifically is an operator summary, typically of operators in the active picture, taken in the ground state (lowest energy state) of the quantum system. Let us now establish how vevs are related to the path integral. Let $\hat{Q}_A^a(t)$ represent our coordinate operators in the active picture. As always, the $\hat{Q}_A^a(t)$ are related to our usual coordinate operators by:

$$\hat{Q}_A^a(t) = \hat{U}^\dagger(t,t_0)\hat{Q}^a\hat{U}(t,t_0) \tag{3.3.43}$$

where $t_0$ is the time at which we choose to define the passive picture coordinate operators $\hat{Q}^a$. At each $t$, we can define normalized eigenstates $|Q(t)\rangle$ of the $\hat{Q}_A^a(t)$ operators:

$$\hat{Q}_A^a(t)|Q(t)\rangle = Q^a|Q(t)\rangle \tag{3.3.44}$$

$$\implies \hat{Q}^a(\hat{U}(t,t_0)|Q(t)\rangle) = Q^a(\hat{U}^(t,t_0)|Q(t)\rangle) \tag{3.3.45}$$

$$\implies |Q(t_0)\rangle = \hat{U}(t,t_0)|Q(t)\rangle. \tag{3.3.46}$$

Immediately it follows that:

$$\langle Q(t)|Q_0(t_0)\rangle = U(t,Q,t_0,Q_0) \tag{3.3.47}$$

and so these time-dependent eigenstates are closely related to the path integral.

Now suppose we have some list of operators in the active picture $\hat{\mathcal{O}}_n(t_n),...,\hat{\mathcal{O}}_1(t_1)$ where $t_n > t_{n-1} > ... > t_2 > t_1$. We can use the path integral to compute the matrix elements of the time ordered product of

these operators. Let $T > t_n$ and $T_0 = t_0 < t_1$. Then:

$$\langle Q(T)|\hat{\mathcal{O}}_n(t_n)...\hat{\mathcal{O}}_1(t_1)|Q_0(T_0)\rangle$$
$$= \langle Q(t_0)|\hat{U}(T,t_0)\hat{U}^\dagger(t_n,t_0)\hat{\mathcal{O}}_n(t_0)\hat{U}(t_n,t_0)...\hat{U}^\dagger(t_1,t_0)\hat{\mathcal{O}}_1(t_0)\hat{U}(t_1,t_0)\hat{U}^\dagger(T_0,t_0)|Q_0(t_0)\rangle \tag{3.3.48}$$
$$= \langle Q|\hat{U}(T,t_n)\hat{\mathcal{O}}_n\hat{U}(t_n,t_{n-1})...\hat{\mathcal{O}}_2\hat{U}(t_2,t_1)\hat{\mathcal{O}}_1\hat{U}(t_1,t_0)|Q_0\rangle \tag{3.3.49}$$

where in the last line we stop writing that the operators or eigenstates are evaluated at $t_0$ because we can simply view them as in the passive picture at that time. We can now insert the Wigner representation of each operator in this product to achieve:

$$\langle Q_f(T)|\hat{\mathcal{O}}_n(t_n)...\hat{\mathcal{O}}_1(t_1)|Q_0(T_0)\rangle$$
$$= \int U(T,Q_f,t_n,Q'_n)\prod_{k=1}^{n}\left(e^{\frac{i}{\hbar}P_{ka}(Q'^a_k-Q^a_k)}\mathcal{O}_k\left(\frac{Q_k+Q'_k}{2},P_k\right)U(t_k,Q_k,t_{k-1},Q_{k-1})\right)$$
$$\cdot\, d^dQ'_1\frac{d^dQ_1d^dP_1}{(2\pi\hbar)^d}...d^dQ'_n\frac{d^dQ_nd^dP_n}{(2\pi\hbar)^d}. \tag{3.3.50}$$

Each factor of $U$ in this expression will give a path integral between the different times of operator evaluation. We then see that between these path integrals the operator's Wigner function is evaluated at the averaged coordinate and that the momentum and coordinate different are integrated over. Therefore, we see that ultimately this is also of the form of an integration over all paths between $Q_0$ at $T_0$ and $Q_f$ at $T$ in which the operators to be evaluated are inserted as Wigner functions at the correct time step. The exponential factor with the momentum is simply the correct contribution to the Wigner action functional appropriate for change between $Q_k$ and $Q'_k$ without a change in time. Thus, this formula when fully expressed in terms of the path integral simply becomes:

$$\langle Q_f(T)|\mathscr{T}:\hat{\mathcal{O}}_n(t_n)...\hat{\mathcal{O}}_1(t_1)|Q_0(T_0)\rangle = \int_{Q(T_0)=Q_0}^{Q(T)=Q_f} e^{\frac{i}{\hbar}S[Q,P]}\mathcal{O}_n(Q(t_n),P(t_n))...\mathcal{O}_1(Q(t_1),P(t_1))\mathcal{D}Q\mathcal{D}P \tag{3.3.51}$$

where we have manually written in the time ordering for clarity of when this formula is valid.

Equation (3.3.51) is not obviously a vev, but we will now see that it is very closely related to one. Suppose we have a quantum system with a time independent Hamiltonian operator $\hat{H}$ and let $|n\rangle$ be an orthonormal basis of eigenstates of $\hat{H}$ and $\psi_n(Q) = \langle Q|n\rangle$ be the corresponding eigenstate wavefunctions. Let $|\Omega\rangle$ be used to denote the ground state (lowest energy state) among these and let its energy eigenvalue be $E_\Omega$. Assume our system is such that this ground state is unique. Let $\hat{\mathcal{O}} = \hat{\mathcal{O}}_n(t_n)...\hat{\mathcal{O}}_1(t_1)$ be a shorthand for our operator from before. Let $T, t, T_0, t_0$ be times such that $T > t > t_n$, $t_1 > t_0 > T_0$. Then, time evolution from $T_0$ to $T$ can be broken into three sections. In the first, from $T_0$ to $t_0$, we simply exponentiate the Hamiltonian operator. In particular, we have:

$$\langle Q(T)|Q'(t)\rangle = \langle Q(t_0)|\hat{U}(T,t_0)\hat{U}^\dagger(t,t_0)|Q'(t_0)\rangle = \langle Q(t_0)|e^{-\frac{i}{\hbar}\hat{H}(T-t)}|Q'(t_0)\rangle = \sum_n e^{-\frac{i}{\hbar}E_n(T-t)}\psi_n(Q)\psi^*_n(Q'). \tag{3.3.52}$$

In the second, from $t_0$ to $t$, we will not explicitly evaluate the time evolution. Finally, from $t$ to $T$ we again exponentiate. Thus:

$$\langle Q(T)|\hat{\mathcal{O}}|Q_0(T_0)\rangle = \int\langle Q(T)|Q'(t)\rangle\langle Q'(t)|\hat{\mathcal{O}}|Q'_0(t_0)\rangle\langle Q'_0(t_0)|Q_0(T_0)\rangle d^dQ'_0d^dQ' \tag{3.3.53}$$
$$= \int\sum_{nn'} e^{-\frac{i}{\hbar}E_n(T-t)-\frac{i}{\hbar}E_{n'}(t_0-T_0)}\psi_n(Q)\psi^*_n(Q')\psi_{n'}(Q'_0)\psi^*_{n'}(Q_0)\langle Q'(t)|\hat{\mathcal{O}}|Q'_0(t_0)\rangle d^dQ'_0d^dQ'. \tag{3.3.54}$$

We are now going to do a subtle trick which is possible specifically because the Hamiltonian is bounded below (so that it is capable of thermodynamic equilibrium, which is required by the experimental success of thermodynamics). This trick is called a **Wick rotation**. The Schrödinger equation, or equivalently the von Neumann equation, gives time evolution in a way which formally depends analytically on time, even if the time parameter is viewed as a complex number. So, for times between $t$ and $T$ and between $t_0$ and $T_0$ we can "tilt the time axis" in the complex plane so that our ultimate initial time is $T_0e^{-i\delta}$ and our final time is $Te^{-i\delta}$ with $\delta$ a small positive real number. By *can* do this, I mean that the path integral formula (3.3.51) will still

hold with these replacements because the realness of $t$ was never used in the calculation. That is, we have:

$$\langle Q_f(Te^{-i\delta})|\mathscr{T}:\hat{\mathcal{O}}_n(t_n)...\hat{\mathcal{O}}_1(t_1)|Q_0(T_0e^{-i\delta})\rangle = \int_{Q(T_0e^{-i\delta})=Q_0}^{Q(Te^{-i\delta})=Q_f} e^{\frac{i}{\hbar}S[Q,P]}\mathcal{O}_n...\mathcal{O}_1\mathcal{D}Q\mathcal{D}P. \tag{3.3.55}$$

Applying this principle to (3.3.54) we have:

$$\frac{\langle Q(Te^{-i\delta})|\hat{\mathcal{O}}|Q_0(T_0e^{-i\delta})\rangle}{\langle Q(Te^{-i\delta})|Q_0(T_0e^{-i\delta})\rangle} = \frac{\int \sum_{nn'} e^{-\frac{i}{\hbar}E_n(Te^{-i\delta}-t)-\frac{i}{\hbar}E_{n'}(t_0-T_0e^{-i\delta})}\psi_n(Q)\psi_n^*(Q')\psi_{n'}(Q_0')\psi_{n'}^*(Q_0)\langle Q'(t)|\hat{\mathcal{O}}|Q_0'(t_0)\rangle d^dQ_0'd^dQ'}{\sum_n e^{-\frac{i}{\hbar}E_n(T-T_0)e^{-i\delta}}\psi_n(Q)\psi_n^*(Q_0)}. \tag{3.3.56}$$

From here we can take the limits $T\to\infty$ and $T_0\to-\infty$ and because of $\delta$, all but the ground state terms in our sum will exponentially decay. (The ground state terms also decay, but the decay rate cancels between then numerator and denominator while all higher energy terms decay more quickly). Thus, taking the limit:

$$\lim_{\substack{T\to\infty\\T_0\to-\infty}} \frac{\langle Q(Te^{-i\delta})|\hat{\mathcal{O}}|Q_0(T_0e^{-i\delta})\rangle}{\langle Q(Te^{-i\delta})|Q_0(T_0e^{-i\delta})\rangle} = \int e^{\frac{i}{\hbar}E_\Omega(t-t_0)}\psi_\Omega^*(Q')\psi_\Omega(Q_0')\langle Q'(t)|\hat{\mathcal{O}}|Q_0'(t_0)\rangle d^dQ_0'd^dQ' \tag{3.3.57}$$

$$= \int e^{\frac{i}{\hbar}E_\Omega(t-t_0)}\langle\Omega|Q'\rangle\langle Q'|\hat{U}(t,t_0)\hat{\mathcal{O}}|Q_0'\rangle\langle Q_0'|\Omega\rangle d^dQ_0'd^dQ' \tag{3.3.58}$$

$$= \langle\Omega|\hat{\mathcal{O}}|\Omega\rangle \tag{3.3.59}$$

which is precisely a vev. Because vevs will be common, we introduce the shorthand for the time ordered vacuum expectation value:

$$\boxed{\texttt{Vev}(\hat{\mathcal{O}}_n(t_n)...\hat{\mathcal{O}}_1(t_1)) = \langle\Omega|\mathscr{T}:\hat{\mathcal{O}}_n(t_n)...\hat{\mathcal{O}}_1(t_1))|\Omega\rangle.} \tag{3.3.60}$$

Because of the time ordering used in the definition of the `Vev`, the operators within it commute if evaluated at different times.

By combining (3.3.55) and (3.3.59) we have:

$$\texttt{Vev}(\hat{\mathcal{O}}_n(t_n)...\hat{\mathcal{O}}_1(t_1)) = \lim_{\delta\to0^+}\lim_{\substack{T\to\infty\\T_0\to-\infty}} \frac{\int_{Q(T_0e^{-i\delta})=Q_0}^{Q(Te^{-i\delta})=Q_f} e^{\frac{i}{\hbar}S[Q,P]}\mathcal{O}_n...\mathcal{O}_1\mathcal{D}Q\mathcal{D}P}{\int_{Q(T_0e^{-i\delta})=Q_0}^{Q(Te^{-i\delta})=Q_f} e^{\frac{i}{\hbar}S[Q,P]}\mathcal{D}Q\mathcal{D}P} \tag{3.3.61}$$

This equation is very interesting because the left hand side does not depend on $Q_f$ or $Q_0$ while the right hand side does. Thus, the boundary conditions $Q_f$ and $Q_0$ are arbitrary, provided they are consistent between the numerator and denominator. Almost all practical calculations involving the path integral are done as perturbations around a quantum harmonic oscillator, and the standard boundary conditions to take in such a case are to set the initial and final coordinates to the equilibrium position of the oscillator. Just as with the Fourier transform, the $\delta$ damping is typically not written explicitly unless it is of specific computational importance for a particular equation. That said, you should not forget that it is there when it is time to be careful. In fact, the $\delta$ damping can be interpreted directly in terms of the Fourier transform. Consider the use of a new time coordinate $t'$ for times outside of the interval $(t,t_0)$ which is related to our previous time coordinate by $t'=e^{i\delta}t$. Now the initial and final values of $t'$ are $T_0$ and $T$ respectively, which are purely real. A function of frequency $\omega$ when written in terms of $t$, meaning containing a component $e^{-i\omega t}$, then becomes a function of frequency $\omega'=e^{-i\delta}\omega$ so that $e^{-i\omega t}=e^{-i\omega' t'}$. Thus the Wick rotation can also be viewed as adding a small negative imaginary component to the frequency, precisely how the usual Fourier damping works. Understanding these finer points of the meaning of (3.3.61), we write more easily and less clearly:

$$\texttt{Vev}(\hat{\mathcal{O}}_n(t_n)...\hat{\mathcal{O}}_1(t_1)) = \frac{\int e^{\frac{i}{\hbar}S[Q,P]}\mathcal{O}_n(Q(t_n),P(t_n))...\mathcal{O}_1(Q(t_1),P(t_1))\mathcal{D}Q\mathcal{D}P}{\int e^{\frac{i}{\hbar}S[Q,P]}\mathcal{D}Q\mathcal{D}P}. \tag{3.3.62}$$

Because the path integral integrates over the entire path through phase space, we can nicely reexpress this result in terms of our usual $\Phi(t)=(Q(t),P(t))$ condensed phase space coordinates as:

$$\boxed{\texttt{Vev}(\hat{\mathcal{O}}_n(t_n)...\hat{\mathcal{O}}_1(t_1)) = \frac{\int e^{\frac{i}{\hbar}S[\Phi]}\mathcal{O}_n(\Phi(t_n))...\mathcal{O}_1(\Phi(t_1))\mathcal{D}\mathbb{P}}{\int e^{\frac{i}{\hbar}S[\Phi]}\mathcal{D}\mathbb{P}}.} \tag{3.3.63}$$

As usual, generators of interest are typically analytic functions of our phase space variables. If all of the

$\hat{\mathcal{O}}_k$ above are such, we can replace them by their series representations in terms of the $\Phi^a$ and thus develop a series representation of the above vev in terms of vevs of products of the $\Phi^a$ evaluated at various times. Such vevs are called the **Green's functions** of the quantum theory. So, the typical vevs of most interest are:

$$\boxed{G_n^{a_1...a_n}(t_1,...,t_n) = \texttt{Vev}(\hat{\Phi}^{a_1}(t_1)...\hat{\Phi}^{a_n}(t_n)).} \tag{3.3.64}$$

(The Green's function with $n$ copies of $\Phi^a(t)$ is typically called the $n$-point Green's function or $n$-point correlator.) All of these Green's functions can be recovered from a single functional which is called the generating functional of the quantum theory. Let $J_a(t)$ be a smooth function of compact support which transforms inversely to the $\Phi^a(t)$ under a canonical transformation. Then, define the functional:

$$\mathcal{Z}[J] = \langle\Omega|\mathscr{T} : e^{\frac{i}{\hbar}\int_{-\infty}^{\infty} J_a(t)\hat{\Phi}^a(t)dt}|\Omega\rangle. \tag{3.3.65}$$

Series expanding in powers of $J_a(t)$, we find:

$$\mathcal{Z}[J] = \sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{i}{\hbar}\right)^n\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}J_{a_1}(t_1)...J_{a_n}(t_n)\langle\Omega|\mathscr{T} : \hat{\Phi}^{a_1}(t_1)...\hat{\Phi}^{a_n}(t_n)|\Omega\rangle dt_1...dt_n. \tag{3.3.66}$$

Thus:

$$\boxed{G_n^{a_1...a_n}(t_1,...,t_n) = \left(\frac{\hbar}{i}\right)^n\frac{\delta^n\mathcal{Z}[J]}{\delta J_{a_1}(t_1)...\delta J_{a_n}(t_n)}\bigg|_{J=0}.} \tag{3.3.67}$$

As well, by using (3.3.63) in the definition of $\mathcal{Z}[J]$, we have:

$$\mathcal{Z}[J] = \frac{\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{i}{\hbar}\right)^n\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}J_{a_1}(t_1)...J_{a_n}(t_n)\int e^{\frac{i}{\hbar}S[\Phi]}\Phi^{a_1}(t_1)...\Phi^{a_n}(t_n)\mathcal{D}\mathbb{P}dt_1...dt_n}{\int e^{\frac{i}{\hbar}S[\Phi]}\mathcal{D}\mathbb{P}} \tag{3.3.68}$$

$$\implies \boxed{\mathcal{Z}[J] = \frac{\int e^{\frac{i}{\hbar}\left(S[\Phi]+\int_{-\infty}^{\infty}J_a(t)\Phi^a(t)dt\right)}\mathcal{D}\mathbb{P}}{\int e^{\frac{i}{\hbar}S[\Phi]}\mathcal{D}\mathbb{P}}.} \tag{3.3.69}$$

The denominator is simply the numerator evaluated at $J_a(t) = 0$. Thus, by completing the single path integral for general $J_a(t)$ in the numerator we can determine $\mathcal{Z}[J]$. From $\mathcal{Z}[J]$ we can determine all of the Green's functions of the theory by using (3.3.67). From those Green's functions we can calculate (3.3.63) for any $\hat{\mathcal{O}}_k$ which have Wigner functions that are analytic in the coordinates and momenta. The vevs in (3.3.63) can be used to determine almost all quantities of interest in many quantum mechanical models. Thus, the ultimate utility of the path integral is in setting up $\mathcal{Z}[J]$ as the central quantity of interest for further calculation.

The quantum generating functional isn't really computed by literally evaluating the path integral. Instead, in practical terms it is computed by perturbing a few known path integral results (Gaussian – the quantum harmonic oscillator and other very simple systems) systematically. Suppose that the Wigner action functional can be split into a solvable part $S_0$ and a perturbation $S_1$ which may reasonably be considered small (solvable here means that if that were the only contribution to the Wigner action functional then the resulting path integral would be doable explicitly):

$$S[\Phi] = S_0[\Phi] + S_1[\Phi]. \tag{3.3.70}$$

Let $\mathcal{Z}_0[J]$ be the generating functional computed when using $S_0[\Phi]$. Then, recklessly commuting limits:

$$\int e^{\frac{i}{\hbar}\left(S[\Phi]+\int_{-\infty}^{\infty}J_a(t)\Phi^a(t)dt\right)}\mathcal{D}\mathbb{P} = \int e^{\frac{i}{\hbar}S_1\left[\frac{\hbar}{i}\frac{\delta}{\delta J}\right]}e^{\frac{i}{\hbar}\left(S_0[\Phi]+\int_{-\infty}^{\infty}J_a(t)\Phi^a(t)dt\right)}\mathcal{D}\mathbb{P} \tag{3.3.71}$$

$$\implies \mathcal{Z}[J] = \mathcal{N}e^{\frac{i}{\hbar}S_1\left[\frac{\hbar}{i}\frac{\delta}{\delta J}\right]}\mathcal{Z}_0[J] \tag{3.3.72}$$

where by evaluating the functional $S_1$ at $\frac{\hbar}{i}\frac{\delta}{\delta J}$ we mean that we express $S_1$ formally as a series and replace its argument with this functional derivative at each order and $\mathcal{N}$ is a normalization factor necessary because by definition $\mathcal{Z}[0] = 1$. Many limits were commuted in this calculation and the effect of this is that (3.3.72) is not typically exactly true. Instead, the right hand side of (3.3.72) can be expanded in powers of $S_1$ to give a reasonable perturbation series and that will be the correct perturbative expansion of the left hand side. The issue is that it is unlikely that the full series converges to the true exact $\mathcal{Z}[J]$, that is, there is no reason to believe $\mathcal{Z}[J]$ depends analytically on whatever series expansion parameter is used when considering $S_1$ as small. Perturbation theory is almost always this way and the resulting series are called **asymptotic series**. Asymptotic series are extremely useful for getting excellent approximations to true answers, even if the series does not converge to the true answer (and possibly doesn't converge at all) because in any practical

calculation we will only compute to finite order of perturbation theory (and typically only a very low order).

The generating functional satisfies a sort of "equation of motion" called the **Schwinger-Dyson equation**. Consider:

$$\int e^{\frac{i}{\hbar}\left(S[\Phi]+\int_{-\infty}^{\infty} J_b\Phi^b dt'\right)}\frac{\delta S}{\delta\Phi^a(t)}[\Phi]\mathcal{D}\mathbb{P} = \frac{\hbar}{i}\int e^{\frac{i}{\hbar}\int_{-\infty}^{\infty} J_b\Phi^b dt'}\frac{\delta}{\delta\Phi^a(t)}e^{\frac{i}{\hbar}S[\Phi]}\mathcal{D}\mathbb{P}. \tag{3.3.73}$$

We now perform functional integration by parts. The boundary terms of the integration are necessarily negligible for the path integral to converge. This situation is helped by the $\delta$ damping from before. For large $\Phi$, $S$ is typically also large. Then, because $t$ has been slightly rotated it develops a small imaginary part so that the argument of the exponential develops a negative real part and so decays for large $\Phi$. Thus we have:

$$\int e^{\frac{i}{\hbar}\left(S[\Phi]+\int_{-\infty}^{\infty} J_b\Phi^b dt'\right)}\frac{\delta S}{\delta\Phi^a(t)}[\Phi]\mathcal{D}\mathbb{P} = -\frac{\hbar}{i}\int e^{\frac{i}{\hbar}S[\Phi]}\frac{\delta}{\delta\Phi^a(t)}e^{\frac{i}{\hbar}\int_{-\infty}^{\infty} J_b\Phi^b dt'}\mathcal{D}\mathbb{P} \tag{3.3.74}$$

$$= -J_a(t)\int e^{\frac{i}{\hbar}\left(S[\Phi]+\int_{-\infty}^{\infty} J_b\Phi^b dt'\right)}\mathcal{D}\mathbb{P}. \tag{3.3.75}$$

Alternatively:

$$\frac{\delta S}{\delta\Phi^a}\left[\frac{\hbar}{i}\frac{\delta}{\delta J}\right]\int e^{\frac{i}{\hbar}\left(S[\Phi]+\int_{-\infty}^{\infty} J_b\Phi^b dt'\right)}\mathcal{D}\mathbb{P} = \int e^{\frac{i}{\hbar}\left(S[\Phi]+\int_{-\infty}^{\infty} J_b\Phi^b dt'\right)}\frac{\delta S}{\delta\Phi^a(t)}[\Phi]\mathcal{D}\mathbb{P}. \tag{3.3.76}$$

With these together then, we have the Schwinger-Dyson equation:

$$\boxed{\frac{\delta S}{\delta\Phi^a(t)}\left[\frac{\hbar}{i}\frac{\delta}{\delta J}\right]\mathcal{Z}[J] = -J_a\mathcal{Z}[J].} \tag{3.3.77}$$

This functional differential equation can be solved order by order in $J_a$ and order by order in perturbation theory to determine the quantum generating functional. Doing so explicitly we will save for later chapters.

### 3.3.4 Path Integral for the Quantum Harmonic Oscillator

One of the only path integrals that can be computed exactly is the quantum harmonic oscillator. Consequently, perturbation theory for the path integral is essentially only useful when perturbing near the qho limit. Suppose we have a system of interacting quantum harmonic oscillators. The resulting Wigner Hamiltonian function is then a quadratic polynomial in the coordinates and momenta. The Wigner action functional then is automatically a quadratic functional of the coordinates and momenta (as the only additional contribution beyond the integral of the Wigner Hamiltonian is the integral of $P_a\dot{Q}^a$ which is also quadratic). By a quadratic functional, I mean that the third functional derivative vanishes everywhere:

$$\frac{\delta^3 S}{\delta\Phi^a(t)\delta\Phi^b(t')\delta\Phi^c(t'')} = 0. \tag{3.3.78}$$

We will see that the entire path integral for the qho is determined by its 2-point Green's function, which is given a special name as the **Feynman propagator** and written as $D^{ab}(t,t')$:

$$\boxed{D^{ab}(t,t') = G_2^{ab}(t,t') = \langle\Omega|\mathscr{T}:\hat{\Phi}^a(t)\hat{\Phi}^b(t')|\Omega\rangle.} \tag{3.3.79}$$

The reason for designating the Feynman propagator with a special symbol is that when doing perturbative calculations the Feynman propagator of the unperturbed theory will appear all over the place.

Define $S_J[\Phi]$ as the functional:

$$S_J[\Phi] = S[\Phi] + \int_{-\infty}^{\infty} J_a(t)\Phi^a(t)dt. \tag{3.3.80}$$

Automatically it follows that: $S_J$ is also a quadratic functional:

$$\frac{\delta^3 S_J}{\delta\Phi^a(t)\delta\Phi^b(t')\delta\Phi^c(t'')} = 0. \tag{3.3.81}$$

When $S_J$ is viewed as an action, there will be some classical solution $\phi_J(t)$ achieved by taking the stationary variation of $S_J$ with initial and final conditions $\phi_J = 0$ at the times $T_0$ and $T$ (which we eventually will take to $-\infty, +\infty$) and solving the Euler-Lagrange equations of motion:

$$\frac{\delta S_J}{\delta\Phi^a}[\phi_J] = 0. \tag{3.3.82}$$

We can write any trajectory $\Phi^a(t)$ as a deviation from that solution as:

$$\Phi^a(t) = \phi_J^a(t) + \delta\Phi^a(t). \tag{3.3.83}$$

Decomposed this way, we view $\phi_J^a(t)$ as a shift of our integration variables and $\delta\Phi^a(t)$ as our new integration variables. Because $S_J[\Phi]$ is a quadratic functional, its functional Taylor series terminates after only 3 terms:

$$S_J[\Phi] = S_J[\phi_J+\delta\Phi] = S_J[\phi_J] + \int_{-\infty}^{\infty} \delta\Phi^a(t)\frac{\delta S_J}{\delta\Phi^a(t)}[\phi_J]dt + \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta\Phi^a(t)\delta\Phi^b(t')\frac{\delta^2 S_J}{\delta\Phi^a(t)\delta\Phi^b(t')}[\phi_J]dt'dt. \tag{3.3.84}$$

The second term on the right hand side vanishes identically by (3.3.82). As well, the second functional derivative in the third term is automatically independent of both $J$. This is because it is independent of being evaluated at $\phi_J$ as by taking another functional derivative we achieve 0 by (3.3.81) and the only other $J$ dependent piece in $S_J$ is linear in $\Phi$ (as introduced in (3.3.80)) which doesn't contribute to the second functional derivative. Therefore, for some $M_{ab}(t,t')$:

$$M_{ab}(t,t') = \frac{\delta^2 S_J}{\delta\Phi^a(t)\delta\Phi^b(t')}[\phi_J] = \frac{\delta^2 S}{\delta\Phi^a(t)\delta\Phi^b(t')}[\Phi]. \tag{3.3.85}$$

With these simplifications we have:

$$S_J[\Phi] = S_J[\phi_J] + \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta\Phi^a\delta\Phi^b M_{ab}(t,t')dt'dt. \tag{3.3.86}$$

With this, we can now evaluate the path integral very simply, letting $\phi_0$ mean $\phi_J$ with $J = 0$:

$$\mathcal{Z}[J] = \frac{\int e^{\frac{i}{\hbar}S_J[\Phi]}\mathcal{D}\Phi}{\int e^{\frac{i}{\hbar}S[\Phi]}\mathcal{D}\Phi} = \frac{e^{\frac{i}{\hbar}S_J[\phi_J]}\int e^{\frac{i}{2\hbar}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta\Phi^a\delta\Phi^b M_{ab}(t,t')dt'dt}\mathcal{D}\delta\Phi}{e^{\frac{i}{\hbar}S[\phi_0]}\int e^{\frac{i}{2\hbar}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta\Phi^a\delta\Phi^b M_{ab}(t,t')dt'dt}\mathcal{D}\delta\Phi} = e^{\frac{i}{\hbar}(S_J[\phi_J]-S[\phi_0])}. \tag{3.3.87}$$

The actual integrations completely cancel!

In order to fully simplify $\mathcal{Z}[J]$, we need to evaluate $S_J[\phi_J]$. Let us assume we have defined our harmonic oscillator variables so that the minimum of the Wigner Hamiltonian is at the origin of phase space and let us set its minimum value to 0. Then, automatically the linear in $\Phi$ pieces of $S[\Phi]$ are 0. From (3.3.85) then we have:

$$S[\Phi] = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} M_{ab}(t,t')\Phi^a(t)\Phi^b(t')dtdt'. \tag{3.3.88}$$

From this it follows that (3.3.82) becomes:

$$\int_{-\infty}^{\infty} M_{ab}(t,t')\phi_J^b(t')dt' = -J_a(t). \tag{3.3.89}$$

This is a linear equation, which in actual practice becomes a linear differential equation ($M_{ab}(t,t')$ will be a sum of Dirac deltas and derivatives of deltas of $t-t'$). The initial and final condition on $\phi_J$ are arbitrary (as long as they are the same for $\phi_0$) and we take $\phi_J(T_0) = \phi_J(T) = 0$ for convenience. This differential equation with those boundary conditions will have some solution and due to linearity, it will be linear in $J_a(t)$. Therefore, for some function $\mathcal{G}^{ab}(t,t')$:

$$\phi_J^a(t) = \int_{-\infty}^{\infty} \mathcal{G}^{ab}(t,t')J_b(t')dt'. \tag{3.3.90}$$

We can view $M_{ab}(t,t')$ and $\mathcal{G}^{ab}(t,t')$ as operators on functions (which act by multiplication followed by integration over $t'$) and in this view $\mathcal{G}^{ab}(t,t')$ is the operator inverse (up to a minus sign) of $M_{ab}(t,t')$ on the space of functions with the specified boundary conditions. Plainly, $\mathcal{G}^{ab}$ is the Green's function of the differential equation. Define the functional $W[J]$ by:

$$\mathcal{Z}[J] = e^{\frac{i}{\hbar}W[J]}. \tag{3.3.91}$$

For our qho it automatically follows that:

$$W[J] = S_J[\phi_J] - S[\phi_0] = S[\phi_J] - S[\phi_0] + \int_{-\infty}^{\infty} J_a(t)\phi_J^a(t)dt. \tag{3.3.92}$$

Because $\phi_J$ is linear in $J_a(t)$ and $S[\Phi]$ is quadratic in $\Phi$, we immediately have that $W[J]$ is quadratic in $J_a(t)$.

Therefore, its functional Taylor series truncates after only three terms:

$$W[J] = W[0] + \int_{-\infty}^{\infty} \frac{\delta W}{\delta J_a(t)}[0] J_a(t) dt + \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\delta^2 W}{\delta J_a(t)\delta J_b(t')}[0] J_a(t) J_b(t') dt dt'. \tag{3.3.93}$$

From the definition of $W[J]$ it is immediate that $W[0] = 0$. Differentiating the definition of $W[J]$ gives:

$$\frac{\delta W}{\delta J_a(t)}[J] = \phi_J^a(t) + \int_{-\infty}^{\infty} \frac{\delta \phi_J^b(t')}{\delta J_a(t)} \left( \frac{\delta S}{\delta \Phi^b(t')}[\phi_J] + J_b(t') \right) dt' = \phi_J^a(t) = \int_{-\infty}^{\infty} \mathcal{G}^{ab}(t,t') J_b(t') dt' \tag{3.3.94}$$

after using (3.3.82) and (3.3.90). Thus:

$$\frac{\delta W}{\delta J_a(t)}[0] = 0, \qquad \frac{\delta^2 W}{\delta J_a(t)\delta J_b(t')} = \mathcal{G}^{ab}(t,t'). \tag{3.3.95}$$

Returning to (3.3.93) we have:

$$W[J] = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{G}^{ab}(t,t') J_a(t) J_b(t') dt dt'. \tag{3.3.96}$$

Using (3.3.67), (3.3.79), and (3.3.91) we see:

$$D^{ab}(t,t') = -i\hbar \mathcal{G}^{ab}(t,t'). \tag{3.3.97}$$

Therefore, for a system of interacting quantum harmonic oscillators, the entire path integral is given by:

$$\boxed{\mathcal{Z}[J] = e^{-\frac{1}{2\hbar^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} D^{ab}(t,t') J_a(t) J_b(t') dt' dt}} \tag{3.3.98}$$

where $D^{ab}(t,t')$ is just the Feynman propagator of the oscillators.

We will now explicitly compute the Feynman propagator for a system of harmonic oscillators. The general Wigner Hamiltonian function for a system of quantum harmonic oscillators takes the form of Orig. (1.15.6) without the need for higher order terms (and using that we have set the origin of the phase space as the minimum of the Wigner Hamiltonian). By the theorem of normal modes (Orig. Theorem 1.14) there exists a constant symplectic matrix $\hat{\Lambda}$ which relate our coordinates and momenta $(Q,P)$ to new coordinates and momenta $(\mathfrak{q},\mathfrak{p})$ by:

$$\Phi = \begin{pmatrix} Q \\ P \end{pmatrix} = \hat{\Lambda} \begin{pmatrix} \mathfrak{q} \\ \mathfrak{p} \end{pmatrix} = \hat{\Lambda}\phi \tag{3.3.99}$$

so that the Wigner Hamiltonian expressed in terms of these new coordinates takes the form:

$$H = \sum_n \frac{\mathfrak{p}_n^2 + \omega_n^2 \mathfrak{q}_n^2}{2}. \tag{3.3.100}$$

Let $d^{ab}(t,t')$ be the Feynman propagator for the normal mode coordinates $(\mathfrak{q},\mathfrak{p})$:

$$d^{ab}(t,t') = \langle \Omega | \mathscr{T} : \hat{\phi}^a(t) \hat{\phi}^b(t') | \Omega \rangle. \tag{3.3.101}$$

Immediately then:

$$D^{ab}(t,t') = \Lambda^a{}_c \Lambda^b{}_d d^{cd}(t,t'). \tag{3.3.102}$$

Because each of the oscillators $(\mathfrak{q},\mathfrak{p})$ are decoupled, $d^{ab}(t,t')$ simply decomposes into a list of Feynman propagators each for an individual oscillator. So, all we need to do is compute the Feynman propagator for a single oscillator.

For an individual oscillator of angular frequency $\varpi$ the Feynman propagator can be computed easily using standard quantum mechanical techniques. The Hamiltonian operator is:

$$\hat{H} = \frac{\hat{\mathfrak{p}}^2}{2} + \frac{1}{2}\varpi^2 \hat{\mathfrak{q}}^2 = \hbar\omega \left( \hat{a}^\dagger \hat{a} + \frac{1}{2}\mathbb{1} \right) \tag{3.3.103}$$

where we as always define the lowering operator as:

$$\hat{a} = \sqrt{\frac{\varpi}{2\hbar}}\hat{\mathfrak{q}} + \frac{i}{\sqrt{2\varpi\hbar}}\hat{\mathfrak{p}}. \tag{3.3.104}$$

Using the Schrödinger equation it quickly follows that the active picture lowering operator $\hat{a}(t)$ satisfies:

$$i\hbar \frac{d\hat{a}(t)}{dt} = [\hat{a}(t), \hat{H}] = \hbar\omega \hat{a}(t) \implies \hat{a}(t) = e^{-i\omega t}\hat{a}_0. \tag{3.3.105}$$

The coordinate and momenta operators become simply:

$$\hat{\mathfrak{q}}(t) = \sqrt{\frac{\hbar}{2\varpi}}(\hat{a}(t) + \hat{a}^\dagger(t)), \qquad \hat{\mathfrak{p}}(t) = \frac{d\mathfrak{q}(t)}{dt}. \tag{3.3.106}$$

The 2-point vevs of $\hat{a}(t)$ and $\hat{a}^\dagger(t)$ are:

$$\texttt{Vev}(\hat{a}(t)\hat{a}(t')) = \vartheta(t-t')e^{-i\varpi(t+t')}\langle 0|\hat{a}_0^2|0\rangle + \vartheta(t'-t)e^{-i\varpi(t+t')}\langle 0|\hat{a}_0^2|0\rangle = 0 \tag{3.3.107}$$

$$\texttt{Vev}(\hat{a}(t)\hat{a}^\dagger(t')) = \vartheta(t-t')e^{-i\varpi(t-t')}\langle 0|\hat{a}_0\hat{a}_0^\dagger|0\rangle + \vartheta(t'-t)e^{-i\varpi(t-t')}\langle 0|\hat{a}_0^\dagger\hat{a}_0|0\rangle = e^{-i\varpi(t-t')}\vartheta(t-t') \tag{3.3.108}$$

$$\texttt{Vev}(\hat{a}^\dagger(t)\hat{a}^\dagger(t')) = \vartheta(t-t')e^{i\varpi(t+t')}\langle 0|\left(\hat{a}_0^\dagger\right)^2|0\rangle + \vartheta(t'-t)e^{i\varpi(t+t')}\langle 0|\left(\hat{a}_0^\dagger\right)^2|0\rangle = 0. \tag{3.3.109}$$

Only the mixed $\hat{a}$, $\hat{a}^\dagger$ vev is nonzero. Due to time translational symmetry, it is only a function of $t-t'$ and so we can define the function $f(t)$:

$$f(t) = e^{-i\varpi t}\vartheta(t), \qquad \texttt{Vev}(\hat{a}(t)\hat{a}^\dagger(t')) = f(t-t'). \tag{3.3.110}$$

When doing real calculations, the Fourier transform of $f(t)$ tends to be more useful. Remember the $\varepsilon$ damping always used in the definition of the Fourier transform as it is important here:

$$\int_{-\infty}^{\infty} e^{i\omega t}f(t)dt = \lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty} e^{-\varepsilon|t|-i\varpi t}\vartheta(t)dt = \lim_{\varepsilon\to 0^+}\frac{i}{\omega-\varpi+i\varepsilon} \tag{3.3.111}$$

$$\implies f(t) = \lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\frac{i}{\omega-\varpi+i\varepsilon}e^{-i\omega t}\frac{d\omega}{2\pi}. \tag{3.3.112}$$

We now compute the vevs of our coordinates and momenta in terms of this representation of $f(t)$:

$$\texttt{Vev}(\hat{\mathfrak{q}}(t)\hat{\mathfrak{q}}(t')) = \frac{\hbar}{2\varpi}\texttt{Vev}(\hat{a}(t)\hat{a}(t') + \hat{a}^\dagger(t)\hat{a}(t') + \hat{a}(t)\hat{a}^\dagger(t') + \hat{a}^\dagger(t)\hat{a}^\dagger(t')) \tag{3.3.113}$$

$$= \frac{\hbar}{2\varpi}\lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\left(\frac{i}{-\omega-\varpi+i\varepsilon} + \frac{i}{\omega-\varpi+i\varepsilon}\right)e^{-i\omega(t-t')}\frac{d\omega}{2\pi} \tag{3.3.114}$$

$$= \lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\frac{i\hbar\left(1-i\frac{\varepsilon}{\varpi}\right)}{\omega^2-\varpi^2+2i\varpi\varepsilon+\varepsilon^2}e^{-i\omega(t-t')}\frac{d\omega}{2\pi}. \tag{3.3.115}$$

The $\varepsilon$ piece in the numerator makes a vanishing contribution to the integral in to $\varepsilon\to 0^+$ limit, as does the $\varepsilon^2$ piece in the denominator. The only $\varepsilon$ dependent piece that makes a nonzero contribution to the result is the $+2i\varpi\varepsilon$ and so it is the only one necessary to keep (this can be checked by simply carrying out the integration). As well, because we only require $\varepsilon\to 0^+$ we can rescale the definition of $\varepsilon$ to absorb the factor of $2\varpi$. So:

$$\boxed{\texttt{Vev}(\hat{\mathfrak{q}}(t)\hat{\mathfrak{q}}(t')) = \lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\frac{i\hbar}{\omega^2-\varpi^2+i\varepsilon}e^{-i\omega(t-t')}\frac{d\omega}{2\pi}.} \tag{3.3.116}$$

The appearance of $+i\varepsilon$ in the denominator here is often called the Feynman $i\varepsilon$ or Feynman prescription. The vevs of other combinations of momenta and coordinates are simply:

$$\texttt{Vev}(\hat{\mathfrak{q}}(t)\hat{\mathfrak{p}}(t')) = \frac{\partial}{\partial t'}\texttt{Vev}(\hat{\mathfrak{q}}(t)\hat{\mathfrak{q}}(t')), \qquad \texttt{Vev}(\hat{\mathfrak{p}}(t)\hat{\mathfrak{p}}(t')) = \frac{\partial^2}{\partial t\partial t'}\texttt{Vev}(\hat{\mathfrak{q}}(t)\hat{\mathfrak{q}}(t')). \tag{3.3.117}$$

Now let $k$ and $l$ be indices which run over $1,2$ with 1 indicating a coordinate $\mathfrak{q}$ and 2 indicating a momentum $\mathfrak{p}$ and let $n$ run over the different pairs of normal modes. So, we write ${\Lambda^a}_{n,k}$ to mean that $a$ is a component of our original $Q,P$ coordinates and $n,k$ specifies a particular coordinate or momentum of our normal mode coordinates. With this notation, we have finally that the Feynman propagator of a general system of interacting quantum harmonic oscillators is given by (with implicit summation over $k,l$ and explicit summation over $n$):

$$\boxed{D^{ab}(t,t') = \sum_n {\Lambda^a}_{n,k}{\Lambda^b}_{n,l}\begin{pmatrix}1 & \frac{\partial}{\partial t'}\\ \frac{\partial}{\partial t} & \frac{\partial^2}{\partial t\partial t'}\end{pmatrix}^{kl}\lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\frac{i\hbar}{\omega^2-\omega_n^2+i\varepsilon}e^{-i\omega(t-t')}\frac{d\omega}{2\pi}.} \tag{3.3.118}$$

In the path integral perspective, the $+i\varepsilon$ in the denominator appears the way it does because of the boundary conditions on the path integral (fixing the initial and final coordinates and applying the time Wick rotation). If you use different boundary conditions for computing which $\mathcal{G}^{ab}$ satisfies (3.3.90), you will find different appearances of $i\varepsilon$. The most common useful alternative boundary conditions in some calculations

are delayed boundary conditions (also called causal boundary conditions) in which the initial coordinates and initial momenta are specified. These boundary conditions lead to the familiar causal Green's function for the harmonic oscillator. Importantly, the use of the path integral (and equivalently the fact that vevs are time ordered) imposes the Feynman $+i\varepsilon$ upon us; it is not a choice. Other boundary conditions are only ever useful when computing something other than a time ordered vev.

### 3.3.5 Fermionic Path Integral

## Exercises

**Problem 3.1: Be careful, even with quadratic path integrals.** The ability to put $U(t_f, Q_f, t_0, Q_0)$ into the form of (3.3.36) depends fully on the special form of the Hamiltonian Wigner function (3.3.29). It has often been tempting to others to attempt to use (3.3.36) even when $W^{ab}$ in (3.3.29) is $Q$ dependent. Here we will see how this goes wrong.

(a) Show that if $W^{ab} = W^{ab}(t, Q)$ then by following the steps leading to (3.3.32) we have:

$$U(t_f, Q_f, t_0, Q_0) = \lim_{N\to\infty} \int e^{\frac{i}{\hbar}\sum_{n=0}^{N-1} \delta S_{\mathcal{L}}} \prod_{n=0}^{N-1} \left( \frac{1}{(2\pi i\hbar\delta t_n)^{\frac{d}{2}} \sqrt{\det W(t_n, \bar{Q}_n)}} \right) d^dQ_1 ... d^dQ_{N-1}.$$

From this show that the $\sqrt{\det W}$ factors now no longer cancel in calculations of operator summaries such as in (3.3.40); they matter!

(b) Show formally (because the notation is plainly only schematic) that the above $U(t_f, Q_f, t_0, Q_0)$ can be put into the form:

$$U(t_f, Q_f, t_0, Q_0) = \mathcal{N}(t_f, t_0) \int e^{\frac{i}{\hbar}(S_{\mathcal{L}} + \frac{i\hbar}{2}\int_{t_0}^{t_f} \delta(0) \ln\det W(t,Q) dt)} \mathcal{D}Q$$

for some function $\mathcal{N}(t_f, t_0)$ and where $\delta(0)$ is understood as the $N \to \infty$ limit of $\frac{1}{\delta t_n}$. Thus the effective action functional $S'$ receives a correction which is neither real nor finite. Again, the details of (3.3.29) matter for justifying (3.3.36).

**Problem 3.2: Leading classical behavior of the Propagator.** The path integral provides an easily angle to see how the propagator behaves near the classical limit.

(a) Argue that in the $\hbar \to 0^+$ limit, the phase of the path integral oscillates wildly in any region of integration in which the exponent is not stationary. Further, argue that this rapid oscillation causes these regions to have vanishing contribution to the result. Thus, explain that we expect that in that limit $U(t_f, Q_f, t_0, Q_0)$ should contain a prefactor $e^{\frac{i}{\hbar}S_{\text{cl}}(t_f, Q_f, t_0, Q_0)}$ where $S_{\text{cl}}(t_f, Q_f, t_0, Q_0)$ is the classical on-shell action with boundary conditions $Q(t_f) = Q_f$ and $Q(t_0) = Q_0$. The integral measure of the path integral contains many powers of $\hbar$ (formally infinite), and it is not clear a priori how many of these factors will survive in the classical limit. Thus, we also may reasonably expect that $U$ carries some overall power of $\hbar$. That is, we expect that for some choice of $\alpha$ the function:

$$M(t_f, Q_f, t_0, Q_0) = \lim_{\hbar\to 0^+} \hbar^\alpha e^{-\frac{i}{\hbar}S_{\text{cl}}(t_f, Q_f, t_0, Q_0)} U(t_f, Q_f, t_0, Q_0)$$

should be well-defined and finitely nonzero. In this problem we will show that these expectations are correct, determine $\alpha$, and compute $M$.

(b) Let $H$ have $\hbar$ expansion:

$$H(t, Q, P) = \mathcal{H}(t, Q, P) + \hbar H_1(t, Q, P) + \mathcal{O}(\hbar^2)$$

and let $\bar{Q}(t)$ and $\bar{P}(t)$ be the classical trajectories which make the classical action stationary. Define $\delta Q(t)$ and $\delta P(t)$ by:

$$Q(t) = \bar{Q}(t) + \sqrt{\hbar}\delta Q(t), \qquad P(t) = \bar{P}(t) + \sqrt{\hbar}\delta P(t).$$

Expand the Wigner action functional in $\hbar$ to show that:

$$S[Q, P] = S_{\text{cl}}(t_f, Q_f, t_0, Q_0) + \hbar \int_{t_0}^{t_f} \left( \delta P_a \delta\dot{Q}^a - \widetilde{\mathcal{H}}(t, \delta Q, \delta P) + H_1(t, \bar{Q}(t), \bar{P}(t)) \right) dt + \mathcal{O}(\hbar^{\frac{3}{2}})$$

where we have defined the function:

$$\widetilde{\mathcal{H}}(t, \delta Q, \delta P) = \frac{1}{2}\frac{\partial^2\mathcal{H}}{\partial\bar{Q}^a\partial\bar{Q}^b}\delta Q^a\delta Q^b + \frac{\partial^2\mathcal{H}}{\partial\bar{Q}^a\bar{P}_b}\delta Q^a\delta P_b + \frac{1}{2}\frac{\partial^2\mathcal{H}}{\partial\bar{P}_a\partial\bar{P}_b}\delta P_a\delta P_b$$

and everywhere $\mathcal{H}$ and its derivatives appear in this definition, they are evaluated at $(t, \bar{Q}(t), \bar{P}(t))$.

(c) Return the above to the path integral to show that:

$$U = e^{\frac{i}{\hbar}S_{\text{cl}} - i\int_{t_0}^{t_f} H_1 dt} \lim_{N\to\infty} \int e^{i\sum_{n=0}^{N-1}\left((\delta Q^a_{n+1} - \delta Q^a_n)\delta P_{na} - \frac{t_f - t_0}{N}\widetilde{\mathcal{H}}\left(t_0 + \frac{t_f-t_0}{N}, \frac{\delta Q_{n+1}+\delta Q_n}{2}, \delta P_n\right)\right) + \mathcal{O}(\sqrt{\hbar})} \cdot \frac{d^d\delta P_0}{(4\pi^2\hbar)^{\frac{d}{2}}} \frac{d^d\delta Q_1 d^d \delta P_1}{(2\pi)^d} \cdots \frac{d^d \delta Q_{N-1} d^d \delta P_{N-1}}{(2\pi)^d}.$$

Based on this, explain that $\alpha = \frac{d}{2}$ and show that:

$$Me^{i\int_{t_0}^{t_f} H_1 dt} = \lim_{N\to\infty} \int e^{i\sum_{n=0}^{N-1}\left((\delta Q^a_{n+1} - \delta Q^a_n)\delta P_{na} - \frac{t_f-t_0}{N}\widetilde{\mathcal{H}}\right)} \frac{d^d\delta P_0}{(2\pi)^d} \frac{d^d\delta Q_1 d^d\delta P_1}{(2\pi)^d} \cdots \frac{d^d\delta Q_{N-1} d^d \delta P_{N-1}}{(2\pi)^d}.$$

(d) Show that the right hand side of the above is exactly the path integral expression for the propagator of a quantum system with Wigner Hamiltonian $\widetilde{\mathcal{H}}$, phase space $(\delta Q, \delta P)$, for which $\hbar = 1$, and with boundary conditions $\delta Q(t_0) = \delta Q(t_f) = 0$. Let $\widetilde{S}(t_f, \delta Q_f, t_0, \delta Q_0)$ be the classical on-shell action corresponding to such a system and let $\widetilde{U}(t_f, \delta Q_f, t_0, \delta Q_0)$ be the quantum propagator associated to such a system. That is, show that:

$$Me^{i\int_{t_0}^{t_f} H_1 dt} = \widetilde{U}(t_f, 0, t_0, 0)$$

and because $\widetilde{\mathcal{H}}$ is a quadratic polynomial in $\delta Q, \delta P$, explain that exactly:

$$\widetilde{U}(t_f, \delta Q_f, t_0, \delta Q_0) = \left(\frac{-1}{2\pi i}\right)^{\frac{d}{2}} \sqrt{\det \frac{\partial^2 \widetilde{S}}{\partial \delta Q_f \partial \delta Q_0}} e^{i\widetilde{S}}$$

(e) Using the classical equations of motion for $\delta Q$ and $\delta P$, show that the on-shell for generic boundary conditions is given by:

$$\widetilde{S}(t_f, \delta Q_f, t_0, \delta Q_0) = \frac{1}{2}\left(\delta Q^a_f \delta P_{fa} - \delta Q^a_0 \delta P_{0a}\right)$$

and thus that:

$$\frac{\partial^2 \widetilde{S}}{\partial \delta Q_f \partial \delta Q_0}(t_f, 0, t_0, 0) = \frac{1}{2}\left(\frac{\partial \delta P_{fa}}{\partial \delta Q^b_0} - \frac{\partial \delta P_{0b}}{\partial \delta Q^a_f}\right).$$

By returning to the change of variables that defined $\delta Q$ and $\delta P$, show that at $\delta Q_f = \delta Q_0 = 0$:

$$\frac{\partial \delta P_{fa}}{\partial \delta Q^b_0} = -\frac{\partial \delta P_{0b}}{\partial \delta Q^a_f} = \frac{\partial^2 S_{\text{cl}}}{\partial Q^a_f \partial Q^b_0}.$$

(f) Collect all of the results of this problem to show that for a generic quantum system with a well-behaved classical limit, its propagator satisfies:

$$\lim_{\hbar\to 0^+} \hbar^{\frac{d}{2}} e^{-\frac{i}{\hbar}S_{\text{cl}}(t_f, Q_f, t_0, Q_0)} U(t_f, Q_f, t_0, Q_0) = \left(\frac{-1}{2\pi i}\right)^{\frac{d}{2}} \sqrt{\det \frac{\partial^2 S_{\text{cl}}}{\partial Q_f \partial Q_0}}.$$

As we had previously shown, this formula holds exactly for the quantum harmonic oscillator without taking the $\hbar \to 0^+$ limit on the left-hand side but we see it remains useful for a generic system near the classical limit.

**Problem 3.3: Time Domain Representation of the Feynman Propagator.** Complete the Fourier transform on the right hand side of (3.3.116) to show explicitly that:

$$\texttt{Vev}(\hat{\mathfrak{q}}(t)\hat{\mathfrak{q}}(t')) = \frac{\hbar}{2\varpi} e^{-i\varpi|t-t'|}.$$

You may be initially surprised that this vev is not real valued, given that $\hat{\mathfrak{q}}(t)$ is self-adjoint at each $t$. However, because $\hat{\mathfrak{q}}(t)$ and $\hat{\mathfrak{q}}(t')$ do not commute for $t \neq t'$, the product $\hat{\mathfrak{q}}(t)\hat{\mathfrak{q}}(t')$ is not self-adjoint and so the vev is fully allowed to be complex valued.

## 3.4 Scalar Theory - Klein-Gordon Equation

Now that we have an idea of what is going on in Lorentz field theory, let's start writing some down. The most basic possible Lorentz transformation property we could have for a field is when $s_l = s_r = 0$ and so by Theorem 1.11 the field $\phi$ doesn't transform at all under Lorentz transformation:

$$\phi'(X') = \phi(X). \tag{3.4.1}$$

Then, the field is said to be a scalar field. Suppose we are looking for a theory of a single real scalar field which is interacting with no other fields. Then, expanding the Lagrangian density for weak $\phi$ the first few terms in the field we can add without derivative are linear and quadratic in $\phi$. The derivatives $\partial_\mu\phi$ make a 4-vector given that $\phi$ is a scalar and so to add a deriative term to the Lagrangian density we need to contract it with another 4-vector. However, because there are no other fields at this time the lowest order possibility is to contract it with itself as $\partial_\mu\phi\partial^\mu\phi$. Thus, the lowest order possible Lagrangian density is:

$$\mathfrak{L} = -\frac{1}{2}\sigma\partial_\mu\phi\partial^\mu\phi - \alpha\phi - \frac{1}{2}\beta\phi^2 \tag{3.4.2}$$

for some constants $\sigma, \alpha, \beta$. As usual, higher order terms containing higher powers of $\phi$ may become relevant for strong fields and higher derivative terms may appear on small scales but at lowest order this is it. We can redefine $\phi$ to make the theory nicer. First, we can shift the definition of $\phi$ so that the linear term vanishes:

$$\phi' = \phi + c \implies \mathfrak{L} = -\frac{1}{2}\sigma\partial_\mu\phi'\partial^\mu\phi' - \alpha(\phi' - c) - \frac{1}{2}\beta(\phi' - c)^2 \tag{3.4.3}$$

$$\implies \mathfrak{L} = -\frac{1}{2}\sigma\partial_\mu\phi'\partial^\mu\phi' - (\alpha + \beta c)\phi' - \frac{1}{2}\beta\phi'^2 - \frac{1}{2}\beta c^2 + \alpha c. \tag{3.4.4}$$

The constant terms we can drop as they are just surface terms. If we choose $c = -\alpha/\beta$ then the linear term drops out and so with this new definition of $\phi$ (meaning we no longer write the primes) we have:

$$\mathfrak{L} = -\frac{1}{2}\sigma\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\beta\phi^2. \tag{3.4.5}$$

We can also rescale our definition of $\phi$:

$$\phi' = \frac{1}{Z}\phi \implies \mathfrak{L} = -\frac{1}{2}\sigma Z^2\partial_\mu\phi'\partial^\mu\phi' - \frac{1}{2}\beta Z^2\phi'^2. \tag{3.4.6}$$

This changes the values of $\sigma$ and $\beta$ to $\sigma' = \sigma Z^2$ and $\beta' = \beta Z^2$. So, if we choose $Z = \frac{1}{\sqrt{|\sigma|}}$ then we can reduce $\sigma'$ to being either $+1$ or $-1$. For convenience, also define $b = \sigma'\beta'$. Then, with this new definition of $\phi, \sigma, b$ (meaning we no longer write primes):

$$\mathfrak{L} = -\frac{1}{2}\sigma\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\sigma b\phi^2 = \frac{1}{2}\sigma\dot{\phi}^2 - \frac{1}{2}\sigma\partial_a\phi\partial^a\phi - \frac{1}{2}b\phi^2. \tag{3.4.7}$$

with $\sigma = \pm 1$. From this we can compute the Hamiltonian:

$$\varphi = \frac{\delta\mathcal{L}}{\delta\dot{\phi}} = \sigma\dot{\phi} \implies H = \varphi\dot{\phi} - \mathfrak{L} = \frac{1}{2}\sigma(\varphi^2 + \partial_a\phi\partial^a\phi + b\phi^2) \tag{3.4.8}$$

$$\implies \mathcal{H} = \frac{\sigma}{2}\int_{\mathbb{S}}(\varphi^2 + \partial_a\phi\partial^a\phi + b\phi^2)d\mathbb{S}. \tag{3.4.9}$$

Because this Hamiltonian is quadratic we can put it into into the Schrödinger form with the right coordinate transformation. To do this we need to rewrite this Hamiltonian in the form $\langle\psi|\hat{H}|\psi\rangle$. It is easiest to do this by first taking the Fourier transform of the fields above:

$$\widetilde{\phi}(\vec{k}) = \int_{\mathbb{S}}\frac{e^{-i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}}\phi(\vec{x})d\mathbb{S} \implies \phi(\vec{x}) = \int_{\mathbb{K}}\frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}}\widetilde{\phi}(\vec{k})d\mathbb{K} \tag{3.4.10}$$

$$\widetilde{\varphi}(\vec{k}) = \int_{\mathbb{S}}\frac{e^{-i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}}\varphi(\vec{x})d\mathbb{S} \implies \varphi(\vec{x}) = \int_{\mathbb{K}}\frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}}\widetilde{\varphi}(\vec{k})d\mathbb{K}. \tag{3.4.11}$$

The Fourier transform of a real valued function satisfies:

$$\widetilde{\phi}^*(\vec{k}) = \int_{\mathbb{S}}\frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}}\phi(\vec{x})d\mathbb{S} = \int_{\mathbb{S}}\frac{e^{-i(-\vec{k})\cdot\vec{x}}}{(2\pi)^{3/2}}\phi(\vec{x})d\mathbb{S} = \widetilde{\phi}(-\vec{k}). \tag{3.4.12}$$

With this, we can rewrite our Hamiltonian as:

$$\mathcal{H} = \frac{\sigma}{2}\int_{\mathbb{SKK}'}\left(\frac{e^{-i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}}\widetilde{\varphi}^*(\vec{k})\frac{e^{i\vec{k}'\cdot\vec{x}}}{(2\pi)^{3/2}}\widetilde{\varphi}(\vec{k}') + \partial_a\frac{e^{-i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}}\widetilde{\phi}^*(\vec{k})\partial^a\frac{e^{i\vec{k}'\cdot\vec{x}}}{(2\pi)^{3/2}}\widetilde{\phi}(\vec{k}') + b\frac{e^{-i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}}\widetilde{\phi}^*(\vec{k})\frac{e^{i\vec{k}'\cdot\vec{x}}}{(2\pi)^{3/2}}\widetilde{\phi}(\vec{k}')\right)d\mathbb{S}d\mathbb{K}d\mathbb{K}' \tag{3.4.13}$$

$$= \frac{\sigma}{2}\int_{\mathbb{SKK}'}\frac{e^{i(\vec{k}'-\vec{k})\cdot\vec{x}}}{(2\pi)^3}\left(\widetilde{\varphi}^*(\vec{k})\widetilde{\varphi}(\vec{k}') + (-ik_a)\widetilde{\phi}^*(\vec{k})(ik'^a)\widetilde{\phi}(\vec{k}') + b\widetilde{\phi}^*(\vec{k})\widetilde{\phi}(\vec{k}')\right)d\mathbb{S}d\mathbb{K}d\mathbb{K}' \tag{3.4.14}$$

$$= \frac{\sigma}{2}\int_{\mathbb{SKK}'}\delta(\vec{k}-\vec{k}')\left(\widetilde{\varphi}^*(\vec{k})\widetilde{\varphi}(\vec{k}') + (-ik_a)\widetilde{\phi}^*(\vec{k})(ik'^a)\widetilde{\phi}(\vec{k}') + b\widetilde{\phi}^*(\vec{k})\widetilde{\phi}(\vec{k}')\right)d\mathbb{K}d\mathbb{K}' \tag{3.4.15}$$

$$= \frac{\sigma}{2}\int_{\mathbb{SKK}'}\left(\widetilde{\varphi}^*(\vec{k})\widetilde{\varphi}(\vec{k}) + (|\vec{k}|^2+b)\widetilde{\phi}^*(\vec{k})\widetilde{\phi}(\vec{k})\right)d\mathbb{K}. \tag{3.4.16}$$

Now we are almost in the form of the quantum Hamiltonian, we just have one final step to make according to our procedure from before. Define $\omega_k$ and $\widetilde{\psi}(t,\vec{k})$ as:

$$\omega_k = \sqrt{|\vec{k}|^2+b} \tag{3.4.17}$$

$$\widetilde{\psi}(\vec{k}) = \sqrt{\frac{\omega_k}{2}}\widetilde{\phi}(\vec{k}) + \frac{i}{\sqrt{2\omega_k}}\widetilde{\varphi}(\vec{k}), \qquad \widetilde{\pi}(\vec{k}) = i\widetilde{\psi}^*(\vec{k}) \tag{3.4.18}$$

$$\implies \widetilde{\psi}^*(\vec{k}) = \sqrt{\frac{\omega_k}{2}}\widetilde{\phi}^*(\vec{k}) - \frac{i}{\sqrt{2\omega_k}}\widetilde{\varphi}^*(\vec{k}) = \sqrt{\frac{\omega_k}{2}}\widetilde{\phi}(-\vec{k}) - \frac{i}{\sqrt{2\omega_k}}\widetilde{\varphi}(-\vec{k}) \tag{3.4.19}$$

$$\implies \widetilde{\phi}(\vec{k}) = \frac{\widetilde{\psi}(\vec{k})+\widetilde{\psi}^*(-\vec{k})}{\sqrt{2\omega_k}}, \qquad \widetilde{\varphi}(\vec{k}) = \frac{1}{i}\sqrt{\frac{\omega_k}{2}}(\widetilde{\psi}(\vec{k}) - \widetilde{\psi}^*(-\vec{k})). \tag{3.4.20}$$

Now putting the Hamiltonian in terms of $\psi$:

$$\mathcal{H} = \frac{\sigma}{2}\int_{\mathbb{K}}\left(\frac{-1}{i}\sqrt{\frac{\omega_k}{2}}(\widetilde{\psi}^*(\vec{k}) - \widetilde{\psi}(-\vec{k}))\frac{1}{i}\sqrt{\frac{\omega_k}{2}}(\widetilde{\psi}(\vec{k}) - \widetilde{\psi}^*(-\vec{k})) + \omega_k^2\frac{\widetilde{\psi}(\vec{k})+\widetilde{\psi}^*(-\vec{k})}{\sqrt{2\omega_k}}\frac{\widetilde{\psi}^*(\vec{k})+\widetilde{\psi}(-\vec{k})}{\sqrt{2\omega_k}}\right)d\mathbb{K} \tag{3.4.21}$$

$$= \frac{\sigma}{2}\int_{\mathbb{K}}\frac{\omega_k}{2}\Big(\widetilde{\psi}^*(\vec{k})\widetilde{\psi}(\vec{k}) - \widetilde{\psi}(-\vec{k})\widetilde{\psi}(\vec{k}) - \widetilde{\psi}^*(-\vec{k})\widetilde{\psi}(\vec{k}) + \widetilde{\psi}^*(-\vec{k})\widetilde{\psi}(-\vec{k}) + \widetilde{\psi}^*(\vec{k})\widetilde{\psi}(\vec{k}) + \widetilde{\psi}(-\vec{k})\widetilde{\psi}(\vec{k}) + \widetilde{\psi}^*(-\vec{k})\widetilde{\psi}(\vec{k}) + \widetilde{\psi}^*(-\vec{k})\widetilde{\psi}(-\vec{k})\Big)d\mathbb{K} \tag{3.4.22}$$

$$= \frac{\sigma}{2}\int_{\mathbb{K}}\omega_k(\widetilde{\psi}^*(\vec{k})\widetilde{\psi}(\vec{k}) + \widetilde{\psi}^*(-\vec{k})\widetilde{\psi}(-\vec{k}))d\mathbb{K} \tag{3.4.23}$$

$$= \sigma\int_{\mathbb{K}}\widetilde{\psi}^*(\vec{k})(\sqrt{|\vec{k}|^2+b})\widetilde{\psi}(\vec{k})d\mathbb{K}. \tag{3.4.24}$$

Now by rewriting using the Dirac notation we arrive at the Schrödinger form of this Hamiltonian:

$$\mathcal{H} = \sigma\int_{\mathbb{K}}\langle\psi|\vec{k}\rangle(\sqrt{|\vec{k}|^2+b}\langle\vec{k}|\psi\rangle d\mathbb{K} = \sigma\langle\psi|\sqrt{|\hat{\vec{p}}|^2+b}|\psi\rangle. \tag{3.4.25}$$

We can now see the meaning of the parameters $\sigma$ and $b$ by considering the non-relativistic (low speed) limit of this Hamiltonian. Because the Hamiltonian operator only depends on $\hat{\vec{p}}$, its eigenstates are simply the momentum eigenstates:

$$\hat{H}|\vec{p}\rangle = \sqrt{|\vec{p}|^2+b}|\vec{p}\rangle. \tag{3.4.26}$$

As well, the velocity operator is given by:

$$\hat{\vec{v}} = \frac{1}{i}[\hat{\vec{x}},\hat{H}] = \frac{\partial\hat{H}}{\partial\hat{\vec{p}}} = \frac{\sigma\hat{\vec{p}}}{\sqrt{|\hat{\vec{p}}|^2+b}} \tag{3.4.27}$$

and so the velocity eigenstates are also the momentum eigenstates:

$$\hat{\vec{v}}|\vec{p}\rangle = \frac{\sigma\vec{p}}{\sqrt{|\vec{p}|^2+b}}|\vec{p}\rangle. \tag{3.4.28}$$

If $b<0$ then there exists a $\vec{p}$ at which the denominator diverges and there does not exists a $\vec{p}$ at which the eigenvalue on the right goes to 0. Thus, in order for the low speed limit of this theory to even exist,

$b \geq 0$. Then, with $b > 0$, the velocity eigenvalue is 0 if and only if $\vec{p} = 0$ and thus the low speed expansion is equivalent to a low $|\vec{p}|$ expansion. Performing this expansion:

$$\hat{\vec{v}}|\vec{p}\rangle = \left(\frac{\sigma\vec{p}}{\sqrt{b}} + \text{H.O.T.}\right)|\vec{p}\rangle. \tag{3.4.29}$$

For a Galilean free particle, we had:

$$\hat{H} = \frac{|\hat{\vec{p}}|^2}{2m} \implies \hat{\vec{v}} = \frac{1}{i}[\hat{\vec{x}}, \hat{H}] = \frac{\hat{\vec{p}}}{m} \tag{3.4.30}$$

$$\implies \hat{\vec{v}}|\vec{p}\rangle = \frac{\vec{p}}{m}|\vec{p}\rangle. \tag{3.4.31}$$

Thus, matching these two results we must have $\sigma = 1$ and $b = m^2$. Putting all of this back into our Lagrangian density, we have that the lowest order Lagrangian density for scalar matter is:

$$\mathfrak{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\phi^2 \tag{3.4.32}$$

where $m$ is the mass of the field (meaning in the classical limit it becomes the mass of the corresponding particle). This is called the Klein-Gordon Lagrangian density (LD). Computing the equation of motion we get the Klein-Gordon equation of motion:

$$\frac{\partial\mathfrak{L}}{\partial\phi} - \partial_\mu\frac{\partial\mathfrak{L}}{\partial\partial_\mu\phi} = -m^2\phi + \partial_\mu\partial^\mu\phi = 0. \tag{3.4.33}$$

This combination $\partial_\mu\partial^\mu$ is sometimes abbreviated by $\partial^2$ giving:

$$(\partial^2 - m^2)\phi = 0. \tag{3.4.34}$$

Because this $\phi$ is only a single real scalar field we can't make any U(1) transformation on it the way we did for the wavefunction in nonrelativistic quantum mechanics. So, we can't couple this field to the vector potential in the way we did before. This may be surprising as you may expect that because this $\widetilde{\psi}(t, \vec{k})$ is exactly in the Hamiltonian in the way the wavefunction was previously that we could simply Fourier transform it to a spatial wavefunction $\psi(t, \vec{x})$ and apply U(1) transformations to it, just as before. However, while $\widetilde{\psi}$ and $\widetilde{\phi}$ are nicely related in Fourier space, the way they are related in real space is in fact non-local. The non-locality enters when one tries to take the Fourier transform involving the necessary factors of $\sqrt{\omega_k}$ appearing in (3.4.20). Due to this non-locality, $\psi(t, \vec{x})$ will not transform in any simple way at all under Lorentz transformation. It's Lorentz transformation $\psi'(t', \vec{x}')$ ends up depending on $\psi(t, \vec{x})$ for all $t$ and $\vec{x}$. Consequently, a local U(1) transformation of $\psi$ does not correspond to a local transformation at all of our scalar field $\phi$. It turns out that in the special case of $m = 0$ there is a different way to make $\phi$ transform under the $U(1)$ transformation, as covered in the homework. For the general case of $m \neq 0$, a real local scalar field cannot transform at all under the U(1) gauge transformation. Another transparent way to see that the behavior of $\psi$ is non-local is to look at its equation of motion. From (3.4.24):

$$\mathcal{H} = \langle\psi|\sqrt{|\hat{\vec{p}}|^2 + m^2}|\psi\rangle \tag{3.4.35}$$

$$\implies i\frac{\partial\psi}{\partial t} = \sqrt{m^2 - \nabla^2}\psi. \tag{3.4.36}$$

Here the meaning of $\sqrt{m^2 - \nabla^2}$ is given either by the series expansion of the $\sqrt{m^2 + x}$ around $x = 0$ and replacing $x$ by $\nabla^2$ or equivalently by defining it by its eigenvalue in the Fourier domain and then taking the inverse Fourier transform afterward. Using the Fourier domain expression, one can see that this operator is in fact non-local as it causes the time-derivative of $\psi(t, \vec{x})$ to depend on $\psi(t, \vec{x}')$ for all $\vec{x}'$. To be clear, it is possible to make interaction terms between $\phi$ and the electromagnetic field, for example we could add a term to the Lagrangian density:

$$\mathfrak{L}_{\text{int}} = g\phi\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu}. \tag{3.4.37}$$

This certainly has all of the necessary symmetries and thus is a valid term to consider in the Lagrangian. This term is automatically cubic (because $\mathcal{F}_{\mu\nu}$ is linear in the potentials so we have two powers of the potentials and one power of $\phi$) and so is "higher order" than our free equations of motion. But, it is absolutely allowed. A fact one must reconcile with in field theory is that actually the total number of possible terms is enormous once you have more than one field.

While a single real massive scalar field can't couple directly to the potentials, two such fields can. Consider now a theory of two real massive scalar fields. By repeating our previous argument from the case of

a single such field, our general quadratic order Lagrangian density is simply (by writing all Lorentz invariant quadratic terms with one or fewer derivatives on each field):

$$\mathfrak{L} = -\frac{a}{2}\partial_\mu\phi_1\partial^\mu\phi_1 - b\partial_\mu\phi_1\partial^\mu\phi_2 - \frac{c}{2}\partial_\mu\phi_2\partial^\mu\phi_2 - d\phi_1 - e\phi_2 - \frac{f}{2}\phi_1^2 - g\phi_1\phi_2 - \frac{h}{2}\phi_2^2. \tag{3.4.38}$$

All of these terms are okay in principle. However, if we want to couple our fields to the potentials in the same way as we did in Chapter 1, then we will think about these two fields as the components of a field $\Phi(t,\vec{x})$ defined as an element of the internal space at $(t,\vec{x})$, $\mathbb{V}(t,\vec{x})$. Then, the U(1) transformation is as we found:

$$\begin{pmatrix}\phi_1'\\ \phi_2'\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\phi_1\\ \phi_2\end{pmatrix} \implies \begin{pmatrix}\mathscr{D}\phi_1\\ \mathscr{D}\phi_2\end{pmatrix} = \begin{pmatrix}-\phi_2\\ \phi_1\end{pmatrix}. \tag{3.4.39}$$

To have local symmetry under this transformation (meaning $\theta = \theta(t,\vec{x})$) we will again need the potentials to act as a connection. However, we can have global symmetry (meaning $\theta$ does not depend on $t,\vec{x}$) without adding in the potentials. We can do so because the potentials change under gauge transformation by the derivatives of $\theta$ and so if the potentials are not present (and thus effectively 0) and we apply a constant uniform $\theta$ then they will still not be present (as they will still be 0) afterward. Then, gauge symmetry will mean that we need symmetry under this global transformation and thus require:

$$\mathfrak{L}' = \mathfrak{L} + \partial_\mu f^\mu(\theta, x, \phi_1, \phi_2) \tag{3.4.40}$$

$$\implies \left.\frac{d\mathfrak{L}'}{d\theta}\right|_{\theta=0} = \partial_\mu \mathscr{D} f(x,\phi_1,\phi_2) \tag{3.4.41}$$

$$\implies \frac{\partial\mathfrak{L}}{\partial\phi_1}\mathscr{D}\phi_1 + \frac{\partial\mathfrak{L}}{\partial\phi_2}\mathscr{D}\phi_2 + \frac{\partial\mathfrak{L}}{\partial\partial_\mu\phi_1}\mathscr{D}\partial_\mu\phi_1 + \frac{\partial\mathfrak{L}}{\partial_\mu\phi_2}\mathscr{D}\partial_\mu\phi_2 = \eth_\mu\mathscr{D}f^\mu + \frac{\partial\mathscr{D}f^\mu}{\partial\phi_1}\partial_\mu\phi_1 + \frac{\partial\mathscr{D}f^\mu}{\partial\phi_2}\partial_\mu\phi_2 \tag{3.4.42}$$

$$\implies (d + f\phi_1 + g\phi_2)\phi_2 + (-e - g\phi_1 - h\phi_2)\phi_1 + (a\partial^\mu\phi_1 + b\partial^\mu\phi_2)\partial_\mu\phi_2 + (-b\partial^\mu\phi_1 - c\partial^\mu\phi_2)\partial_\mu\phi_1$$
$$= \eth_\mu\mathscr{D}f^\mu + \frac{\partial\mathscr{D}f^\mu}{\partial\phi_1}\partial_\mu\phi_1 + \frac{\partial\mathscr{D}f^\mu}{\partial\phi_2}\partial_\mu\phi_2 \tag{3.4.43}$$

$$\implies d\phi_2 - e\phi_1 - g\phi_1^2 + (f-h)\phi_1\phi_2 + g\phi_2^2 - b\partial^\mu\phi_1\partial_\mu\phi_1 + (a-c)\partial^\mu\phi_1\partial_\mu\phi_2 + b\partial^\mu\phi_2\partial_\mu\phi_2$$
$$= \eth_\mu\mathscr{D}f^\mu + \frac{\partial\mathscr{D}f^\mu}{\partial\phi_1}\partial_\mu\phi_1 + \frac{\partial\mathscr{D}f^\mu}{\partial\phi_2}\partial_\mu\phi_2. \tag{3.4.44}$$

Matching the terms with two factors of derivatives, we see:

$$-b\partial^\mu\phi_1\partial_\mu\phi_1 + (a-c)\partial^\mu\phi_1\partial_\mu\phi_2 + b\partial^\mu\phi_2\partial_\mu\phi_2 = 0 \implies a = c, \qquad b = 0. \tag{3.4.45}$$

Matching the terms with one factor of the derivative, we see:

$$\frac{\partial\mathscr{D}f^\mu}{\partial\phi_1}\partial_\mu\phi_1 + \frac{\partial\mathscr{D}f^\mu}{\partial\phi_2}\partial_\mu\phi_2 = 0 \implies \frac{\partial\mathscr{D}f^\mu}{\partial\phi_1} = \frac{\mathscr{D}f^\mu}{\partial\phi_2} = 0. \tag{3.4.46}$$

Finally matching the terms with zero factors of the derivative, we see:

$$d\phi_2 - e\phi_1 - g\phi_1^2 + (f-h)\phi_1\phi_2 + g\phi_2^2 = \frac{\partial f^\mu(x)}{\partial x^\mu} \implies f = h, \qquad d = e = g = 0. \tag{3.4.47}$$

Then, the Lagrangian density becomes:

$$\mathfrak{L} = -\frac{a}{2}\partial_\mu\phi_1\partial^\mu\phi_1 - \frac{f}{2}\phi_1^2 - \frac{a}{2}\partial_\mu\phi_2\partial^\mu\phi_2 - \frac{f}{2}\phi_2^2. \tag{3.4.48}$$

Then, we can see that the theory becomes two copies of a real scalar field. So, again let $\sigma = \frac{a}{|a|}$ and $\beta = \frac{\sigma f}{a^2}$. Then, defining the new fields:

$$\phi_{1r} = \sqrt{|a|}\phi_1, \qquad \phi_{2r} = \sqrt{|a|}\phi_2. \tag{3.4.49}$$

These new fields satisfy the same U(1) transformation:

$$\begin{pmatrix}\phi_{1r}'\\ \phi_{2r}'\end{pmatrix} = \begin{pmatrix}\sqrt{|a|}\phi_1'\\ \sqrt{|a|}\phi_2'\end{pmatrix} = \sqrt{|a|}\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\phi_1\\ \phi_2\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\phi_{1r}\\ \phi_{2r}\end{pmatrix}. \tag{3.4.50}$$

So, because the transformation properties are preserved we take $\phi_{1r}$ and $\phi_{2r}$ to be defined as our new fields and drop the r's. Then, our Lagrangian density in these new fields is:

$$\mathfrak{L} = -\frac{\sigma}{2}\partial_\mu\phi_1\partial^\mu\phi_1 - \frac{\sigma\beta}{2}\phi_1^2 - \frac{\sigma}{2}\partial_\mu\phi_2\partial^\mu\phi_2 - \frac{\sigma\beta}{2}\phi_2^2. \tag{3.4.51}$$

Then, in the classical limit as argued previously we must have $\sigma = 1$ and $\beta = m^2$ where $m$ is the mass of the

corresponding classical particle. Thus, the Lagrangian density becomes:

$$\mathfrak{L} = -\frac{1}{2}\partial_\mu\phi_1\partial^\mu\phi_1 - \frac{1}{2}\partial_\mu\phi_2\partial^\mu\phi_2 - \frac{m^2}{2}\phi_1^2 - \frac{m^2}{2}\phi_2^2. \tag{3.4.52}$$

This can be organized very nicely by defining the complex scalar field $\phi$ analogously to the complex wavefunction in nonrelativistic quantum mechanics:

$$\phi \stackrel{\text{def}}{=} \frac{\phi_1 + i\phi_2}{\sqrt{2}} \implies \phi_1 = \frac{\phi + \phi^*}{\sqrt{2}}, \qquad \phi_2 = \frac{\phi - \phi^*}{i\sqrt{2}}. \tag{3.4.53}$$

Written in terms of $\phi$, the Lagrangian density becomes:

$$\mathfrak{L} = -\frac{1}{2}\partial_\mu\left(\frac{\phi+\phi^*}{\sqrt{2}}\right)\partial^\mu\left(\frac{\phi+\phi^*}{\sqrt{2}}\right) - \frac{1}{2}\partial_\mu\left(\frac{\phi-\phi^*}{i\sqrt{2}}\right)\partial^\mu\left(\frac{\phi-\phi^*}{i\sqrt{2}}\right) - \frac{m^2}{2}\left(\frac{\phi+\phi^*}{\sqrt{2}}\right)^2 - \frac{m^2}{2}\left(\frac{\phi-\phi^*}{i\sqrt{2}}\right)^2 \tag{3.4.54}$$

$$= -\frac{1}{4}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\partial_\mu\phi^*\partial^\mu\phi - \frac{1}{4}\partial_\mu\phi^*\partial^\mu\phi^* + \frac{1}{4}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\partial_\mu\phi^*\partial^\mu\phi + \frac{1}{4}\partial_\mu\phi^*\partial^\mu\phi^*$$
$$- \frac{m^2}{4}\phi^2 - \frac{m^2}{2}\phi^*\phi - \frac{m^2}{4}(\phi^*)^2 + \frac{m^2}{4}\phi^2 - \frac{m^2}{2}\phi^*\phi + \frac{m^2}{2}(\phi^*)^2 \tag{3.4.55}$$

$$= -\partial^\mu\phi^*\partial_\mu\phi - m^2\phi^*\phi. \tag{3.4.56}$$

Then, the U(1) transformation becomes:

$$\phi' = \frac{\phi_1' + i\phi_2'}{\sqrt{2}} = \frac{\phi_1\cos\theta - \phi_2\sin\theta + i\phi_1\sin\theta + i\phi_2\cos\theta}{\sqrt{2}} = (\cos\theta + i\sin\theta)\frac{\phi_1 + i\phi_2}{\sqrt{2}} = e^{i\theta}\phi. \tag{3.4.57}$$

The complex scalar field again exactly satisfies the Klein-Gordon equation:

$$\frac{\partial\mathfrak{L}}{\partial\phi^*} - \partial_\mu\frac{\partial\mathfrak{L}}{\partial\partial_\mu\phi^*} = -m^2\phi + \partial_\mu\partial^\mu\phi = (\partial^2 - m^2)\phi = 0. \tag{3.4.58}$$

Now that we have global U(1) symmetry it is very straight forward to put in an interaction with the vector potential exactly as we did before by defining the local $U(1)$ gauge transformation by:

$$\phi'(X) = e^{iqf(X)}\phi(X), \qquad A'_\mu(X) = A_\mu(X) + \partial_\mu f(X). \tag{3.4.59}$$

Then, the gauge covariant derivative is given as:

$$\mathscr{D}_\mu\phi \stackrel{\text{def}}{=} (\partial_\mu - iqA_\mu)\phi \implies \mathscr{D}_\mu\phi' = (\partial_\mu - iqA_\mu - iq\partial_\mu f)(e^{iqf}\phi) = e^{iqf}\mathscr{D}_\mu\phi. \tag{3.4.60}$$

This lets us easily write down the gauge-invariant LD:

$$\mathfrak{L} = -(\mathscr{D}_\mu\phi)^*(\mathscr{D}^\mu\phi) - m^2|\phi|^2. \tag{3.4.61}$$

This is the minimal way in which we can couple the potentials to our field $\phi$ (minimal it the sense that we must do this due to gauge invariance; there are other terms we can add but are not forced to by gauge invariance). This theory has equation of motion:

$$\frac{\partial\mathfrak{L}}{\partial\phi^*} - \partial_\mu\frac{\partial\mathfrak{L}}{\partial\partial_\mu\phi^*} = -iqA_\mu\mathscr{D}^\mu\phi - m^2\phi + \partial_\mu(\mathscr{D}^\mu\phi) = \mathscr{D}_\mu\mathscr{D}^\mu\phi - m^2\phi = (\mathscr{D}^2 - m^2)\phi \stackrel{!}{=} 0 \tag{3.4.62}$$

$$\implies (\mathscr{D}^2 - m^2)\phi = 0. \tag{3.4.63}$$

Further this theory has the conserved current:

$$J^\mu = \frac{\partial\mathfrak{L}}{\partial A_\mu} = iq\phi(\mathscr{D}^\mu\phi)^* - iq\phi^*\mathscr{D}^\mu\phi \tag{3.4.64}$$

$$\partial_\mu J^\mu = iq\partial_\mu\phi(\mathscr{D}^\mu\phi)^* + iq\phi\partial_\mu\mathscr{D}^\mu\phi^* - iq\partial_\mu\phi^*\mathscr{D}^\mu\phi - iq\phi^*\partial_\mu\mathscr{D}^\mu\phi \tag{3.4.65}$$

$$= iq(\mathscr{D}_\mu\phi + iqA_\mu\phi)(\mathscr{D}^\mu\phi)^* + iq\phi(\mathscr{D}_\mu^* - iqA_\mu)(\mathscr{D}^\mu\phi)^* - iq(\mathscr{D}_\mu^* - iqA_\mu)\phi^*\mathscr{D}^\mu\phi - iq\phi^*(\mathscr{D}_\mu + iqA_\mu)\mathscr{D}^\mu\phi \tag{3.4.66}$$

$$= iq\mathscr{D}_\mu\phi(\mathscr{D}^\mu\phi)^* - q^2(\mathscr{D}^\mu\phi)^*A_\mu\phi + iq\phi(\mathscr{D}^2\phi)^* + q^2(\mathscr{D}^\mu\phi)^*A_\mu\phi \tag{3.4.67}$$
$$- iq(\mathscr{D}_\mu\phi)^*\mathscr{D}^\mu\phi - q^2\phi^*A_\mu\mathscr{D}^\mu\phi - iq\phi^*\mathscr{D}^2\phi + q^2\phi^*A_\mu\mathscr{D}^\mu\phi \tag{3.4.68}$$

$$= iq\phi(\mathscr{D}^2\phi)^* - iq\phi^*(\mathscr{D}^2\phi) = iq\phi(\mathscr{D}^2\phi - m^2\phi + m^2\phi)^* - iq\phi^*(\mathscr{D}^2\phi - m^2\phi + m^2\phi) \tag{3.4.69}$$

$$= iqm^2\phi^*\phi - iqm^2\phi^*\phi = 0. \tag{3.4.70}$$

Hence, we have our first full interacting theory of relativistic fields:

$$\mathfrak{L} = -\frac{1}{4}\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu} - (\mathscr{D}_\mu\phi)^*(\mathscr{D}^\mu\phi) - m^2|\phi|^2. \tag{3.4.71}$$

Now let us study the behavior of our field $\phi$ under parity and time reversal. There is no unique way to make $\phi$ transform under parity. One could take $\phi$ to transform as:

$$\phi_{\mathcal{P}}(t,\vec{x}) = \phi(t,-\vec{x}) \tag{3.4.72}$$

in which case $\phi$ is called a proper scalar or one could take the same expression with an overall minus sign (in which case $\phi$ is called a pseudoscalar). Similarly, one could take $\phi$ under time reversal to transform as:

$$\phi_{\mathcal{T}}(t,\vec{x}) = \phi(-t,\vec{x}) \tag{3.4.73}$$

or with a minus sign as well. We will take the time reversal as shown with no sign change by convention. There is a third interesting discrete transformation we have not yet studied: charge conjugation. Charge conjugation as a transformation is defined to invert all transformations on the internal space and not affect spacetime. For example, for the vector potential it has the internal transformation of the gauge transformation:

$$A'_\mu = A_\mu + \partial_\mu f \tag{3.4.74}$$

and so the charge conjugated vector potential $A_{\mathcal{C}\mu}$ must satisfy:

$$A'_{\mathcal{C}\mu} = A_{\mathcal{C}\mu} - \partial_\mu f \tag{3.4.75}$$

As such, the charge conjugated vector potential is simply defined as:

$$A_{\mathcal{C}\mu}(X) = -A_\mu(X). \tag{3.4.76}$$

Immediately then, the field strength is also odd under charge conjugation:

$$\mathcal{F}_{\mathcal{C}\mu\nu} = \partial_\mu A_{\mathcal{C}\nu} - \partial_\nu A_{\mathcal{C}\mu} = -\partial\mu A_\nu - \partial_\nu(-A_\mu) = -(\partial_\mu A_\nu - \partial_\nu A_\mu) = -\mathcal{F}_{\mu\nu}. \tag{3.4.77}$$

On $\phi(t,\vec{x})$ then the U(1) transformation must be inverted and so the charge conjugated field must satisfy:

$$\phi'_{\mathcal{C}} = e^{-i\theta}\phi_{\mathcal{C}}. \tag{3.4.78}$$

As such, the charge conjugated $\phi$ is simply defined as:

$$\phi_{\mathcal{C}}(X) = \phi^*(X). \tag{3.4.79}$$

Under parity transformation, we can see:

$$(\mathscr{D}_0\phi)_{\mathcal{P}} = \partial_{\mathcal{P}0}\phi_{\mathcal{P}} - iqA_{\mathcal{P}0}\phi_{\mathcal{P}} = \partial_0\phi(t,-\vec{x}) - iqA_0(t,-\vec{x})\phi(t,-\vec{x}) = \mathscr{D}_0\phi(t,-\vec{x}) \tag{3.4.80}$$

$$(\mathscr{D}_a\phi)_{\mathcal{P}} = \partial_{\mathcal{P}a}\phi_{\mathcal{P}} - iqA_{\mathcal{P}a}\phi_{\mathcal{P}} = -(\partial_a\phi)(t,-\vec{x}) - iq(-A_a(t,-\vec{x}))\phi(t,-\vec{x}) = -\mathscr{D}_a\phi(t,-\vec{x}) \tag{3.4.81}$$

$$\implies (\mathscr{D}_\mu\phi)_{\mathcal{P}} = \mathcal{P}^\nu{}_\mu\mathscr{D}_\nu\phi(t,-\vec{x}). \tag{3.4.82}$$

Consequently, the Lagrangian automatically has parity symmetry:

$$\mathfrak{L}_{\mathcal{P}}(t,-\vec{x}) = -\frac{1}{4}\left(\mathcal{P}\mathcal{F}\mathcal{P}\right)^{\mu\nu}\left(\mathcal{P}\mathcal{F}\mathcal{P}\right)_{\mu\nu} - \mathcal{P}^\nu{}_\mu\mathscr{D}_\nu\phi^*\mathcal{P}^{\rho\mu}\mathscr{D}_\rho\phi - m^2|\phi_{\mathcal{P}}|^2 \tag{3.4.83}$$

$$= -\frac{1}{4}\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu} - (\mathscr{D}_\mu\phi)^*(\mathscr{D}^\mu\phi) - m^2|\phi|^2 = \mathfrak{L}(t,\vec{x}) \tag{3.4.84}$$

$$\implies \mathcal{L}_{\mathcal{P}} = \int_{\mathbb{S}}\mathfrak{L}_{\mathcal{P}}(t,\vec{x})d^3\vec{x} = \int_{\mathbb{S}}\mathfrak{L}_{\mathcal{P}}(t,-\vec{x})d^3\vec{x} = \int_{\mathbb{S}}\mathfrak{L}(t,\vec{x})d^3\vec{x} = \mathcal{L}. \tag{3.4.85}$$

Recall the current density of classical particles:

$$J^0(x) = \rho(t,\vec{X}) = \sum_{n=1}^{N} q_n\delta(\vec{x}-\vec{x}_n(t)), \qquad J^a(X) = \sum_{n=1}^{N} q_n\dot{x}^a_n(t)\delta(\vec{x}-\vec{x}_n(t)). \tag{3.4.86}$$

The current density of classical particles transforms under parity as:

$$J^0_{\mathcal{P}}(X) = \sum_{n=1}^{N} q_n\delta(\vec{x}+\vec{x}_n(t)) = \sum_{n=1}^{N} q_n\delta((-\vec{x})-\vec{x}_n(t)) = J^0(t,-\vec{x}) \tag{3.4.87}$$

$$J^a_{\mathcal{P}}(X) = \sum_{n=1}^{N} q_n(-\dot{x}^a_n(t))\delta(\vec{x}+\vec{x}_n(t)) = -\sum_{n=1}^{N} q_n\dot{x}^a_n(t)\delta((-\vec{x})-\vec{x}_n(t)) = -J^a(t,-\vec{x}) \tag{3.4.88}$$

$$\implies J^\mu_{\mathcal{P}}(X') = (\mathcal{P}J)^\mu(X). \tag{3.4.89}$$

So, this is taken as the defining transformation property of the current density under parity. We can see that

in fact our derived current density has the correct property:

$$J^\mu_{\mathcal{P}}(X') = iq\phi_{\mathcal{P}}(\mathscr{D}^\mu\phi)^*_{\mathcal{P}} - iq\phi^*_{\mathcal{P}}\mathscr{D}^\mu\phi_{\mathcal{P}} = \mathcal{P}^\mu{}_\nu(iq\phi(\mathscr{D}^\nu\phi)^* - iq\phi^*(\mathscr{D}^\nu\phi)) = (\mathcal{P}J)^\mu(X). \tag{3.4.90}$$

The time reversal transformation has a very important property we have not yet discussed. By definition, the time reversal transformation in addition to transforming the fields takes the complex conjugate of any coefficients. This is due to Wigner's theorem. Under time reversal then:

$$(\mathscr{D}_\mu\phi)_{\mathcal{T}} = \mathcal{T}(\partial_\mu\phi - iqA_\mu\phi) = \partial_{\mathcal{T}\mu}\phi_{\mathcal{T}} + iqA_{\mathcal{T}\mu}\phi_{\mathcal{T}} \tag{3.4.91}$$

$$\implies (\mathscr{D}_0\phi)_{\mathcal{T}} = -\partial_0\phi(-t,\vec{x}) + iqA_0(-t,\vec{x})\phi(-t,\vec{x}) = -\mathscr{D}_0\phi(-t,\vec{x}) \tag{3.4.92}$$

$$\implies (\mathscr{D}_a\phi)_{\mathcal{T}} = \partial_a\phi(-t,\vec{x}) + iq(-A_a(-t,\vec{x}))\phi(-t,\vec{X}) = \mathscr{D}_a\phi(-t,\vec{x}) \tag{3.4.93}$$

$$\implies (\mathscr{D}_\mu\phi)_{\mathcal{T}}(X') = \mathcal{T}^\nu{}_\mu\mathscr{D}_\nu\phi(X). \tag{3.4.94}$$

Consequently, the Lagrangian density automatically has time reversal symmetry:

$$\mathfrak{L}_{\mathcal{T}}(-t,\vec{x}) = -\frac{1}{4}(-\mathcal{T}\mathcal{F}\mathcal{T})^{\mu\nu}(-\mathcal{T}\mathcal{F}\mathcal{T})_{\mu\nu} - \mathcal{T}^\nu{}_\mu\mathscr{D}_\nu\phi(X)^*\mathcal{T}^{\rho\mu}\mathscr{D}_\rho\phi - m^2|\phi_{\mathcal{T}}|^2 \tag{3.4.95}$$

$$= -\frac{1}{4}\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu} - \mathscr{D}_\mu\phi^*\mathscr{D}^\mu\phi - m^2|\phi|^2 = \mathfrak{L}(t,\vec{x}) \tag{3.4.96}$$

$$\implies S_{\mathcal{T}} = \int_{-t_f}^{-t_0}\int_{\mathbb{S}}\mathfrak{L}_{\mathcal{T}}(t,\vec{x})d\mathbb{S}dt = \int_{-t_f}^{-t_0}\int_{\mathbb{S}}\mathfrak{L}(-t,\vec{x})d\mathbb{S}dt = \int_{t_0}^{t_f}\int_{\mathbb{S}}\mathfrak{L}(t,\vec{x})d\mathbb{S}dt = S. \tag{3.4.97}$$

The current density of classical particles transforms under time reversal as:

$$J^0_{\mathcal{T}}(X) = \sum_{n=1}^{n} q_n\delta(\vec{x} - \vec{x}_n(-t)) = J^0(-t,\vec{x}) \tag{3.4.98}$$

$$J^a_{\mathcal{T}}(X) = \sum_{n=1}^{n} q_n\frac{d}{dt}(x^a_n(-t))\delta(\vec{x} - \vec{x}_n(-t)) = -\sum_{n=1}^{n} q_n\dot{x}^a_n(-t)\delta(\vec{x} - \vec{x}_n(-t)) = -J^a(-t,\vec{x}) \tag{3.4.99}$$

$$\implies J^\mu_{\mathcal{T}}(X') = -(\mathcal{T}J)^\mu(X). \tag{3.4.100}$$

The current density we derived matches this transformation property:

$$J^\mu_{\mathcal{T}}(X') = \mathcal{T}(iq\phi(\mathscr{D}^\mu\phi)^* - iq\phi^*\mathscr{D}^\mu\phi) = -iq\phi_{\mathcal{T}}(\mathscr{D}^\mu\phi_{\mathcal{T}})^* + iq\phi^*_{\mathcal{T}}\mathscr{D}^\mu\phi_{\mathcal{T}} \tag{3.4.101}$$

$$= -iq\phi\mathcal{T}^\mu{}_\nu(\mathscr{D}^\nu\phi)^* + iq\phi^*\mathcal{T}^\mu{}_\nu(\mathscr{D}^\nu\phi) = -\mathcal{T}^\mu{}_\nu J^\nu(X) = -(\mathcal{T}J)^\mu(X). \tag{3.4.102}$$

Under charge conjugation:

$$(\mathscr{D}_\mu\phi)_{\mathcal{C}} = \partial_\mu\phi_{\mathcal{C}} - iqA_{\mathcal{C}\mu}\phi_{\mathcal{C}} = \partial_\mu\phi^* + iqA_\mu\phi^* = \mathscr{D}_\mu\phi^*. \tag{3.4.103}$$

Consequently, the Lagrangian density automatically has charge conjugation symmetry:

$$\mathfrak{L}_{\mathcal{C}} = -\frac{1}{4}\mathcal{F}^{\mu\nu}_{\mathcal{C}}\mathcal{F}_{\mathcal{C}\mu\nu} - (\mathscr{D}_\mu\phi_{\mathcal{C}})^*(\mathscr{D}^\mu\phi_{\mathcal{C}}) - m^2|\phi_{\mathcal{C}}|^2 \tag{3.4.104}$$

$$= -\frac{1}{4}(-\mathcal{F}^{\mu\nu})(-\mathcal{F}_{\mu\nu}) - \mathscr{D}_\mu\phi(\mathscr{D}^\mu\phi)^* - m^2|\phi^*|^2 = \mathfrak{L}. \tag{3.4.105}$$

When we considered classical particles we never had internal degrees of freedom to invert and so we never defined charge conjugation. Under charge conjugation the current density we derived transforms as:

$$J^\mu_{\mathcal{C}} = iq\phi_{\mathcal{C}}(\mathscr{D}^\mu\phi_{\mathcal{C}})^* - iq\phi^*_{\mathcal{C}}\mathscr{D}^\mu\phi_{\mathcal{C}} = iq\phi^*(\mathscr{D}^\mu\phi) - iq\phi(\mathscr{D}^\mu\phi)^* = -J^\mu. \tag{3.4.106}$$

Thus, this is taken as the defining property of the current under charge conjugation.

## Exercises

**Problem 3.4: CPT transformation of scalar fields.**
(a)

**Problem 3.5: creation annihilation operators**
(a)

**Problem 3.6: hamiltonian for this theory**
(a)

**Problem 3.7: energy momentum tensor for scalar**
(a)

**Problem 3.8: solve equations of motion/diagonalize hamiltonian**
(a)

**Problem 3.9: stuckleberg lagrangian and proca equations**
(a)

## 3.5 Spinor Theory - Dirac Equation

The next possibilities after the scalar representation $s_l = s_r = 0$ are $s_l = \frac{1}{2}, \quad s_r = 0$ and $s_l = 0, \quad s_r = \frac{1}{2}$. We often notate these as an ordered pair $(s_l, s_r)$ and so we will now consider $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$. The first of these is then a two component field $\psi_L$ under which we have (using Theorem 1.11):

$$\hat{L}_a = \frac{\sigma_a}{2}, \qquad \hat{R}_a = 0 \implies \hat{J}_a = \frac{\sigma_a}{2}, \qquad \hat{K}_a = -\frac{i}{2}\sigma_a. \tag{3.5.1}$$

This $\psi_L$ is necessarily complex as the given $\hat{J}_a$ and $\hat{K}_a$ generate its Lorentz transformations are complex. The expression for $\hat{J}_a$ then reveals that $\psi_L$ is genuinely a spin 1/2 field as these $\hat{J}_a$ are the spin angular momentum matrices in general and for this field they correspond to those for spin 1/2 which we found in theorem 1.11. This field $\psi_L$ is called a left-handed Weyl spinor (because it transforms under a left-handed representation of the Lorentz group, meaning $s_r = 0$ and $s_l \neq 0$; this is called its chirality). Similarly, the second of these fields is then a two component field $\psi_R$ under which we have by the same theorem:

$$\hat{L}_a = 0, \qquad \hat{R}_a = \frac{\sigma_a}{2} \implies \hat{J}_a = \frac{\sigma_a}{2}, \qquad \hat{K}_a = \frac{i}{2}\sigma_a \tag{3.5.2}$$

and so $\psi_R$ is also spin 1/2 and necessarily complex. The field $\psi_R$ is called a right-handed Weyl spinor (it is right-handed because $s_l = 0$ and $s_r \neq 0$). We could study the field theory of either of these individually but it will be useful to study the field theory of both of them together. If one wants to then consider only one of the fields one can then simply set the other one to 0. So, define the 4 component complex spinor field $\psi$ in block matrix notation as:

$$\psi(X) = \begin{pmatrix} \psi_L(X) \\ \psi_R(X) \end{pmatrix}. \tag{3.5.3}$$

This $\psi$ is called a Dirac spinor. This $\psi$ then lives in the vector space in which the space of $\psi_L$'s is linearly combined with the space of $\psi_R$'s as we see and so this $\psi$ is an element of the direct sum of the two spaces: $\psi \in (\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$. Then, acting on $\psi$ in this block matrix notation we have:

$$\hat{J}_a = \begin{pmatrix} \frac{\sigma_a}{2} & 0 \\ 0 & \frac{\sigma_a}{2} \end{pmatrix}, \qquad \hat{K}_a = \begin{pmatrix} -\frac{i}{2}\sigma_a & 0 \\ 0 & \frac{i}{2}\sigma_a \end{pmatrix} \tag{3.5.4}$$

and so a general Lorentz transformation on $\psi$ is given by the operator:

$$U = e^{-i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a} = e^{-i\theta^{\mu\nu}\hat{S}_{\mu\nu}} \tag{3.5.5}$$

with these expressions for $\hat{J}_a$ and $\hat{K}_a$. The general Lorentz transformation decomposes into a left-handed and right-handed transformation as:

$$U = e^{-i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a} = \begin{pmatrix} e^{-\frac{i}{2}(\theta^a - i\varphi^a)\sigma_a} & 0 \\ 0 & e^{-\frac{i}{2}(\theta^a + i\varphi^a)\sigma_a} \end{pmatrix} = \begin{pmatrix} U_L & 0 \\ 0 & U_R \end{pmatrix}. \tag{3.5.6}$$

So, the left and right handed components transform separately as:

$$\psi'_L = U_L \psi_L, \qquad \psi'_R = U_R \psi_R. \tag{3.5.7}$$

Now that we have the explicit form of the Lorentz transformation of our field, we can beginning building a Lagrangian density with symmetry under it. As well, because we are dealing with a complex field we know we will ultimately be able to implement the U(1) transformation on it to couple it to electromagnetism through the transformation for some real $\theta$:

$$\psi' = e^{i\theta}\psi. \tag{3.5.8}$$

As such, we will need our Lagrangian density for $\psi$ to have a global U(1) symmetry, just as we did for the complex scalar field. Then, no terms linear in $\psi$ or $\psi^\dagger$ are possible as they are not preserved under U(1) symmetry. As well, no quadratic terms containing two powers of $\psi$ or two powers of $\psi^\dagger$ are possible as they also are not preserved under the U(1) transformation. Thus, to quadratic order, each term in the Lagrangian density must contain one power of $\psi^\dagger$ (with possible differentiation) and one power of $\psi$ (with possible differentiation) so that the factors of $e^{i\theta}$ cancel between the two of them so as to give global U(1) symmetry. Thus, we would like to study quantities of the form $\psi^\dagger \text{Stuff} \psi$, called bilinears. In order for our

Lagrangian density to have such a form, all transformations of our field $\psi$ must be linear combinations of $\psi$ and $\psi^*$ otherwise the resulting transformed quantities would no longer be bilinear in $\psi$ and $\psi^*$. In particular, it will be useful for us to study the parity transformation and because of the linear combination requirement under parity transformation the transformed $\psi$ will be given as:

$$\psi_{\mathcal{P}}(t,\vec{x}) = \hat{\mathcal{P}}\psi(t,-\vec{x}) + \hat{\mathcal{Q}}\psi^*(t,-\vec{x}) \tag{3.5.9}$$

for some $4\times 4$ matrices $\hat{\mathcal{P}}$ and $\hat{\mathcal{Q}}$ as this is a general linear combination of $\psi$ and $\psi^*$ with space inverted. In order for the parity transformation to be compatible with the U(1) transformation (meaning for the transformations to close under composition), if we parity transform a U(1) transformed $\psi$ it must be some U(1) transformation of the parity transformed $\psi$. That is, under the transformation in (3.5.8) there must be some real $\varphi$ so that:

$$\psi'_{\mathcal{P}} = e^{i\varphi}\psi_{\mathcal{P}} \tag{3.5.10}$$

$$\implies e^{i\theta}\hat{\mathcal{P}}\psi + e^{-i\theta}\hat{\mathcal{Q}}\psi^* = e^{i\varphi}\hat{\mathcal{P}}\psi + e^{i\varphi}\hat{\mathcal{Q}}\psi^*. \tag{3.5.11}$$

Differentiating with respect to $\theta$ at $\theta = 0$:

$$\hat{\mathcal{P}}\psi - \hat{\mathcal{Q}}\psi^* = \frac{d\varphi}{d\theta}(0)\hat{\mathcal{P}}\psi + \frac{d\varphi}{d\theta}(0)\hat{\mathcal{Q}}\psi^*. \tag{3.5.12}$$

This can only hold true for all $\psi$ if either $\hat{\mathcal{P}}\psi = 0$ or $\hat{\mathcal{Q}}\psi^*$ for all $\psi$. Consequently, either $\hat{\mathcal{P}} = 0$ or $\hat{\mathcal{Q}} = 0$. If $\hat{\mathcal{P}} = 0$ then we see that $\varphi = -\theta$ by (3.5.11). Then, the direction of the U(1) transformation is inverted by such a transformation. So, with $\hat{\mathcal{P}} = 0$ we can see that not only would such a transformation invert space but also be a charge conjugation. By definition, parity only inverts space and does not transform the internal degrees of freedom. As such, this possibility is excluded by the definition of parity. Then, we must have $\hat{\mathcal{Q}} = 0$ giving a parity transformation of the form:

$$\psi_{\mathcal{P}}(t,\vec{x}) = \hat{\mathcal{P}}\psi(t,-\vec{x}). \tag{3.5.13}$$

For a classical particle, under parity transformation its position is inverted:

$$\vec{x}_{\mathcal{P}}(t) = -\vec{x}(t). \tag{3.5.14}$$

Consequently, its velocity is inverted:

$$\dot{\vec{x}}_{\mathcal{P}}(t) = \frac{d\vec{x}_{\mathcal{P}}(t)}{dt} = -\frac{d\vec{x}}{dt} = -\dot{\vec{x}}, \tag{3.5.15}$$

its momentum is inverted:

$$\vec{p}_{\mathcal{P}} = \frac{\partial \mathcal{L}_{\mathcal{P}}}{\partial \dot{\vec{x}}_{\mathcal{P}}} = -\frac{\partial \mathcal{L}}{\partial \dot{\vec{x}}} = -\vec{p}, \tag{3.5.16}$$

its Hamiltonian is preserved:

$$\mathcal{H}_{\mathcal{P}} = \vec{p}_{\mathcal{P}}\cdot\dot{\vec{x}}_{\mathcal{P}} - \mathcal{L}_{\mathcal{P}} = \vec{p}\cdot\dot{\vec{x}} - \mathcal{L} = \mathcal{H}, \tag{3.5.17}$$

its angular momentum is preserved:

$$\vec{\mathrm{L}}_{\mathcal{P}} = \vec{x}_{\mathcal{P}}\times\vec{p}_{\mathcal{P}} = (-\vec{x})\times(-\vec{p}) = \vec{x}\times\vec{p} = \vec{\mathrm{L}} \tag{3.5.18}$$

and its generator of boosts is inverted:

$$\vec{\mathrm{K}}_{\mathcal{P}} = \vec{p}_{\mathcal{P}}t_{\mathcal{P}} - \mathcal{H}_{\mathcal{P}}\vec{x}_{\mathcal{P}} = (-\vec{p})t - \mathcal{H}(-\vec{x}) = -\vec{\mathrm{K}}. \tag{3.5.19}$$

Because our wave theory will have to match the particle theory in the classical limit, the same parity transformation properties are inherited for the total linear momentum, energy, angular momentum, and boost generators of our wave theory. Because $\hat{J}_a$ and $\hat{K}_a$ generate the Lorentz transformations among the components of $\psi$, under parity then we must have:

$$\hat{J}_{\mathcal{P}a} = \hat{J}_a, \qquad \hat{K}_{\mathcal{P}a} = -\hat{K}_a \tag{3.5.20}$$

so that the parity transformed $\psi$ transforms as:

$$\psi'_{\mathcal{P}} = U_{\mathcal{P}}\psi_{\mathcal{P}} = e^{-i\theta^a\hat{J}_{\mathcal{P}a} - i\varphi^a\hat{K}_{\mathcal{P}a}}\psi_{\mathcal{P}} = e^{-i\theta^a\hat{J}_a + i\varphi^a\hat{K}_a}\hat{\mathcal{P}}\psi. \tag{3.5.21}$$

We can directly compute how $\psi$ transforms under parity as:

$$\psi'_{\mathcal{P}} = \hat{\mathcal{P}}\psi' = \hat{\mathcal{P}}U\psi = \hat{\mathcal{P}}e^{-i\theta^a\hat{J}_a - i\varphi^a\hat{K}_a}\psi. \tag{3.5.22}$$

Matching these two:

$$e^{-i\theta^a \hat{J}_a + i\varphi^a \hat{K}_a}\hat{\mathcal{P}}\psi = \hat{\mathcal{P}}e^{-i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}\psi, \qquad \forall\psi \tag{3.5.23}$$

$$\implies e^{-i\theta^a \hat{J}_a + i\varphi^a \hat{K}_a}\hat{\mathcal{P}} = \hat{\mathcal{P}}e^{-i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}. \tag{3.5.24}$$

By differentiating both sides with respect to $\theta$ and $\varphi$ and evaluating at $(0,0)$ we find:

$$\frac{\partial}{\partial\theta^a}\bigg|_0 e^{-i\theta^b \hat{J}_b + i\varphi^b \hat{K}_b}\hat{\mathcal{P}} = \hat{\mathcal{P}}\,\frac{\partial}{\partial\theta^a}\bigg|_0 e^{-i\theta^b \hat{J}_b - i\varphi^b \hat{K}_b} \tag{3.5.25}$$

$$\implies \hat{J}_a\hat{\mathcal{P}} = \hat{\mathcal{P}}\hat{J}_a \tag{3.5.26}$$

$$\frac{\partial}{\partial\varphi^a}\bigg|_0 e^{-i\theta^b \hat{J}_b + i\varphi^b \hat{K}_b}\hat{\mathcal{P}} = \hat{\mathcal{P}}\,\frac{\partial}{\partial\varphi^a}\bigg|_0 e^{-i\theta^b \hat{J}_b - i\varphi^b \hat{K}_b} \tag{3.5.27}$$

$$\implies \hat{K}_a\hat{\mathcal{P}} = -\hat{\mathcal{P}}\hat{K}_a. \tag{3.5.28}$$

The $4\times 4$ matrix $\hat{\mathcal{P}}$ can be written in block notation as a $2\times 2$ matrix of $2\times 2$ matrices for some $2\times 2$ matrices $A, B, C, D$ as:

$$\hat{\mathcal{P}} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{3.5.29}$$

Any $2\times 2$ matrix can be written as a linear combination of the identity matrix and the Pauli matrices:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{a+d}{2}\mathbb{1} + \frac{b+c}{2}\sigma_1 + i\frac{b-c}{2}\sigma_2 + \frac{a-d}{2}\sigma_3. \tag{3.5.30}$$

Hence, we can write each of the matrices $A$, $B$, $C$, $D$ for some complex coefficients $A^\mu$, $B^\mu$, $C^\mu$, $D^\mu$ as:

$$A = A^0\mathbb{1} + A^a\sigma_a, \qquad B = B^0\mathbb{1} + B^a\sigma_a, \qquad C = C^0\mathbb{1} + C^a\sigma_a, \qquad D = D^0\mathbb{1} + D^a\sigma_a. \tag{3.5.31}$$

Then, we have from (3.5.26):

$$\hat{J}_a\hat{\mathcal{P}} = \hat{\mathcal{P}}\hat{J}_a \tag{3.5.32}$$

$$\implies \begin{pmatrix} \frac{\sigma_a}{2} & 0 \\ 0 & \frac{\sigma_a}{2} \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} \frac{\sigma_a}{2} & 0 \\ 0 & \frac{\sigma_a}{2} \end{pmatrix} \tag{3.5.33}$$

$$\implies \begin{pmatrix} \sigma_a A & \sigma_a B \\ \sigma_a C & \sigma_a D \end{pmatrix} = \begin{pmatrix} A\sigma_a & B\sigma_a \\ C\sigma_a & D\sigma_a \end{pmatrix}. \tag{3.5.34}$$

Hence $A, B, C, D$ have to commute with all of the $\sigma_a$. Because each Pauli matrix fails to commute with the other two, this means that all of the Pauli matrix terms for $A, B, C, D$ in (3.5.29) must be 0 and so we are just left with that each of them is proportional to the identity:

$$A = A^0\mathbb{1}, \qquad B = B^0\mathbb{1}, \qquad C = C^0\mathbb{1}, \qquad D = D^0\mathbb{1}. \tag{3.5.35}$$

Now applying this to (3.5.28):

$$-\hat{K}_a\hat{\mathcal{P}} = \hat{\mathcal{P}}\hat{K}_a \tag{3.5.36}$$

$$\implies -\begin{pmatrix} -i\frac{\sigma_a}{2} & 0 \\ 0 & i\frac{\sigma_a}{2} \end{pmatrix}\begin{pmatrix} A^0\mathbb{1} & B^0\mathbb{1} \\ C^0\mathbb{1} & D^0\mathbb{1} \end{pmatrix} = \begin{pmatrix} A^0\mathbb{1} & B^0\mathbb{1} \\ C^0\mathbb{1} & D^0\mathbb{1} \end{pmatrix}\begin{pmatrix} -i\frac{\sigma_a}{2} & 0 \\ 0 & i\frac{\sigma_a}{2} \end{pmatrix} \tag{3.5.37}$$

$$\implies \begin{pmatrix} A^0\sigma_a & B^0\sigma_a \\ -C^0\sigma_a & -D^0\sigma_a \end{pmatrix} = \begin{pmatrix} -A^0\sigma_a & B^0\sigma_a \\ -C^0\sigma_a & D^0\sigma_a \end{pmatrix} \tag{3.5.38}$$

$$\implies A^0 = 0, \quad D^0 = 0. \tag{3.5.39}$$

Hence we have that $\hat{\mathcal{P}}$ must be of the form:

$$\hat{\mathcal{P}} = \begin{pmatrix} 0 & B^0\mathbb{1} \\ C^0\mathbb{1} & 0 \end{pmatrix} \tag{3.5.40}$$

Then, parity transformation becomes:

$$\psi_{L\mathcal{P}} = B\psi_R, \qquad \psi_{R\mathcal{P}} = C\psi_L. \tag{3.5.41}$$

So far we have worked to define the parity transformation but we have not assumed parity is a symmetry of our system. Regardless of whether parity is a symmetry of our system, parity$^2$ is always a symmetry as inverting space twice leaves no inversion at all in the end. Because the Lagrangian density consists of bilinear terms in $\psi^\dagger$ and $\psi$, under parity$^2$ any such term must automatically have symmetry. So, for any matrix $M$

which appears in our Lagrangian density we must have:

$$\psi^\dagger_{\mathcal{P}^2} M \psi_{\mathcal{P}^2} = \psi^\dagger M \psi \tag{3.5.42}$$

$$\implies \psi^\dagger (BC)^* M (BC) \psi = \psi^\dagger M \psi \tag{3.5.43}$$

$$\implies (|BC|^2 - 1)\psi^\dagger M \psi = 0. \tag{3.5.44}$$

Hence if $\psi^\dagger M \psi$ is ever non-zero (which it must be otherwise our Lagrangian density trivially vanishes) we must have:

$$|BC|^2 = 1. \tag{3.5.45}$$

Thus, for some phase $\alpha$:

$$BC = e^{i\alpha}. \tag{3.5.46}$$

If we now define new fields $\psi_L^{\text{new}}$ and $\psi_R^{\text{new}}$ by:

$$\psi_L^{\text{new}} = \psi_L, \qquad \psi_R^{\text{new}} = e^{-\frac{i}{2}\alpha} B \psi_R. \tag{3.5.47}$$

Then, under Lorentz transformation:

$$\psi_L'^{\text{new}} = \psi_L' = U_L \psi_L = U_L \psi_L^{\text{new}}, \qquad \psi_R'^{\text{new}} = e^{-\frac{i}{2}\alpha} B \psi_R' = e^{-\frac{i}{2}\alpha} B U_R \psi_R = U_R \psi_R^{\text{new}}. \tag{3.5.48}$$

Thus, these new fields satisfy the same Lorentz transformation properties as the old fields. As well, they have the same U(1) transformation:

$$\psi_L'^{\text{new}} = \psi_L' = e^{i\theta} \psi_L = e^{i\theta} \psi_L^{\text{new}}, \qquad \psi_R'^{\text{new}} = e^{-\frac{i}{2}\alpha} B \psi_R' = e^{-\frac{i}{2}\alpha} B e^{i\theta} \psi_R = e^{i\theta} \psi_R^{\text{new}}. \tag{3.5.49}$$

But, the parity transformation is now:

$$\psi_{L\mathcal{P}}^{\text{new}} = \psi_{L\mathcal{P}} = B \psi_R = e^{\frac{i}{2}\alpha} \psi_R^{\text{new}}, \qquad \psi_{R\mathcal{P}}^{\text{new}} = e^{-\frac{i}{2}\alpha} B \psi_{R\mathcal{P}} = e^{-\frac{i}{2}\alpha} BC \psi_L = e^{\frac{i}{2}\alpha} \psi_L^{\text{new}}. \tag{3.5.50}$$

And so no longer writing new for these fields, we always choose our fields so that our parity transformation is of the form:

$$\psi_{L\mathcal{P}} = e^{\frac{i}{2}\alpha} \psi_R, \qquad \psi_{R\mathcal{P}} = e^{\frac{i}{2}\alpha} \psi_L. \tag{3.5.51}$$

We can also define a new parity transformation which is the composition of this parity transformation and a U(1) transformation by $e^{-\frac{i}{2}\alpha}$. This is allowed because of the U(1) symmetry and because the resulting transformation still inverts space. So, with this appropriately phase adjusted definition of the parity transformation of our field, we have that we can always define our Dirac spinor so that under parity:

$$\psi_{\mathcal{P}} = \hat{\mathcal{P}} \psi = \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} \implies \hat{\mathcal{P}} = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \tag{3.5.52}$$

This matrix appearing here for the parity transformation will continue to be important and so it is given a special name:

$$\gamma^0 \stackrel{\text{def}}{=} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \tag{3.5.53}$$

We can immediately see that $(\gamma^0)^2 = \mathbb{1}$. In particular, this matrix is helpful if we additionally define the Dirac adjoint $\bar{\psi}$:

$$\bar{\psi} \stackrel{\text{def}}{=} \psi^\dagger \gamma^0. \tag{3.5.54}$$

The Lorentz transformation properties of $\bar{\psi}$ are quite nice. By direct calculation:

$$\bar{\psi}' = \psi'^\dagger \gamma^0 = \psi^\dagger U^\dagger \gamma^0 = \psi^\dagger \gamma^0 \gamma^0 U^\dagger \gamma^0 = \bar{\psi} \gamma^0 U^\dagger \gamma^0. \tag{3.5.55}$$

This can simplify significantly by using from (3.1.46) that $\hat{J}_a^\dagger = \hat{J}_a$ and $\hat{K}_a^\dagger = -\hat{K}_a$ as well as (3.5.26) and (3.5.28). First, note that because $(\gamma^0)^2 = \mathbb{1}$ for any two matrices $A$ and $B$:

$$(\gamma^0 A \gamma^0)(\gamma^0 B \gamma^0) = \gamma^0 A B \gamma^0. \tag{3.5.56}$$

It immediately follows from $n$-fold application of this that:

$$(\gamma^0 A \gamma^0)^n = (\gamma^0 A \gamma^0)...(\gamma^0 A \gamma^0) = \gamma^0 A^n \gamma^0. \tag{3.5.57}$$

This, for any function of the matrix $A$ which is analytic near 0:

$$f(\gamma^0 A \gamma^0) = \sum_{n=0}^{\infty} c_n (\gamma^0 A \gamma^0)^n = \sum_{n=0}^{\infty} c_n \gamma^0 A^n \gamma^0 = \gamma^0 f(A) \gamma^0. \tag{3.5.58}$$

Then:

$$\gamma^0 U^\dagger \gamma^0 = \gamma^0 e^{i\theta^a \hat{J}_a^\dagger + i\varphi^a \hat{K}_a^\dagger}\gamma^0 = \gamma^0 e^{i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}\gamma^0 = e^{\gamma^0(i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a\gamma^0)} \tag{3.5.59}$$

$$= e^{i\theta^a \hat{\mathcal{P}}\hat{J}_a\hat{\mathcal{P}} - i\varphi^a \hat{\mathcal{P}}\hat{K}_a\hat{\mathcal{P}}} = e^{i\theta^a \hat{J}_a\hat{\mathcal{P}}\hat{\mathcal{P}} - i\varphi^a(-\hat{K}_a)\hat{\mathcal{P}}\hat{\mathcal{P}}} \tag{3.5.60}$$

$$= e^{i\theta^a \hat{J}_a + i\varphi^a \hat{K}_a} = U^{-1}. \tag{3.5.61}$$

So, under Lorentz transformation:

$$\bar{\psi}' = \bar{\psi}U^{-1}. \tag{3.5.62}$$

Because of this, instead of studying quantities of the form $\psi^\dagger \text{Stuff}\psi$, we will equivalently study $\bar{\psi}\text{Stuff}\psi$ because the Lorentz transformation properties will generally be clearer. Because $\psi^\dagger$ and $\bar{\psi}$ are invertibly related, one could of course do either. This is just a choice of convenience.

Now that we understand the transformation properties of $\psi$, let us try to find any scalar quantities of the form $\bar{\psi}\text{Stuff}\psi$. That is, let us try to identify $M$ so that under Lorentz transformation:

$$(\bar{\psi}M\psi)' \stackrel{!}{=} \bar{\psi}M\psi \tag{3.5.63}$$

for some $4\times 4$ matrix $M$. Then:

$$\implies \bar{\psi}'M\psi' = \psi^\dagger M\psi \tag{3.5.64}$$

$$\implies \bar{\psi}U^{-1}MU\psi = \bar{\psi}M\psi, \quad \forall\psi \tag{3.5.65}$$

$$\implies U^{-1}MU = M \tag{3.5.66}$$

$$\implies MU = UM. \tag{3.5.67}$$

$$\implies [M,U] = 0. \tag{3.5.68}$$

Recalling our expression for $U$,

$$[M, e^{-i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}] = 0 \tag{3.5.69}$$

By differentiating both sides with respect to $\theta$ or $\varphi$ and evaluating at $(0,0)$ we find:

$$[M, \frac{\partial}{\partial\theta^a}\Big|_0 e^{-i\theta^b \hat{J}_b - i\varphi^b \hat{K}_b}] = 0 \tag{3.5.70}$$

$$\implies [M, \hat{J}_a] = 0 \tag{3.5.71}$$

$$[M, \frac{\partial}{\partial\varphi^a}\Big|_0 e^{-i\theta^b \hat{J}_b - i\varphi^b \hat{K}_b}] = 0 \tag{3.5.72}$$

$$\implies [M, \hat{K}_a] = 0. \tag{3.5.73}$$

The requirement that $M$ commute with all of the $\hat{J}_a$ is exactly what $\hat{\mathcal{P}}$ had to satisfy in (1.6.26). That condition alone was used to arrive at (3.5.35) and so repeating those same steps we know that this means that $M$ must be of the form (for some coefficients $A, B, C, D$):

$$M = \begin{pmatrix} A\mathbb{1} & B\mathbb{1} \\ C\mathbb{1} & D\mathbb{1} \end{pmatrix}. \tag{3.5.74}$$

This only took into account that $M$ commutes with $\hat{J}_a$. Now requiring that $M$ commute with $\hat{K}_a$:

$$\begin{pmatrix} A\mathbb{1} & B\mathbb{1} \\ C\mathbb{1} & D\mathbb{1} \end{pmatrix}\begin{pmatrix} -\frac{i}{2}\sigma_a & 0 \\ 0 & \frac{i}{2}\sigma_a \end{pmatrix} = \begin{pmatrix} -\frac{i}{2}\sigma_a & 0 \\ 0 & \frac{i}{2}\sigma_a \end{pmatrix}\begin{pmatrix} A\mathbb{1} & B\mathbb{1} \\ C\mathbb{1} & D\mathbb{1} \end{pmatrix} \tag{3.5.75}$$

$$\implies \begin{pmatrix} -\frac{i}{2}A\sigma_a & \frac{i}{2}B\sigma_a \\ -\frac{i}{2}C\sigma_a & \frac{i}{2}D\sigma_a \end{pmatrix} = \begin{pmatrix} -\frac{i}{2}A\sigma_a & -\frac{i}{2}B\sigma_a \\ \frac{i}{2}C\sigma_a & \frac{i}{2}D\sigma_a \end{pmatrix}. \tag{3.5.76}$$

Thus, we must have $B = C = 0$ while $A$ and $D$ are unrestricted. So, any $M$ of the form:

$$M = \begin{pmatrix} A\mathbb{1} & 0 \\ 0 & D\mathbb{1} \end{pmatrix} \tag{3.5.77}$$

will satisfy (3.5.68). Let us see how such an $M$ behaves under parity transformation:

$$(\bar{\psi}M\psi)' = \psi'^\dagger\gamma^0 M\psi' = \psi^\dagger\hat{\mathcal{P}}\gamma^0 M\hat{\mathcal{P}}\psi = \psi^\dagger\gamma^0\gamma^0 M\gamma^0\psi = \bar{\psi}\gamma^0 M\gamma^0\psi. \tag{3.5.78}$$

Hence, under a parity transformation, the matrix $M$ is replaced by the combination $\gamma^0 M\gamma^0$. Explicitly this

gives:

$$\gamma^0 M \gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} A\mathbb{1} & 0 \\ 0 & D\mathbb{1} \end{pmatrix} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} = \begin{pmatrix} D\mathbb{1} & 0 \\ 0 & A\mathbb{1} \end{pmatrix}. \tag{3.5.79}$$

If we decompose the matrix $M$ as:

$$M = \begin{pmatrix} A\mathbb{1} & 0 \\ 0 & D\mathbb{1} \end{pmatrix} = \frac{A+D}{2} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} + \frac{D-A}{2} \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \tag{3.5.80}$$

then we can see that under parity:

$$\gamma^0 M \gamma^0 = \begin{pmatrix} D\mathbb{1} & 0 \\ 0 & A\mathbb{1} \end{pmatrix} = \frac{A+D}{2} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} - \frac{D-A}{2} \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \tag{3.5.81}$$

So, the first term in this decomposition does not change under parity while the second term is inverted. A quantity which is invariant under connected Lorentz transformations and which is parity even is called a proper Lorentz scalar. A quantity which is invariant under connected Lorentz transformations and which is parity odd is called a Lorentz pseudoscalar. So, we can see that:

$$\bar{\psi} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \psi \text{ is a proper scalar and } \bar{\psi} \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \psi \text{ is a pseudoscalar.}$$

Moreover, any scalar bilinear in $\psi$ and $\bar{\psi}$ is just a linear combination of these two. The proper scalar matrix here is just the identity. The pseudoscalar matrix will keep appearing and so it is given a name:

$$\gamma^5 \stackrel{\text{def}}{=} \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \tag{3.5.82}$$

Thus, any scalar bilinear is of the form (for some coefficients $m_1$ and $m_2$):

$$\bar{\psi} M \psi = m_1 \bar{\psi}\psi + m_2 \bar{\psi}\gamma^5\psi \tag{3.5.83}$$

with $\bar{\psi}\psi$ a proper scalar and $\bar{\psi}\gamma^5\psi$ a pseudoscalar. As well, the pseduoscalarity of the second term is encapsulated in the fact:

$$\gamma^0\gamma^5\gamma^0 = -\gamma^5 \implies \gamma^0\gamma^5 + \gamma^5\gamma^0 = 0. \tag{3.5.84}$$

So, they are said to anti-commute. Next let us look for bilinears which are 4-vectors. That is, we want to find matrices $\Gamma^\mu$ so that $\bar{\psi}\Gamma^\mu\psi$ is a 4-vector. Then, under a Lorentz transformation with matrix components $L^\mu{}_\nu$, we want:

$$(\bar{\psi}\Gamma^\mu\psi)' = L^\mu{}_\nu \bar{\psi}\Gamma^\nu\psi \tag{3.5.85}$$

$$\implies \bar{\psi}U^{-1}\Gamma^\mu U\psi = L^\mu{}_\nu\bar{\psi}\Gamma^\nu\psi \qquad \forall\psi \tag{3.5.86}$$

$$\implies U^{-1}\Gamma^\mu U = L^\mu{}_\nu\Gamma^\nu. \tag{3.5.87}$$

By the exponential expressions for $U$ and $L$ and by differentiating this condition on $\Gamma^\mu$ with respect to $\theta$ and $\varphi$ at the identity, we can find the possible $\Gamma^\mu$:

$$e^{i\theta^{\alpha\beta}\hat{S}_{\alpha\beta}}\Gamma^\mu e^{-i\theta^{\alpha\beta}\hat{S}_{\alpha\beta}} = (e^{\theta^{\alpha\beta}J_{\alpha\beta}})^\mu{}_\nu\Gamma^\nu \tag{3.5.88}$$

$$\implies \frac{d}{d\theta^{\rho\sigma}}\left(e^{i\theta^{\alpha\beta}\hat{S}_{\alpha\beta}}\Gamma^\mu e^{-i\theta^{\alpha\beta}\hat{S}_{\alpha\beta}}\right)\Big|_0 = \frac{d}{d\theta^{\rho\sigma}}\,(e^{\theta^{\alpha\beta}J_{\alpha\beta}})^\mu{}_\nu\Gamma^\nu\Big|_0 \tag{3.5.89}$$

$$\implies i\hat{S}_{\alpha\beta}\Gamma^\mu - i\Gamma^\mu\hat{S}_{\alpha\beta} = (J_{\alpha\beta})^\mu{}_\nu\Gamma^\nu \tag{3.5.90}$$

$$\implies [\Gamma^\mu, \hat{S}_{\alpha\beta}] = i(J_{\alpha\beta})^\mu{}_\nu\Gamma^\nu. \tag{3.5.91}$$

Now using our component expression for $J_{\alpha\beta}$ and rewriting $\hat{S}_{\alpha\beta}$ in terms of $\hat{J}_a$ and $\hat{K}_a$ we find:

$$[\Gamma^\mu, \hat{S}_{\alpha\beta}] = i(\delta^\mu_\beta\eta_{\alpha\nu} - \delta^\mu_\alpha\eta_{\beta\nu})\Gamma^\nu \tag{3.5.92}$$

$$\implies [\Gamma^0, \hat{K}_a] = [\Gamma^0, \hat{S}_{a0}] = i(\delta^0_0\eta_{a\nu} - \delta^0_a\eta_{0\nu})\Gamma^\nu = i\Gamma^a \tag{3.5.93}$$

$$\implies [\Gamma^0, \hat{J}_a] = \frac{1}{2}\epsilon_a{}^{bc}[\Gamma^0, \hat{S}_{bc}] = \frac{1}{2}\epsilon_a{}^{bc}i(\delta^0_c\eta_{b\nu} - \delta^0_b\eta_{c\nu})\Gamma^\nu = 0 \tag{3.5.94}$$

$$\implies [\Gamma^a, \hat{K}_b] = [\Gamma^a, \hat{S}_{b0}] = i(\delta^a_0\eta_{b\nu} - \delta^a_b\eta_{0\nu})\Gamma^\nu = i\delta^a_b\Gamma^0 \tag{3.5.95}$$

$$\implies [\Gamma^a, \hat{J}_b] = \frac{1}{2}\epsilon_b{}^{cd}[\Gamma^a, \hat{S}_{cd}] = \frac{1}{2}\epsilon_b{}^{cd}i(\delta^a_d\eta_{c\nu} - \delta^a_c\eta_{d\nu})\Gamma^\nu = i\epsilon^a{}_{bc}\Gamma^c. \tag{3.5.96}$$

Hence we now know what relations the $\Gamma^\nu$ must satisfy to have the correct Lorentz transformation

properties. Equation (3.5.93) is convenient as it lets us calculate $\Gamma^a$ directly from the commutator of $\Gamma^0$ with $\hat{K}_a$ and so if we have an expression for $\Gamma^0$ we have an expression for all the $\Gamma^\nu$. Just as before we can write each $\Gamma^\mu$ in block matrix notation in terms of 4 $2\times 2$ matrices as:

$$\Gamma^\mu = \begin{pmatrix} \Gamma^\mu_{LL} & \Gamma^\mu_{LR} \\ \Gamma^\mu_{RL} & \Gamma^\mu_{RR} \end{pmatrix}. \tag{3.5.97}$$

Then, the commutators with $\hat{J}_a$ and $\hat{K}_a$ can be written as:

$$[\Gamma^\mu, \hat{J}_a] = \begin{pmatrix} \Gamma^\mu_{LL} & \Gamma^\mu_{LR} \\ \Gamma^\mu_{RL} & \Gamma^\mu_{RR} \end{pmatrix} \begin{pmatrix} \frac{\sigma_a}{2} & 0 \\ 0 & \frac{\sigma_a}{2} \end{pmatrix} - \begin{pmatrix} \frac{\sigma_a}{2} & 0 \\ 0 & \frac{\sigma_a}{2} \end{pmatrix} \begin{pmatrix} \Gamma^\mu_{LL} & \Gamma^\mu_{LR} \\ \Gamma^\mu_{RL} & \Gamma^\mu_{RR} \end{pmatrix} \tag{3.5.98}$$

$$= \frac{1}{2} \begin{pmatrix} \Gamma^\mu_{LL}\sigma_a - \sigma_a\Gamma^\mu_{LL} & \Gamma^\mu_{LR}\sigma_a - \sigma_a\Gamma^\mu_{LR} \\ \Gamma^\mu_{RL}\sigma_a - \sigma_a\Gamma^\mu_{RL} & \Gamma^\mu_{RR}\sigma_a - \sigma_a\Gamma^\mu_{RR} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} [\Gamma^\mu_{LL}, \sigma_a] & [\Gamma^\mu_{LR}, \sigma_a] \\ [\Gamma^\mu_{RL}, \sigma_a] & [\Gamma^\mu_{RR}, \sigma_a] \end{pmatrix} \tag{3.5.99}$$

$$[\Gamma^\mu, \hat{K}_a] = \begin{pmatrix} \Gamma^\mu_{LL} & \Gamma^\mu_{LR} \\ \Gamma^\mu_{RL} & \Gamma^\mu_{RR} \end{pmatrix} \begin{pmatrix} -i\frac{\sigma_a}{2} & 0 \\ 0 & i\frac{\sigma_a}{2} \end{pmatrix} - \begin{pmatrix} -i\frac{\sigma_a}{2} & 0 \\ 0 & i\frac{\sigma_a}{2} \end{pmatrix} \begin{pmatrix} \Gamma^\mu_{LL} & \Gamma^\mu_{LR} \\ \Gamma^\mu_{RL} & \Gamma^\mu_{RR} \end{pmatrix} \tag{3.5.100}$$

$$= \frac{i}{2} \begin{pmatrix} -\Gamma^\mu_{LL}\sigma_a + \sigma_a\Gamma^\mu_{LL} & \Gamma^\mu_{LR}\sigma_a + \sigma_a\Gamma^\mu_{LR} \\ -\Gamma^\mu_{RL}\sigma_a - \sigma_a\Gamma^\mu_{RL} & \Gamma^\mu_{RR}\sigma_a - \sigma_a\Gamma^\mu_{RR} \end{pmatrix}. \tag{3.5.101}$$

The commutator of $\Gamma^\mu$ with $\hat{K}_a$ can be cleaned up a little by introducing the anticommutator $[\cdot,\cdot]_+$ of two matrices defined as:

$$[A, B]_+ \stackrel{\text{def}}{=} AB + BA \tag{3.5.102}$$

$$\implies [\Gamma^\mu, \hat{K}_a] = \frac{i}{2} \begin{pmatrix} -[\Gamma^\mu_{LL}, \sigma_a] & [\Gamma^\mu_{LR}, \sigma_a]_+ \\ -[\Gamma^\mu_{RL}, \sigma_a]_+ & [\Gamma^\mu_{RR}, \sigma_a] \end{pmatrix}. \tag{3.5.103}$$

Now let us consider $\Gamma^0$ in particular. Equation (3.5.93) just gives us an expression for $\Gamma^a$ from $\Gamma^0$ so the first real restriction on $\Gamma^0$ is (3.5.94). By (3.5.99), requiring $\Gamma^0$ to commute with $\hat{J}_a$ is equivalent to:

$$[\Gamma^0_{LL}, \sigma_a] = 0, \qquad [\Gamma^0_{LR}, \sigma_a] = 0, \qquad [\Gamma^0_{RL}, \sigma_a] = 0, \qquad [\Gamma^0_{RR}, \sigma_a] = 0. \tag{3.5.104}$$

By our previous reasoning, any $2\times 2$ matrix which commutes with all of the Pauli matrices is proportional to the identity and so we have that for complex coefficients $p, q, r, s$:

$$\Gamma^0_{LL} = p\mathbb{1}, \qquad \Gamma^0_{LR} = q\mathbb{1}, \qquad \Gamma^0_{RL} = r\mathbb{1}, \qquad \Gamma^0_{RR} = s\mathbb{1} \tag{3.5.105}$$

$$\implies \Gamma^0 = \begin{pmatrix} p\mathbb{1} & q\mathbb{1} \\ r\mathbb{1} & s\mathbb{1} \end{pmatrix}. \tag{3.5.106}$$

Now using (3.5.93) we find that for $\Gamma^a$:

$$\Gamma^a = \frac{1}{i}[\Gamma^0, \hat{K}_a] = \frac{1}{i}\frac{i}{2} \begin{pmatrix} -[p\mathbb{1}, \sigma_a] & [q\mathbb{1}, \sigma_a]_+ \\ -[r\mathbb{1}, \sigma_a]_+ & [s\mathbb{1}, \sigma_a] \end{pmatrix} = \begin{pmatrix} 0 & q\sigma_a \\ -r\sigma_a & 0 \end{pmatrix}. \tag{3.5.107}$$

Now requiring (3.5.95):

$$[\Gamma^a, \hat{K}_b] = \frac{i}{2} \begin{pmatrix} -[0, \sigma_b] & [q\sigma_a, \sigma_b]_+ \\ -[-r\sigma_a, \sigma_b]_+ & [0, \sigma_b] \end{pmatrix} = i\delta^a_b \begin{pmatrix} 0 & q\mathbb{1} \\ r\mathbb{1} & 0 \end{pmatrix} \stackrel{!}{=} i\delta^a_b \Gamma^0 \tag{3.5.108}$$

$$\implies p = s = 0. \tag{3.5.109}$$

Finally we have (3.5.96):

$$[\Gamma^a, \hat{J}_b] = \frac{1}{2} \begin{pmatrix} [0, \sigma_a] & [q\sigma_a, \sigma_b] \\ [-r\sigma_a, \sigma_b] & [0, \sigma_a] \end{pmatrix} = i\epsilon^a{}_{bc} \begin{pmatrix} 0 & q\sigma_c \\ -r\sigma_c & 0 \end{pmatrix} = i\epsilon^a{}_{bc}\Gamma^c \tag{3.5.110}$$

which is satisfied automatically by the restrictions so far. So, we have reduced the $\Gamma^\mu$ to depending on just two complex coefficients. In particular:

$$\Gamma^0 = \begin{pmatrix} 0 & q\mathbb{1} \\ r\mathbb{1} & 0 \end{pmatrix} = \frac{q+r}{2} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} + \frac{q-r}{2} \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} = \frac{q+r}{2}\gamma^0 + \frac{q-r}{2}\gamma^0\gamma^5 \tag{3.5.111}$$

$$\Gamma^a = \begin{pmatrix} 0 & q\sigma_a \\ -r\sigma_a & 0 \end{pmatrix} = \frac{q+r}{2} \begin{pmatrix} 0 & \sigma_a \\ -\sigma_a & 0 \end{pmatrix} + \frac{q-r}{2} \begin{pmatrix} 0 & \sigma_a \\ \sigma_a & 0 \end{pmatrix} = \frac{q+r}{2}\gamma^a + \frac{q-r}{2}\gamma^a\gamma^5 \tag{3.5.112}$$

where here we have defined the matrices $\gamma^a$ by:

$$\gamma^a = \begin{pmatrix} 0 & \sigma_a \\ -\sigma_a & 0 \end{pmatrix} \tag{3.5.113}$$

so that the expression for $\Gamma^a$ takes the same form as that for $\Gamma^0$. Finally let us consider the parity of the this

4-vector. As we found in (3.5.78), the parity transformation of a bilinear $\bar{\psi}M\psi$ results in $\bar{\psi}\gamma^0 M\gamma^0\psi$. For $\Gamma^0$:

$$\gamma^0\Gamma^0\gamma^0 = \gamma^0\left(\frac{q+r}{2}\gamma^0 + \frac{q-r}{2}\gamma^0\gamma^5\right)\gamma^0 = \frac{q+r}{2}\gamma^0 + \frac{q-r}{2}\gamma^5\gamma^0 = \frac{q+r}{2}\gamma^0 - \frac{q-r}{2}\gamma^0\gamma^5. \tag{3.5.114}$$

So, we can see the $\gamma^0$ piece is parity even while the $\gamma^0\gamma^5$ piece is parity odd. For $\Gamma^a$:

$$\gamma^0\Gamma^a\gamma^0 = \gamma^0\left(\frac{q+r}{2}\gamma^a + \frac{q-r}{2}\gamma^a\gamma^5\right)\gamma^0 = -\frac{q+r}{2}\gamma^a + \frac{q-r}{2}\gamma^0\gamma^a\gamma^0\gamma^0\gamma^5\gamma^0 = -\frac{q+r}{2}\gamma^a + \frac{q-r}{2}\gamma^a\gamma^5 \tag{3.5.115}$$

where here we have used that:

$$\gamma^0\gamma^a\gamma^0 = \begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix}\begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix}\begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix} = \begin{pmatrix}0 & -\sigma_a\\ \sigma_a & 0\end{pmatrix} = -\gamma^a. \tag{3.5.116}$$

So, we can see that the $\gamma^a$ piece is parity odd while the $\gamma^a\gamma^5$ piece is parity even. So, we see that for some coefficients $v_1$ and $v_2$, a general 4-vector bilinear can be written as:

$$\bar{\psi}\Gamma^\mu\psi = v_1\bar{\psi}\gamma^\mu\psi + v_2\bar{\psi}\gamma^\mu\gamma^5\psi \tag{3.5.117}$$

where $\bar{\psi}\gamma^\mu\psi$ has 0 component even under parity and spatial components odd under parity (called a proper vector) and $\bar{\psi}\gamma^\mu\gamma^5\psi$ has 0 component odd under parity and spatial components even under parity (called a pseudovector). The vectorial property of the $\gamma^\mu$ is contained in the identity:

$$U^{-1}\gamma^\mu U = L^\mu{}_\nu\gamma^\nu. \tag{3.5.118}$$

These $\gamma$-matrices (also called Dirac matrices) we have defined satisfy a number of very nice algebraic properties. Computing their commutators, we find:

$$[\gamma^0,\gamma^a] = \begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix}\begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix} - \begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix}\begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix} = 2\begin{pmatrix}-\sigma_a & 0\\ 0 & \sigma_a\end{pmatrix} = -4i\hat{K}_a \tag{3.5.119}$$

$$[\gamma^a,\gamma^b] = \begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix}\begin{pmatrix}0 & \sigma_b\\ -\sigma_b & 0\end{pmatrix} - \begin{pmatrix}0 & \sigma_b\\ -\sigma_b & 0\end{pmatrix}\begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix} = \begin{pmatrix}-\sigma_a\sigma_b+\sigma_b\sigma_a & 0\\ 0 & -\sigma_a\sigma_b+\sigma_b\sigma_a\end{pmatrix} \tag{3.5.120}$$

$$= \begin{pmatrix}-2i\epsilon^{abc}\sigma_c & 0\\ 0 & -2i\epsilon^{abc}\sigma_c\end{pmatrix} = -4i\epsilon^{abc}\hat{J}_c. \tag{3.5.121}$$

These relationships can be nicely rewritten in terms of the spin matrices $\hat{S}_{\mu\nu}$:

$$\hat{S}^{0a} = \eta^{0\alpha}\eta^{a\beta}\hat{S}_{\alpha\beta} = -\hat{S}_{0a} = \hat{K}_a = \frac{i}{4}[\gamma^0,\gamma^a] \tag{3.5.122}$$

$$\hat{S}^{ab} = \eta^{a\alpha}\eta^{b\beta}\hat{S}_{\alpha\beta} = \hat{S}_{ab} = \epsilon_{ab}{}^c\hat{J}_c = \frac{i}{4}[\gamma^a,\gamma^b] \tag{3.5.123}$$

$$\implies \hat{S}^{\mu\nu} = \frac{i}{4}[\gamma^\mu,\gamma^\nu]. \tag{3.5.124}$$

As well, the anticommutators of the $\gamma$-matrices are very nice:

$$[\gamma^0,\gamma^0]_+ = 2\gamma^{0\,2} = 2\mathbb{1} \tag{3.5.125}$$

$$[\gamma^0,\gamma^a]_+ = \gamma^0\gamma^a + \gamma^a\gamma^0 = \begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix}\begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix} + \begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix}\begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix} = 0 \tag{3.5.126}$$

$$[\gamma^a,\gamma^b]_+ = \gamma^a\gamma^b + \gamma^b\gamma^a = \begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix}\begin{pmatrix}0 & \sigma_b\\ -\sigma_b & 0\end{pmatrix} + \begin{pmatrix}0 & \sigma_b\\ -\sigma_b & 0\end{pmatrix}\begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix} \tag{3.5.127}$$

$$= \begin{pmatrix}-\sigma_a\sigma_b-\sigma_b\sigma_a & 0\\ 0 & -\sigma_a\sigma_b-\sigma_b\sigma_a\end{pmatrix} = -2\delta^{ab}\mathbb{1}. \tag{3.5.128}$$

These three rules can be combined nicely into one:

$$[\gamma^\mu,\gamma^\nu]_+ = -2\eta^{\mu\nu}\mathbb{1}. \tag{3.5.129}$$

The $\gamma^\mu$ also all anticommute with $\gamma^5$:

$$[\gamma^0,\gamma^5]_+ = \begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix}\begin{pmatrix}-\mathbb{1} & 0\\ 0 & \mathbb{1}\end{pmatrix} + \begin{pmatrix}-\mathbb{1} & 0\\ 0 & \mathbb{1}\end{pmatrix}\begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix} = 0 \tag{3.5.130}$$

$$[\gamma^a,\gamma^5]_+ = \begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix}\begin{pmatrix}-\mathbb{1} & 0\\ 0 & \mathbb{1}\end{pmatrix} + \begin{pmatrix}-\mathbb{1} & 0\\ 0 & \mathbb{1}\end{pmatrix}\begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix} = 0 \tag{3.5.131}$$

$$\implies [\gamma^\mu, \gamma^5]_+ = 0. \tag{3.5.132}$$

And, $\gamma^5$ can be written by products of the $\gamma^\mu$ by direct calculation:

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{3.5.133}$$

The matrix $\gamma^5$ is named as it is even though there is no $\gamma^4$ because some people count their coordinates as $x^1 = x, \quad x^2 = y, \quad x^3 = z, \quad x^4 = t$ instead of $x^0 = t, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z$. To avoid being confused whether $\gamma^4$ would refer to $\gamma^5$ or $\gamma^0$, both conventions just call it $\gamma^5$. Now that we have understood how to construct vectors out of the bilinears, the next natural object to search for would be rank-2 tensors. However, the spin matrices already form a rank 2 tensor:

$$(\bar\psi\hat S^{\mu\nu}\psi)' = \bar\psi U^{-1}\hat S^{\mu\nu}U\psi = \bar\psi U^{-1}\frac{i}{4}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)U\psi \tag{3.5.134}$$

$$= \bar\psi\frac{i}{4}(U^{-1}\gamma^\mu UU^{-1}\gamma^\nu U - U^{-1}\gamma^\nu UU^{-1}\gamma^\mu U)\psi \tag{3.5.135}$$

$$= \bar\psi\frac{i}{4}(L^\mu{}_\alpha\gamma^\alpha L^\nu{}_\beta\gamma^\beta - L^\nu{}_\beta\gamma^\beta L^\mu{}_\alpha\gamma^\alpha)\psi \tag{3.5.136}$$

$$= L^\mu{}_\alpha L^\nu{}_\beta\bar\psi\frac{i}{4}(\gamma^\alpha\gamma^\beta - \gamma^\beta\gamma^\alpha)\psi \tag{3.5.137}$$

$$= L^\mu{}_\alpha L^\nu{}_\beta\bar\psi\hat S^{\alpha\beta}\psi. \tag{3.5.138}$$

So, yes these bilinear do in fact form a rank 2 tensor. It is not immediately clear though that this should be the only rank 2 tensor possible. However, if we consider a generic $4\times 4$ matrix M, we can decompose it as:

$$M = \begin{pmatrix} A^0\mathbb{1} + A^a\sigma_a & B^0\mathbb{1} + B^a\sigma_a \\ C^0\mathbb{1} + C^a\sigma_a & D^0\mathbb{1} + D^a\sigma_a \end{pmatrix} \tag{3.5.139}$$

$$\begin{aligned} =&\frac{A^0+D^0}{2}\begin{pmatrix}\mathbb{1} & 0\\ 0 & \mathbb{1}\end{pmatrix} + \frac{D^0-A^0}{2}\begin{pmatrix}-\mathbb{1} & 0\\ 0 & \mathbb{1}\end{pmatrix} + \frac{B^0+C^0}{2}\begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix} + \frac{B^0-C^0}{2}\begin{pmatrix}0 & \mathbb{1}\\ -\mathbb{1} & 0\end{pmatrix}\\ &+ \frac{A^a+D^a}{2}\begin{pmatrix}\sigma_a & 0\\ 0 & \sigma_a\end{pmatrix} + \frac{D^a-A^a}{2}\begin{pmatrix}-\sigma^a & 0\\ 0 & \sigma^a\end{pmatrix} + \frac{B^a+C^a}{2}\begin{pmatrix}0 & \sigma_a\\ \sigma_a & 0\end{pmatrix} + \frac{B^a-C^a}{2}\begin{pmatrix}0 & \sigma_a\\ -\sigma_a & 0\end{pmatrix} \end{aligned} \tag{3.5.140}$$

$$\begin{aligned} =&\frac{A^0+D^0}{2}\mathbb{1} + \frac{D^0-A^0}{2}\gamma^5 + \frac{B^0+C^0}{2}\gamma^0 + \frac{B^0-C^0}{2}\gamma^0\gamma^5\\ &+ \frac{A^a+D^a}{2}2\hat J_a + \frac{D^a-A^a}{2}(-2i)\hat K_a + \frac{B^a+C^a}{2}\gamma^a\gamma^5 + \frac{B^a-C^a}{2}\gamma^a. \end{aligned} \tag{3.5.141}$$

So, for some appropriate coefficients any $4\times 4$ matrix M can be written in terms of pieces which respectively become a proper scalar, a pseudoscalar, a proper vector, a pseudovector, and an anti-symmetric rank 2 tensor when put into a bilinear:

$$M = m_1\mathbb{1} + m_2\gamma^5 + v_{1\mu}\gamma^\mu + v_{2\mu}\gamma^\mu\gamma^5 + w_{\mu\nu}\hat S^{\mu\nu}. \tag{3.5.142}$$

Thus, these objects we have defined so far form a complete basis of such matrices.

Because we have a complete basis of matrices for our bilinears and we understand all of their transformation properties it is quite easy to write down the possible terms for the Lagrangian density for our field. In order to make a term involving 0 derivatives of $\psi$ which is still a scalar, we must have:

$$\mathfrak{L}_{0-\text{der}} = -m_1\bar\psi\psi - m_2\bar\psi\gamma^5\psi \tag{3.5.143}$$

for some coefficients $m_1$ and $m_2$. In order to make a term involving 1 derivative of $\psi$ which is a scalar, because $\partial_\mu$ are the components of a 4-vector we must contract it with another 4-vector. Therefore, we must have:

$$\mathfrak{L}_{1-\text{der}} = v_1\bar\psi\gamma^\mu\partial_\mu\psi + v_2\bar\psi\gamma^\mu\gamma^5\psi + w_1\partial_\mu\bar\psi\gamma^\mu\psi + w_2\partial_\mu\bar\psi\gamma^\mu\gamma^5\psi. \tag{3.5.144}$$

We can integrate by parts the derivatives off of $\bar\psi$ onto $\psi$ by modifying the Lagrangian density by a surface term (total divergence):

$$\mathfrak{L}_{1-\text{der}} = v_1\bar\psi\gamma^\mu\partial_\mu\psi + v_2\bar\psi\gamma^\mu\gamma^5\psi + \partial_\mu\left(w_1\bar\psi\gamma^\mu\psi + w_2\bar\psi\gamma^\mu\gamma^5\psi\right) - w_1\bar\psi\gamma^\mu\partial_\mu\psi - w_2\bar\psi\gamma^\mu\gamma^5\partial_\mu\psi \tag{3.5.145}$$

$$\simeq (v_1 - w_1)\bar\psi\gamma^\mu\partial_\mu\psi + (v_2 - w_2)\bar\psi\gamma^\mu\gamma^5\psi. \tag{3.5.146}$$

Thus, we really only need to consider single derivative terms with the derivative directly on $\psi$, not $\bar\psi$. In fact, even considering terms with more than 1 derivative, whatever derivatives are on $\bar\psi$ in the bilinear can be integrated onto $\psi$ by parts by modifying the Lagrangian density by a surface term. Thus, we can

simply assume that all derivatives are placed directly on $\psi$ by definition. Previously, we always had to have at least 2 derivatives in our Lagrangian density to make a scalar and so this spin 1/2 theory is interesting in its ability to only have 1 derivative. Thus, keeping only the lowest number of fields of derivatives as necessary as usual, our lowest order possible Lagrangian density is then:

$$\mathfrak{L} = v_1\bar{\psi}\gamma^\mu\partial_\mu\psi + v_2\bar{\psi}\gamma^\mu\gamma^5\partial_\mu\psi - m_1\bar{\psi}\psi - m_2\bar{\psi}\gamma^5\psi. \tag{3.5.147}$$

These coefficients in principle can be complex and so let us write them in terms of real and imaginary parts.

$$\mathfrak{L} = (v_{1r} + iv_{1i})\bar{\psi}\gamma^\mu\partial_\mu\psi + (v_{2r} + iv_{2i})\bar{\psi}\gamma^\mu\gamma^5\partial_\mu\psi - (m_{1r} + im_{1i})\bar{\psi}\psi - (m_{2r} + im_{2i})\bar{\psi}\gamma^5\psi. \tag{3.5.148}$$

If we were to write the Lagrangian density in terms of the real and imaginary parts of $\psi$, we could treat those real fields as independent variables for the purpose of computing equations of motion. Then, the real part of the Lagrangian density contributes purely real terms to the equations of motion while the imaginary part would contribute purely imaginary terms. But, because those are real fields, there must be no imaginary part in their equations of motion. Therefore, any imaginary part of the Lagrangian density must fail to contribute to the equations of motion and thus must be a surface term. That is, we must have:

$$\mathfrak{L}^* \simeq \mathfrak{L}. \tag{3.5.149}$$

Imposing this:

$$\mathfrak{L}^* = (v_{1r} - iv_{1i})(\bar{\psi}\gamma^\mu\partial_\mu\psi)^\dagger + (v_{2r} - iv_{2i})(\bar{\psi}\gamma^\mu\gamma^5\partial_\mu\psi)^\dagger - (m_{1r} - im_{1i})(\bar{\psi}\psi)^\dagger - (m_{2r} - im_{2i})(\bar{\psi}\gamma^5\psi)^\dagger. \tag{3.5.150}$$

Let us compute these terms individually:

$$(\bar{\psi}\gamma^\mu\partial_\mu\psi)^\dagger = (\psi^\dagger\gamma^0\gamma^\mu\partial_\mu\psi)^\dagger = \partial_\mu\psi^\dagger\gamma^{\mu\dagger}\gamma^0\psi = \partial_\mu\psi^\dagger\gamma^0\gamma^0\gamma^{\mu\dagger}\gamma^0\psi = \partial_\mu\bar{\psi}\gamma^0\gamma^{\mu\dagger}\gamma^0\psi. \tag{3.5.151}$$

For $\mu = 0$, we use that $\gamma^0$ is self-adjoint:

$$\gamma^0\gamma^{0\dagger}\gamma^0 = \gamma^0\gamma^0\gamma^0 = \gamma^0. \tag{3.5.152}$$

For $\mu \neq 0$, we use that $\gamma^a$ is anti-self-adjoint:

$$\gamma^0\gamma^{a\dagger}\gamma^0 = \gamma^0(-\gamma^a)\gamma^0 = -\gamma^0\gamma^a\gamma^0 = \gamma^a\gamma^0\gamma^0 = \gamma^a. \tag{3.5.153}$$

Hence, for general $\mu$:

$$\gamma^0\gamma^{\mu\dagger}\gamma^0 = \gamma^\mu. \tag{3.5.154}$$

So:

$$(\bar{\psi}\gamma^\mu\partial_\mu\psi)^\dagger = \partial_\mu\bar{\psi}\gamma^\mu\psi. \tag{3.5.155}$$

Continuing:

$$(\bar{\psi}\gamma^\mu\gamma^5\partial_\mu\psi)^\dagger = \partial_\mu\psi^\dagger\gamma^{5\dagger}\gamma^{\mu\dagger}\gamma^{0\dagger}\psi = \partial_\mu\psi^\dagger\gamma^5\gamma^0\gamma^0\gamma^{\mu\dagger}\gamma^0\psi = \partial_\mu\psi^\dagger(-\gamma^0\gamma^5)\gamma^\mu\psi = \partial_\mu\bar{\psi}\gamma^\mu\gamma^5\psi \tag{3.5.156}$$

$$(\bar{\psi}\psi)^\dagger = (\psi^\dagger\gamma^0\psi)^\dagger = \psi^\dagger\gamma^0\psi = \bar{\psi}\psi \tag{3.5.157}$$

$$(\bar{\psi}\gamma^5\psi)^\dagger = (\psi^\dagger\gamma^0\gamma^5\psi)^\dagger = \psi^\dagger\gamma^5\gamma^0\psi = -\psi^\dagger\gamma^0\gamma^5\psi = -\bar{\psi}\gamma^5\psi. \tag{3.5.158}$$

So, we have:

$$\mathfrak{L}^* = (v_{1r} - iv_{1i})\partial_\mu\bar{\psi}\gamma^\mu\psi + (v_{2r} - iv_{2i})\partial_\mu\bar{\psi}\gamma^\mu\gamma^5\psi - (m_{1r} - im_{1i})\bar{\psi}\psi + (m_{2r} - im_{2i})\bar{\psi}\gamma^5\psi \tag{3.5.159}$$

$$\simeq -(v_{1r} - iv_{1i})\bar{\psi}\gamma^\mu\partial_\mu\psi - (v_{2r} - iv_{2i})\bar{\psi}\gamma^\mu\gamma^5\partial_\mu\psi - (m_{1r} - im_{1i})\bar{\psi}\psi + (m_{2r} - im_{2i})\bar{\psi}\gamma^5\psi \tag{3.5.160}$$

$$= \mathfrak{L} - 2v_{1r}\bar{\psi}\gamma^\mu\partial_\mu\psi - 2v_{2r}\bar{\psi}\gamma^\mu\gamma^5\partial_\mu\psi + 2im_{1i}\bar{\psi}\psi + 2m_{2r}\bar{\psi}\gamma^t\psi \tag{3.5.161}$$

Consequently:

$$v_{1r} = v_{2r} = m_{1i} = m_{2r} = 0. \tag{3.5.162}$$

So, for real coefficients $v$, $w$, $m$, $n$:

$$\mathfrak{L} = iv\bar{\psi}\gamma^\mu\partial_\mu\psi + iw\bar{\psi}\gamma^\mu\gamma^5\partial_\mu\psi - m\bar{\psi}\psi - in\bar{\psi}\gamma^5\psi. \tag{3.5.163}$$

When we couple this to electromagnetism, we know we must use the gauge covariant derivative:

$$\mathscr{D}_\mu\psi = \partial_\mu\psi - iqA_\mu\psi. \tag{3.5.164}$$

In the presence of the vector potential then this Lagrangian density becomes:

$$\mathfrak{L} = iv\bar{\psi}\gamma^\mu(\partial_\mu\psi - iqA_\mu\psi) + iw\bar{\psi}\gamma^\mu\gamma^5(\partial_\mu\psi - iqA_\mu\psi) - m\bar{\psi}\psi - in\bar{\psi}\gamma^5\psi. \tag{3.5.165}$$

Consequently, the current density produced is:

$$J^\mu = \frac{\partial \mathfrak{L}}{\partial A_\mu} = qv\bar{\psi}\gamma^\mu\psi + qw\bar{\psi}\gamma^\mu\gamma^5\psi. \tag{3.5.166}$$

By our parity transformation, we can see that the $v$ term is a proper vector and the $w$ term is a pseudovector. However, the electric current density is a proper vector. Thus, if this theory is to match the classical particles in the classical limit then the transformation properties of the current density demand that $w = 0$. So, we take as our Lagrangian density:

$$\mathfrak{L} = iv\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - in\bar{\psi}\gamma^5\psi. \tag{3.5.167}$$

Now we are finally faced with a tough physics question. Should we assume parity symmetry? So far it has come to us for free in relativity but now there is this pseudoscalar term which presents itself. There is nothing we can identify here that is problematic about such a term aside from it violating parity symmetry. Let us assume that we have parity symmetry and continue as doing so worked quite well in the Galilean theory. Then, we have:

$$\mathfrak{L} = iv\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi. \tag{3.5.168}$$

Consider the field redefinition:

$$\psi^{\text{new}} = \sqrt{|v|}\psi. \tag{3.5.169}$$

This new field has the same Lorentz transformation properties as the old:

$$\psi'^{\text{new}} = \sqrt{|v|}\psi' = \sqrt{|v|}U\psi = U\sqrt{|v|}\psi = U\psi^{\text{new}} \tag{3.5.170}$$

the same parity transformation as the old:

$$\psi_{\mathcal{P}}^{\text{new}} = \sqrt{|v|}\psi_{\mathcal{P}} = \sqrt{|v|}\hat{\mathcal{P}}\psi = \hat{\mathcal{P}}\sqrt{|v|}\psi = \hat{\mathcal{P}}\psi^{\text{new}} \tag{3.5.171}$$

and the same U(1) transformation as the old:

$$\psi'^{\text{new}} = \sqrt{|v|}\psi' = \sqrt{|v|}e^{i\theta}\psi = e^{i\theta}\sqrt{|v|}\psi = e^{i\theta}\psi^{\text{new}}. \tag{3.5.172}$$

So, this field is every bit as good of a choice of $\psi$ as our old field and so dropping the "new" label, the Lagrangian density in terms of this $\psi$ becomes:

$$\mathfrak{L} = i\frac{v}{|v|}\bar{\psi}\gamma^\mu\partial_\mu\psi - \frac{m}{|v|}\bar{\psi}\psi. \tag{3.5.173}$$

Thus, with this scaling we scale the value of $m$ to a new value $\frac{m}{|v|}$ and the coefficient on the first term in the Lagrangian density is set to either 1 or $-1$. So, assuming we have applied such a scaling we can simply rename $\frac{m}{|v|}$ to $m$ and assume $v$ is $\pm 1$.

Suppose $v = -1$. Then, take the field redefinition:

$$\psi^{\text{new}} = -i\gamma^2\psi^*. \tag{3.5.174}$$

Showing that the redefined field $\psi^{\text{new}}$ has the correct Lorentz transformation property is a little tricky and relies on a special property of $\gamma^2$. Notice that $\sigma_1$ and $\sigma_3$ have entirely real valued entries while $\sigma_2$ has entirely imaginary valued entries. Therefore, $\gamma^\mu$ for $\mu \neq 2$ has real entries and $\gamma^2$ has imaginary entries. Hence:

$$\gamma^{\mu^*} = \gamma^\mu \quad (\mu \neq 2), \qquad \gamma^{2^*} = -\gamma^2. \tag{3.5.175}$$

Using the anticommutation relations for the $\gamma$-matrices:

$$\gamma^2\gamma^{\mu^*}\gamma^2 = \gamma^2\gamma^\mu\gamma^2 \quad (\mu \neq 2) \tag{3.5.176}$$

$$\implies \gamma^2\gamma^{\mu^*}\gamma^2 = -\gamma^\mu\gamma^2\gamma^2 \quad (\mu \neq 2) \tag{3.5.177}$$

$$\implies \gamma^2\gamma^{\mu^*}\gamma^2 = \gamma^\mu \quad (\mu \neq 2). \tag{3.5.178}$$

As well, for $\mu = 2$:

$$\gamma^2\gamma^{2^*}\gamma^2 = -\gamma^2\gamma^2\gamma^2 = \gamma^2. \tag{3.5.179}$$

Hence, for even for $\mu = 2$:

$$\gamma^2\gamma^{\mu^*}\gamma^2 = \gamma^\mu \implies \gamma^{\mu^*} = \gamma^2\gamma^\mu\gamma^2. \tag{3.5.180}$$

Then,

$$\psi'^{\text{new}} = -i\gamma^2\psi'^* = -i\gamma^2U^*\psi^* \tag{3.5.181}$$

$$U^* = \left(e^{-i\theta_{\mu\nu}\hat{S}^{\mu\nu}}\right)^* = e^{i\theta_{\mu\nu}\hat{S}^{\mu\nu *}} \tag{3.5.182}$$

$$\hat{S}^{\mu\nu *} = \left(\frac{i}{4}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)\right)^* = -\frac{i}{4}(\gamma^{\mu *}\gamma^{\nu *} - \gamma^{\nu *}\gamma^{\mu *}) \tag{3.5.183}$$

$$\implies \hat{S}^{\mu\nu *} = -\frac{i}{4}(\gamma^2\gamma^\mu\gamma^2\gamma^2\gamma^\nu\gamma^2 - \gamma^2\gamma^\nu\gamma^2\gamma^2\gamma^\mu\gamma^2) = \gamma^2\frac{i}{4}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)\gamma^2 = \gamma^2\hat{S}^{\mu\nu}\gamma^2 \tag{3.5.184}$$

$$\implies U^* = e^{i\theta_{\mu\nu}\hat{S}^{\mu\nu *}} = e^{i\theta_{\mu\nu}\gamma^2\hat{S}^{\mu\nu}\gamma^2}. \tag{3.5.185}$$

Because $\left(\gamma^2\right)^2 = -\mathbb{1}$, we have that for any matrices $A$ and $B$:

$$(\gamma^2 A\gamma^2)(\gamma^2 B\gamma^2) = -\gamma^2 AB\gamma^2. \tag{3.5.186}$$

It immediately follows by the $n$ fold iteration that:

$$(\gamma^2 A\gamma^2)^n = (-1)^{n-1}\gamma^2 A^n\gamma^2. \tag{3.5.187}$$

This allows a nice simplification of (3.5.185) by using the Taylor series definition of the matrix exponential:

$$U^* = e^{i\theta_{\mu\nu}\gamma^2\hat{S}^{\mu\nu}\gamma^2} = \sum_{n=0}^{\infty}\frac{i^n}{n!}\left(\gamma^2\theta_{\mu\nu}\hat{S}^{\mu\nu}\gamma^2\right)^n = \sum_{n=0}^{\infty}\frac{i^n}{n!}(-1)^{n-1}\gamma^2\left(\theta_{\mu\nu}S^{\mu\nu}\right)^n\gamma^2 \tag{3.5.188}$$

$$= -\gamma^2\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\gamma^2\left(\theta_{\mu\nu}\hat{S}^{\mu\nu}\right)^n\gamma^2 = -\gamma^2 e^{-i\theta_{\mu\nu}\hat{S}^{\mu\nu}}\gamma^2 = -\gamma^2 U\gamma^2 \tag{3.5.189}$$

Now we can see that when a Lorentz transformation is applied to $\psi$, the same transformation is applied to $\psi^{\text{new}}$:

$$\psi'^{\text{new}} = -i\gamma^2\psi'^* = -i\gamma^2 U^*\psi = -i\gamma^2(-\gamma^2 U\gamma^2)\psi = -iU\gamma^2\psi = U\psi^{\text{new}}. \tag{3.5.190}$$

On this new field, the parity transformation becomes:

$$\psi^{\text{new}}_{\mathcal{P}} = -i\gamma^2\psi^*_{\mathcal{P}} = -i\gamma^2\gamma^0\psi^* = \gamma^0 i\gamma^2\psi^* = -\mathcal{P}\psi^{\text{new}}. \tag{3.5.191}$$

So, the parity transformation picks up a minus sign. However, recall that we can simply adjust the definition of the parity transformation by an overall phase using the U(1) symmetry and so we can always absorb this minus sign into the definition of the parity transformation. As well, on this field the U(1) transformation becomes:

$$\psi'^{\text{new}} = -i\gamma^2\psi'^* = -i\gamma^2(e^{i\theta}\psi)^* = -e^{-i\theta}i\gamma^2\psi^* = e^{-i\theta}\psi^{\text{new}}. \tag{3.5.192}$$

So, the U(1) transformation is inverted on this new field (but the symmetry itself is preserved) and the Lorentz transformation is unchanged. Consequently, this $\psi^{\text{new}}$ is actually the charge conjugation of $\psi$. Thus, for the Dirac spinor charge conjugation is defined as:

$$\psi_{\mathcal{C}} = -i\gamma^2\psi^*. \tag{3.5.193}$$

Because the Lorentz transformation properties of $\psi^{\text{new}}$ and $\psi$ are the same and the parity and U(1), transformations are preserved up to convention (phase in the case of parity, direction in the case of U(1)), we can take $\psi^{\text{new}}$ as our redefined field. Then, we can invert to find $\psi$ in terms of $\psi^{\text{new}}$:

$$\psi^{\text{new}} = -i\gamma^2\psi^* \implies \psi = -i\gamma^2\psi^{\text{new}*} \tag{3.5.194}$$

$$\implies \bar{\psi} = \psi^\dagger\gamma^0 = (-i\gamma^2\psi^{\text{new}*})^\dagger\gamma^0 = -i\psi^{\text{new}T}\gamma^2\gamma^0. \tag{3.5.195}$$

Supposing that $v = -1$, this redefinition changes it to a positive sign:

$$\mathfrak{L} = -i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \tag{3.5.196}$$

$$= -ii\psi^{\text{new}T}\gamma^2\gamma^0\gamma^\mu i\gamma^2\partial_\mu\psi^{\text{new}*} - mi\psi^{\text{new}T}\gamma^2\gamma^0 i\gamma^2\psi^{\text{new}*} \tag{3.5.197}$$

$$= i\psi^{\text{new}T}\gamma^2\gamma^0\gamma^\mu\gamma^2\partial_\mu\psi^{\text{new}*} + m\psi^{\text{new}T}\gamma^2\gamma^0\gamma^2\psi^{\text{new}*} \tag{3.5.198}$$

$$= -i\psi^{\text{new}T}\gamma^0\gamma^2\gamma^\mu\gamma^2\partial_\mu\psi^{\text{new}*} - m\psi^{\text{new}T}\gamma^0\gamma^2\gamma^2\psi^{\text{new}*} \tag{3.5.199}$$

$$= -i\psi^{\text{new}T}\gamma^0(\gamma^\mu)^*\partial_\mu\psi^{\text{new}*} + m\psi^{\text{new}T}\gamma^0\psi^{\text{new}*} \tag{3.5.200}$$

$$= -i\left(\psi^{\text{new}T}\gamma^0(\gamma^\mu)^*\partial_\mu\psi^{\text{new}*}\right)^T + m\left(\psi^{\text{new}T}\gamma^0\psi^{\text{new}*}\right)^T \tag{3.5.201}$$

$$= -i\partial_\mu\psi^{\text{new}\dagger}(\gamma^\mu)^\dagger\gamma^0\psi^{\text{new}} + m\psi^{\text{new}\dagger}\gamma^0\psi^{\text{new}} \tag{3.5.202}$$

$$= -i\partial_\mu\bar{\psi}^{\text{new}}\gamma^0(\gamma^\mu)^\dagger\gamma^0\psi^{\text{new}} + m\bar{\psi}^{\text{new}}\psi^{\text{new}} \tag{3.5.203}$$

$$= -i\partial_\mu \bar{\psi}^{\text{new}} \gamma^\mu \psi^{\text{new}} + m\bar{\psi}^{\text{new}}\psi^{\text{new}} \tag{3.5.204}$$
$$\simeq i\bar{\psi}^{\text{new}} \gamma^\mu \partial_\mu \psi^{\text{new}} + m\bar{\psi}^{\text{new}}\psi^{\text{new}}. \tag{3.5.205}$$

The last step is made by integration by parts. Hence this field redefinition changes the signs of both terms and so we can use it to guarantee that the first term has a plus sign.

Once we have fixed the sign of the first term to be plus, we can fix the sign of the second term by another field redefinition. Suppose the second term has $m < 0$. Then, consider the field redefinition:

$$\psi^{\text{new}} = \gamma^5 \psi \tag{3.5.206}$$
$$\implies \psi = \gamma^5 \psi^{\text{new}}, \qquad \bar{\psi} = \psi^\dagger \gamma^0 = \psi^{\text{new}\dagger} \gamma^5 \gamma^0 = -\psi^{\text{new}\dagger} \gamma^0 \gamma^5 = -\bar{\psi}^{\text{new}} \gamma^5. \tag{3.5.207}$$

We can show that this field redefinition doesn't change our Lorentz transformation by using that $\gamma^5$ anti-commutes with all of the $\gamma^\mu$. This means that it commutes with the spin matrices:

$$\hat{S}^{\mu\nu}\gamma^5 = \frac{i}{4}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)\gamma^5 = \frac{i}{4}(\gamma^\mu(-\gamma^5)\gamma^\nu - \gamma^\nu(-\gamma^5)\gamma^\mu) = \gamma^5 \frac{i}{4}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) = \gamma^5 \hat{S}^{\mu\nu} \tag{3.5.208}$$
$$\implies \gamma^5 \hat{S}^{\mu\nu}\gamma^5 = (\gamma^5)^2 \hat{S}^{\mu\nu} = \hat{S}^{\mu\nu}. \tag{3.5.209}$$

Then, because $(\gamma^5)^2 = \mathbb{1}$, we have that for any matrix $A$:

$$\gamma^5 U \gamma^5 = \gamma^5 e^{-i\theta^{\mu\nu}\hat{S}_{\mu\nu}}\gamma^5 = e^{-i\theta^{\mu\nu}\gamma^5\hat{S}_{\mu\nu}\gamma^5} = e^{-i\theta^{\mu\nu}\hat{S}_{\mu\nu}\gamma^5\gamma^5} = e^{-i\theta^{\mu\nu}\hat{S}_{\mu\nu}} = U \tag{3.5.210}$$
$$\implies \psi'^{\text{new}} = \gamma^5\psi' = \gamma^5 U\psi = \gamma^5 U\gamma^5\gamma^5\psi = U\psi^{\text{new}}. \tag{3.5.211}$$

Under parity transformation:

$$\psi^{\text{new}}_{\mathcal{P}} = \gamma^5\psi_{\mathcal{P}} = \gamma^5\gamma^0\psi = -\gamma^0\gamma^5\psi = -\mathcal{P}\psi^{\text{new}}. \tag{3.5.212}$$

So, again up to an arbitrary phase the parity transformation is preserved. As well, under U(1) transformation:

$$\psi'^{\text{new}} = \gamma^5\psi' = \gamma^5 e^{i\theta}\psi = e^{i\theta}\psi^{\text{new}}. \tag{3.5.213}$$

And so, the U(1) transformation is preserved. Thus, we can take this $\psi^{\text{new}}$ as our redefined field. Then, in the Lagrangian density:

$$\mathfrak{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \tag{3.5.214}$$
$$= i(-\bar{\psi}^{\text{new}}\gamma^5)\gamma^\mu\partial_\mu(\gamma^5\psi^{\text{new}}) - m(-\bar{\psi}^{\text{new}}\gamma^5)\gamma^5\psi^{\text{new}} \tag{3.5.215}$$
$$= -i\bar{\psi}^{\text{new}}(-\gamma^\mu\gamma^5)\gamma^5\partial_\mu\psi^{\text{new}} + m\bar{\psi}^{\text{new}}\psi^{\text{new}} \tag{3.5.216}$$
$$= i\bar{\psi}^{\text{new}}\gamma^\mu\partial_\mu\psi^{\text{new}} + m\bar{\psi}^{\text{new}}\psi^{\text{new}}. \tag{3.5.217}$$

Hence it is equivalent to sending $m \to -m$. Now we can say that we choose the definition of $\psi$ so that by convention (through all of the previous field redefinitions) the Lagrangian density of the theory when interacting with no other fields is:

$$\boxed{\mathfrak{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi.} \tag{3.5.218}$$

where $m$ is a non-negative positive constant. This is called the Dirac Lagrangian Density. We can easily find the equation of motion by varying with respect to $\psi^\dagger$. Because $\bar{\psi} = \psi^\dagger\gamma^0$, the variation with respect to $\psi^\dagger$ is proportional to the variation with respect to $\bar{\psi}$ by the chain rule and so one vanishes if and only if the other does:

$$\frac{\partial\mathfrak{L}}{\partial\bar{\psi}} - \partial_\mu\frac{\partial\mathfrak{L}}{\partial\partial_\mu\bar{\psi}} = i\gamma^\mu\partial_\mu\psi - m\psi \stackrel{!}{=} 0. \tag{3.5.219}$$

This equation of motion is called the Dirac equation. By taking the derivative of this equation, we recover the Klein-Gordon equation:

$$i\gamma^\mu\partial_\mu\psi = m\psi \tag{3.5.220}$$
$$\implies i\gamma^\mu\partial_\mu i\gamma^\nu\partial_\nu\psi = im\gamma^\mu\partial_\mu\psi \tag{3.5.221}$$
$$\implies -\gamma^\mu\gamma^\nu\partial^2_{\mu\nu}\psi = m^2\psi \tag{3.5.222}$$
$$\implies -\left(\frac{\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu}{2} + \frac{\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu}{2}\right)\partial^2_{\mu\nu}\psi = m^2\psi \tag{3.5.223}$$
$$\implies \eta^{\mu\nu}\partial^2_{\mu\nu}\psi + 0 = m^2\psi \tag{3.5.224}$$
$$\implies (\partial^2 - m^2)\psi = 0. \tag{3.5.225}$$

So, while in the Dirac equation the individual components of $\psi$ are coupled to each other if one takes a second derivative all of the components decouple into the Klein-Gordon equation for each component separately. We can also give this parameter $m$ here the interpretation of mass just as we did with the parameter $m$ in the scalar field theory by considering the non-relativistic limit. First, notice that this LD is actually already in the Schrödinger form:

$$\mathfrak{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \tag{3.5.226}$$
$$= i\psi^\dagger\gamma^0\gamma^0\dot{\psi} + i\psi^\dagger\gamma^0\gamma^a\partial_a\psi - m\psi^\dagger\gamma^0\psi \tag{3.5.227}$$
$$= i\psi^\dagger\dot{\psi} - \psi^\dagger\gamma^0(-i\gamma^a\partial_a + m)\psi. \tag{3.5.228}$$

Because this LD is in the form of the Schrödinger Lagrangian, we can identify $\psi$ as the Schrödinger wavefunction $\psi$, $\hat{p}_a = -i\partial_a$, and $\hat{H} = \gamma^0\gamma^a\hat{p}_a + m$. Then, by direct calculation:

$$\hat{H}^2 = (\gamma^0\gamma^a\hat{p}_a + \gamma^0 m)^2 \tag{3.5.229}$$
$$= \gamma^0\gamma^a\gamma^0\hat{p}_a\hat{p}_b + (\gamma^0\gamma^a + \gamma^a\gamma^0)m\hat{p}_a + m^2 \tag{3.5.230}$$
$$= -\gamma^a\gamma^b\hat{p}_a\hat{p}_b + 0 + m^2 \tag{3.5.231}$$
$$= -\left(\frac{\gamma^a\gamma^b + \gamma^b\gamma^a}{2} + \frac{\gamma^a\gamma^b - \gamma^b\gamma^a}{2}\right)\hat{p}_a\hat{p}_b + m^2 \tag{3.5.232}$$
$$= \delta^{ab}\hat{p}_a\hat{p}_b + 0 + m^2 \tag{3.5.233}$$
$$= \hat{p}^2 + m^2. \tag{3.5.234}$$

Thus, we can see that $\hat{H}^2$ has the same eigenstates as the momentum operator and on those states it has spectrum $\hat{p}^2 + m^2$. By canonically quantizing the classical relativistic free particle Hamiltonian $\mathcal{H} = \sqrt{|\vec{p}|^2 + m^2}$ we see that the spectrum of $\hat{H}^2$ would be $\hat{p}^2 + m^2$ as well and thus we can identify the $m$ appearing in the Dirac equation as the classical particle's mass. (In principle, this only allows us to identify that the two $m^2$'s are the same, but because we have ensured by definition that both $m$'s are positive this means that they are indeed the same.)

Now that we have a decent picture of the Dirac Lagrangian in isolation, let us couple it to other theories. Of course we would like to make an interaction with electromagnetism. We know that the gauge symmetry of electromagnetism corresponds to a local U(1) gauge symmetry. A U(1) transformation is just multiplication by a complex phase:

$$\psi' = e^{iqf}\psi \tag{3.5.235}$$

We already know that to make our theory invariant under the local gauge symmetry we must use the gauge covariant derivative. So, define:

$$\mathscr{D}_\mu\psi = (\partial_\mu - iqA_\mu)\psi \tag{3.5.236}$$

This way under gauge transformation we have:

$$(\mathscr{D}_\mu\psi)' = (\partial_\mu - iqA'_\mu)\psi' = (\partial_\mu - iqA_\mu - iq\partial_\mu f)\left(e^{iqf}\psi\right) = e^{iqf}\mathscr{D}_\mu\psi. \tag{3.5.237}$$

Hence, this combination uniquely cancels the dependence on the spacetime derivatives of the gauge transformation. Automatically, using this derivative in the LD gives a gauge invariant Lagrangian with the Dirac field coupled to electromagnetism:

$$\mathfrak{L} = i\bar{\psi}\gamma^\mu\mathscr{D}_\mu\psi - m\bar{\psi}\psi = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi + qA_\mu\bar{\psi}\gamma^\mu\psi. \tag{3.5.238}$$

This modification of the derivative term is necessary to couple the Dirac theory to electromagnetism (and so it is called the minimal coupling) but other terms can in principle be added. For example, $\bar{\psi}S^{\mu\nu}\mathcal{F}_{\mu\nu}\psi$ has all of the necessary symmetries. As is generically the case in field theories, many higher order interaction terms are possible. But, equation (3.5.238) is the minimum necessary. This gives the interacting Dirac equation:

$$\frac{\partial\mathfrak{L}}{\partial\bar{\psi}} - \partial_\mu\frac{\partial\mathfrak{L}}{\partial\partial_\mu\bar{\psi}} = i\gamma^\mu\mathscr{D}_\mu\psi - m\psi \stackrel{!}{=} 0. \tag{3.5.239}$$

Taking the covariant derivative of this equation gives:

$$i\gamma^\mu\mathscr{D}_\mu i\gamma^\nu\mathscr{D}_\nu\psi = im\gamma^\mu\mathscr{D}_\mu\psi = m^2\psi \tag{3.5.240}$$
$$\implies -\gamma^\mu\gamma^\nu\mathscr{D}_\mu\mathscr{D}_\nu\psi = m^2\psi \tag{3.5.241}$$

$$\implies -\left(\frac{\gamma^\mu\gamma^\nu+\gamma^\nu\gamma^\mu}{2}+\frac{\gamma^\mu\gamma^\nu-\gamma^\nu\gamma^\mu}{2}\right)\mathscr{D}_\mu\mathscr{D}_\nu\psi=m^2\psi \tag{3.5.242}$$

$$\implies -\left(-\eta^{\mu\nu}\mathbb{1}-2i\hat{S}^{\mu\nu}\right)\mathscr{D}_\mu\mathscr{D}_\nu\psi=m^2\psi \tag{3.5.243}$$

$$\implies \mathscr{D}^2\psi+2i\hat{S}^{\mu\nu}\left(\frac{\mathscr{D}_\mu\mathscr{D}_\nu+\mathscr{D}_\nu\mathscr{D}_\mu}{2}+\frac{\mathscr{D}_\mu\mathscr{D}_\nu-\mathscr{D}_\nu\mathscr{D}_\mu}{2}\right)\psi=m^2\psi \tag{3.5.244}$$

$$\implies \mathscr{D}^2\psi+i\hat{S}^{\mu\nu}[\mathscr{D}_\mu,\mathscr{D}_\nu]\psi=m^2\psi. \tag{3.5.245}$$

Here we have used that the spin matrices are anti-symmetric and so the combination $\hat{S}^{\mu\nu}M_{\mu\nu}$ for any symmetric $M_{\mu\nu}$ is 0. The commutator of the covariant derivatives is given by:

$$[\mathscr{D}_\mu,\mathscr{D}_\nu]\psi=\mathscr{D}_\mu\mathscr{D}_\nu\psi-\mathscr{D}_\nu\mathscr{D}_\mu\psi \tag{3.5.246}$$

$$=\partial_\mu\mathscr{D}_\nu\psi-iqA_\mu\mathscr{D}_\nu\psi-\partial_\nu\mathscr{D}_\mu\psi+iqA_\nu\mathscr{D}_\mu\psi \tag{3.5.247}$$

$$=\partial_\mu(\partial_\nu\psi-iqA_\nu\psi)-iqA_\mu(\partial_\nu\psi-iqA_\nu\psi)-\partial_\nu(\partial_\mu\psi-iqA_\mu\psi)+iqA_\nu(\partial_\mu\psi-iqA_\mu\psi) \tag{3.5.248}$$

$$=\partial^2_{\mu\nu}\psi-iq\partial_\mu(A_\nu\psi)-iqA_\mu\partial_\nu\psi-q^2A_\mu A_\nu\psi-\partial^2_{\mu\nu}\psi+iq\partial_\nu(A_\mu\psi)+iqA_\nu\partial_\mu\psi+q^2A_\mu A_\nu\psi \tag{3.5.249}$$

$$=-iq\partial_\mu A_\nu-iqA_\nu\partial_\mu\psi-iqA_\mu\partial_\nu\psi+iq\partial_\nu A_\mu+iqA_\mu\partial_\nu\psi+iqA_\nu\partial_\mu\psi \tag{3.5.250}$$

$$=-iq\mathcal{F}_{\mu\nu}. \tag{3.5.251}$$

Thus, we have:

$$\mathscr{D}^2\psi+q\hat{S}^{\mu\nu}\mathcal{F}_{\mu\nu}\psi=m^2\psi. \tag{3.5.252}$$

So, in the case of interacting with electromagnetism the individual components of $\psi$ no longer decouple when one takes another derivative. However, the left and right handed parts of $\psi$ still do decouple as the spin matrices are all block diagonal. It is also interesting to look at the equation of motion for $\bar{\psi}$ as it turns out to be nicer than that of $\psi^\dagger$.

$$\frac{\partial\mathfrak{L}}{\partial\psi}-\partial_\mu\frac{\partial\mathfrak{L}}{\partial\partial_\mu\psi}=-m\bar{\psi}+qA_\mu\bar{\psi}\gamma^\mu-\partial_\mu(i\bar{\psi}\gamma^\mu)=-i\mathscr{D}^*_\mu\bar{\psi}\gamma^\mu-m\bar{\psi} \tag{3.5.253}$$

$$\implies i\mathscr{D}^*_\mu\bar{\psi}\gamma^\mu=-m\bar{\psi}. \tag{3.5.254}$$

By expanding the $A_\mu$ dependence of the Lagrangian density explicitly, we can compute the current density due to the Dirac spinor:

$$\mathfrak{L}=i\bar{\psi}\gamma^\mu\partial_\mu\psi-m\bar{\psi}\psi+q\bar{\psi}\gamma^\mu A_\mu\psi \tag{3.5.255}$$

$$\implies J^\mu=\frac{\partial\mathfrak{L}}{\partial A_\mu}=q\bar{\psi}\gamma^\mu\psi. \tag{3.5.256}$$

By expanding explicitly we see that $J^0=q\psi^\dagger\psi$ just as in the non-relativistic case but the expression for $J^a$ has changed. The continuity equation for $J^\mu$ can be checked directly from the Dirac equation to get that it is in fact satisfied:

$$\partial_\mu J^\mu=\partial_\mu(q\bar{\psi}\gamma^\mu\psi)=q\partial_\mu\bar{\psi}\gamma^\mu\psi+q\bar{\psi}\gamma^\mu\partial_\mu\psi \tag{3.5.257}$$

$$=q(\partial_\mu\bar{\psi}+iqA_\mu\bar{\psi}-iqA_\mu\bar{\psi})\gamma^\mu\psi+q\bar{\psi}\gamma^\mu(\partial_\mu\psi-iqA_\mu\psi+iqA_\mu\psi) \tag{3.5.258}$$

$$=q\mathscr{D}^*_\mu\bar{\psi}-iq^2A_\mu\bar{\psi}\gamma^\mu\psi+q\bar{\psi}\gamma^\mu\mathscr{D}_\mu\psi+iq^2A_\mu\bar{\psi}\gamma^\mu\psi \tag{3.5.259}$$

$$=q(im\bar{\psi})\gamma^\mu\psi+q\bar{\psi}\gamma^\mu(-im\psi)=0. \tag{3.5.260}$$

When coupled to electromagnetism, the Dirac theory is still in the Schrödinger form. We can find the explicit Hamiltonian operator:

$$\mathfrak{L}=i\psi^\dagger\dot{\psi}+i\psi^\dagger\gamma^0\gamma^a\partial_a\psi-m\psi^\dagger\gamma^0\psi+q\psi^\dagger A_0\psi+q\psi^\dagger\gamma^0\gamma^a A_a\psi \tag{3.5.261}$$

$$=i\psi^\dagger\dot{\psi}-\psi^\dagger(\gamma^0\gamma^a(\hat{p}_a-qA_a)+m\gamma^0+q\phi)\psi \tag{3.5.262}$$

$$\implies \hat{H}=\gamma^0\gamma^a(\hat{p}_a-qA_a(t,\hat{\vec{x}}))+m\gamma^0+q\phi(t,\hat{\vec{x}}). \tag{3.5.263}$$

Thus, we can define the wavefunction $|\psi(t)\rangle$ as usual by:

$$|\psi\rangle=\int_{\mathbb{S}}\psi(t,\vec{x})|\vec{x}\rangle d^3\vec{x} \tag{3.5.264}$$

and then we immediately have the usual Schrödinger equation:

$$i\frac{d|\psi\rangle}{dt} = \hat{H}|\psi\rangle \tag{3.5.265}$$

with the Hamiltonian operator identified. We can also naturally define $|\psi_L\rangle$ and $|\psi_R\rangle$ as the corresponding wavefunctions to the left and right handed components of $|\psi\rangle$:

$$|\psi\rangle = \begin{pmatrix} |\psi_L\rangle \\ |\psi_R\rangle \end{pmatrix}. \tag{3.5.266}$$

The Hamiltonian then gives the equations of motion for the left and right components:

$$i\begin{pmatrix} |\dot{\psi}_L\rangle \\ |\dot{\psi}_R\rangle \end{pmatrix} = \left(\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^a \\ -\sigma^a & 0 \end{pmatrix}(\hat{p}_a - q\hat{A}_a) + m\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} + q\hat{\phi}\begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}\right)\begin{pmatrix} |\psi_L\rangle \\ |\psi_R\rangle \end{pmatrix} \tag{3.5.267}$$

$$\implies i\begin{pmatrix} |\dot{\psi}_L\rangle \\ |\dot{\psi}_R\rangle \end{pmatrix} = \begin{pmatrix} -\sigma^a(\hat{p}_a - q\hat{A}_a)|\psi_L\rangle + m|\psi_R\rangle + q\hat{\phi}|\psi_L\rangle \\ \sigma^a(\hat{p}_a - q\hat{A}_a)|\psi_R\rangle + m|\psi_L\rangle + q\hat{\phi}|\psi_R\rangle \end{pmatrix}. \tag{3.5.268}$$

It is interesting that for the Galilean spin 1/2 particle there were only two degrees of freedom in the wavefunction but we see here in the Lorentz theory that there are four independent components because each of $|\psi_L\rangle$ and $|\psi_R\rangle$ have two components. Somehow then, in the non-relativistic limit we ought to be able to reduce these four degrees of freedom to only two. We know that (if we momentarily restore factors of $c$) the Lorentz Hamiltonian $\mathcal{H}_L$ is related to the Galilean Hamiltonian $\mathcal{H}_G$ for a mass $m$ by:

$$\mathcal{H}_L = mc^2 + \mathcal{H}_G + \mathcal{O}\left(\frac{1}{c^2}\right). \tag{3.5.269}$$

Consequently, if $\hat{H}_L$ is our Hamiltonian operator corresponding to our Lorentz quantum theory we expect there to be an operator $\hat{H}_0$ which doesn't depend on $c$ and which gives the non-relativistic time evolution so that:

$$\langle\psi|\hat{H}_L|\psi\rangle = mc^2\langle\psi|\psi\rangle + \langle\psi|\hat{H}_0|\psi\rangle + \mathcal{O}\left(\frac{1}{c^2}\right). \tag{3.5.270}$$

This gives the equation of motion:

$$i\frac{d|\psi\rangle}{dt} = mc^2|\psi\rangle + \hat{H}_0|\psi\rangle + \mathcal{O}\left(\frac{1}{c^2}\right) \tag{3.5.271}$$

$$\implies i\frac{d}{dt}\left(e^{imc^2t}|\psi\rangle\right) = \hat{H}_0 e^{imc^2t}|\psi\rangle + \mathcal{O}\left(\frac{1}{c^2}\right). \tag{3.5.272}$$

Hence, the relativistic shift in the Hamiltonian leads to a phase shift in the wavefunction so that the wavefunction which satisfies the non-relativistic equation of motion is:

$$|\psi_0\rangle = e^{imc^2t}|\psi\rangle \tag{3.5.273}$$

$$\implies i\frac{d|\psi_0\rangle}{dt} = \hat{H}_0|\psi_0\rangle + \text{Higher Order Terms.} \tag{3.5.274}$$

Again dropping factors of $c$ and defining $|\psi_{L0}\rangle$ and $|\psi_{R0}\rangle$ as the left and right components of $|\psi_0\rangle$, we see that they satisfy the equations of motion:

$$i|\dot{\psi}_{L0}\rangle = -\sigma^a(\hat{p}_a - q\hat{A}_a)|\psi_{L0}\rangle - m(|\psi_{L0}\rangle - |\psi_{R0}\rangle) + q\hat{\phi}|\psi_{L0}\rangle \tag{3.5.275}$$

$$i|\dot{\psi}_{R0}\rangle = \sigma^a(\hat{p}_a - q\hat{A}_a)|\psi_{R0}\rangle + m(|\psi_{L0}\rangle - |\psi_{R0}\rangle) + q\hat{\phi}|\psi_{R0}\rangle. \tag{3.5.276}$$

To understand the non-relativistic behavior of these equations, recall that we found the two component Galilean spin 1/2 spinor $|\psi_G\rangle$ under parity transformed as:

$$\hat{\mathcal{P}}|\psi_G\rangle = \int_{\mathbb{S}} |-\vec{x}\rangle\langle\vec{x}|d\mathbb{S}|\psi_G\rangle \tag{3.5.277}$$

$$\implies \psi_{G\mathcal{P}}(t,\vec{x}) = \psi_G(t,-\vec{x}). \tag{3.5.278}$$

So, the Galilean wavefunction is parity even. Out of the two component spinors $|\psi_{L0}\rangle$ and $|\psi_{R0}\rangle$ we can make a parity even and parity odd combination and so let us define:

$$|\psi_G\rangle = \frac{|\psi_{L0}\rangle + |\psi_{R0}\rangle}{\sqrt{2}}, \qquad |\psi_A\rangle = \frac{|\psi_{L0}\rangle - |\psi_{R0}\rangle}{\sqrt{2}} \tag{3.5.279}$$

where the $\sqrt{2}$ factors are chosen by convenience. Under parity, we see:

$$\psi_{G\mathcal{P}}(t,\vec{x}) = \psi_G(t,-\vec{x}), \qquad \psi_{A\mathcal{P}}(t,\vec{x}) = -\psi_A(t,-\vec{x}). \tag{3.5.280}$$

The function $\psi_G(t,\vec{x})$ has been named here suggestively based on its parity transformation as we hope to see it somehow become the Galilean wavefunction. Let us compute the equations of motion for $|\psi_G\rangle$ and $|\psi_A\rangle$:

$$i|\dot{\psi}_G\rangle = -\sigma^a(\hat{p}_a - q\hat{A}_a)\left(\frac{|\psi_{L0}\rangle - |\psi_{R0}\rangle}{\sqrt{2}}\right) + q\hat{\phi}\frac{|\psi_{L0}\rangle + |\psi_{R0}\rangle}{\sqrt{2}} \tag{3.5.281}$$

$$= -\sigma^a(\hat{p}_a - q\hat{A}_a)|\psi_A\rangle + q\hat{\phi}|\psi_G\rangle \tag{3.5.282}$$

$$i|\dot{\psi}_A\rangle = -\sigma^a(\hat{p}_a - q\hat{A}_a)\left(\frac{|\psi_{L0}\rangle + |\psi_{R0}\rangle}{\sqrt{2}}\right) - 2m\frac{|\psi_{L0}\rangle - |\psi_{R0}\rangle}{\sqrt{2}} + q\hat{\phi}\frac{|\psi_{L0}\rangle - |\psi_{R0}\rangle}{\sqrt{2}} \tag{3.5.283}$$

$$= -\sigma^a(\hat{p}_a - q\hat{A}_a)|\psi_G\rangle - 2m|\psi_A\rangle + q\hat{\phi}|\psi_A\rangle. \tag{3.5.284}$$

The equation of motion for $|\psi_A\rangle$ is equivalently:

$$\frac{i|\dot{\psi}_A\rangle - q\hat{\phi}|\psi_A\rangle}{2m} + |\psi_A\rangle = -\frac{\sigma^a(\hat{p}_a - q\hat{A}_a)}{2m}|\psi_G\rangle. \tag{3.5.285}$$

We know that for an eigenstate of the Hamiltonian operator $i|\dot{\psi}_A\rangle = E|\psi_a\rangle$ where $E$ is the corresponding energy eigenvalue. We can view the non-relativistic limit as when the energy $mc^2$ completely dominates over the $c$ independent part of the Hamiltonian. In $c=1$ units this means then that the energy scales involved are small compared to $m$. In particular, considering the relativistic single particle Hamiltonian:

$$\mathcal{H} = \sqrt{m^2 + |\vec{p} - q\vec{A}|^2} + q\phi = m + \mathcal{H}_G + \text{H.O.T.} \tag{3.5.286}$$

$$\implies \frac{\mathcal{H}_G - q\phi}{m} = \sqrt{1 + \frac{|\vec{p} - q\vec{A}|^2}{m^2}} - 1 + \text{H.O.T.} = \mathcal{O}\left(\frac{|\vec{p} - q\vec{A}|^2}{m^2}\right) = \mathcal{O}(|\vec{v}|^2). \tag{3.5.287}$$

If one restores factors of $c$ and considers the non-relativistic limit as the large $c$ limit, being of order $|\vec{v}|^2$ becomes being of order $\frac{|\vec{v}|^2}{c^2}$ or equivalently order $\frac{1}{c^2}$. Consequently, for $|\psi\rangle$ which are non-relativistic we have that:

$$\frac{i|\dot{\psi}_A\rangle - q\hat{\phi}|\psi_A\rangle}{2m} = \mathcal{O}\left(\frac{1}{c^2}\right). \tag{3.5.288}$$

On the right hand side of equation (3.5.285) we only have one power of $\frac{\vec{p}-q\vec{A}}{m}$ which is of order of the velocity and so the right hand side is order $\frac{1}{c}$. Thus, it dominates the $\frac{1}{c^2}$ term on the left hand side and to leading order in the non-relativistic approximation:

$$|\psi_A\rangle \approx -\frac{\sigma^a(\hat{p}_a - q\hat{A}_a)}{2m}|\psi_G\rangle. \tag{3.5.289}$$

Thus, we can see how there are only 2 independent components in the non-relativistic theory; it is because two components of the relativistic $|\psi\rangle$ can be computed from the other two in this limit. Then, the equation of motion for $|\psi_G\rangle$ becomes:

$$i|\dot{\psi}_G\rangle = -\sigma^a(\hat{p}_a - q\hat{A}_a)\left(-\frac{\sigma^b(\hat{p}_b - q\hat{A}_b)}{2m}\right)|\psi_G\rangle + q\hat{\phi}|\psi_G\rangle = \hat{H}_G|\psi_G\rangle \tag{3.5.290}$$

where:

$$\hat{H}_G = \frac{(\sigma^a(\hat{p}_a - q\hat{A}_a))^2}{2m} + q\hat{\phi} \tag{3.5.291}$$

$$= \frac{1}{2m}\left(\frac{\sigma^a\sigma^b + \sigma^b\sigma^a}{2} + \frac{\sigma^a\sigma^b - \sigma^b\sigma^a}{2}\right)(\hat{p}_a - q\hat{A}_a)(\hat{p}_b - q\hat{A}_b) + q\hat{\phi} \tag{3.5.292}$$

$$= \frac{1}{2m}\left(\delta^{ab}\mathbb{1} + 2i\epsilon^{ab}{}_c\hat{S}_c\right)(\hat{p}_a - q\hat{A}_a)(\hat{p}_b - a\hat{A}_b) + q\hat{\phi} \tag{3.5.293}$$

$$= \frac{|\hat{\vec{p}} - q\hat{\vec{A}}|^2}{2m} + \frac{i}{m}\epsilon^{ab}{}_c\hat{S}_c\frac{(\hat{p}_a - q\hat{A}_a)(\hat{p}_b - q\hat{A}_b) + (\hat{p}_b - q\hat{A}_b)(\hat{p}_a - q\hat{A}_a)}{2} \tag{3.5.294}$$

$$+\frac{i}{m}\epsilon^{ab}{}_c\hat{S}_c\frac{(\hat{p}_a-q\hat{A}_a)(\hat{p}_b-q\hat{A}_b)-(\hat{p}_b-q\hat{A}_b)(\hat{p}_a-q\hat{A}_a)}{2}+q\hat{\phi} \tag{3.5.295}$$

$$=\frac{|\hat{\vec{p}}-q\hat{\vec{A}}|^2}{2m}+0+\frac{i}{2m}\epsilon^{ab}{}_c\hat{S}^c[\hat{p}_a-q\hat{A}_a,\hat{p}_b-\hat{A}_b]+q\hat{\phi} \tag{3.5.296}$$

$$=\frac{|\hat{\vec{p}}-q\hat{\vec{A}}|^2}{2m}+\frac{i}{2m}\epsilon^{ab}{}_c\hat{S}^c(-i)^2[\hat{\mathscr{D}}_a,\hat{\mathscr{D}}_b]+q\hat{\phi} \tag{3.5.297}$$

$$=\frac{|\hat{\vec{p}}-q\hat{\vec{A}}|^2}{2m}-\frac{i}{2m}\epsilon^{ab}{}_c\hat{S}^c(-iq)\mathcal{F}_{ab}+q\hat{\phi} \tag{3.5.298}$$

$$=\frac{|\hat{\vec{p}}-q\hat{\vec{A}}|^2}{2m}-\frac{q}{2m}\epsilon^{ab}{}_c\hat{S}^c(\partial_aA_b-\partial_bA_a)+q\hat{\phi} \tag{3.5.299}$$

$$=\frac{|\hat{\vec{p}}-q\hat{\vec{A}}|^2}{2m}-\frac{q}{m}\epsilon^{ab}{}_c\hat{S}^c(\partial_aA_b)+q\hat{\phi} \tag{3.5.300}$$

$$=\frac{|\hat{\vec{p}}-q\hat{\vec{A}}|^2}{2m}-\frac{q}{m}\hat{S}^c\hat{B}_c+q\hat{\phi} \tag{3.5.301}$$

$$=\frac{|\hat{\vec{p}}-q\hat{\vec{A}}|^2}{2m}-\frac{q}{m}\hat{\vec{B}}\cdot\hat{\vec{S}}+q\hat{\phi}. \tag{3.5.302}$$

Thus, in the non-relativistic limit $|\psi_G\rangle$ evolves with the Spin 1/2 Hamiltonian we found in nonrelativistic quantum mechanics but with the particular value of $\gamma=\frac{q}{m}$ or equivalently a gyromagnetic ratio of $g=2$. Let us look now at the Lagrangian:

$$\mathcal{L}=\int_{\mathbb{S}}\left(i\bar{\psi}\gamma^\mu\partial_\mu\psi-m\bar{\psi}\psi+qA_\mu\bar{\psi}\gamma^\mu\psi\right)d\mathbb{S} \tag{3.5.303}$$

$$=\int_{\mathbb{S}}\left(i\psi^\dagger\dot{\psi}-\psi^\dagger(\gamma^0\gamma^a(-i\partial_a-q\hat{A}_a)+m\gamma^0+q\phi)\psi\right)d\mathbb{S} \tag{3.5.304}$$

$$=\int_{\mathbb{S}}\left(i\psi_L^\dagger\dot{\psi}_L+i\psi_R^\dagger\dot{\psi}_R+\psi_L^\dagger(\sigma^a(-i\partial_a-qA_a)-q\phi)\psi_L-\psi_R^\dagger(\sigma^a(-i\partial_a-qA_a)+q\phi)\psi_R-m\psi_L^\dagger\psi_R-m\psi_R^\dagger\psi_L\right)d\mathbb{S} \tag{3.5.305}$$

$$=i\langle\psi_L|\dot{\psi_L}\rangle+i\langle\psi_R|\dot{\psi_R}\rangle+\langle\psi_L|(\sigma^a(\hat{p}_a-q\hat{A}_a)-q\hat{\phi})|\psi_L\rangle-\langle\psi_R|(\sigma^a(\hat{p}_a-q\hat{A}_a)+q\hat{\phi})|\psi_R\rangle-m\langle\psi_L|\psi_R\rangle-m\langle\psi_R|\psi_L\rangle \tag{3.5.306}$$

$$\begin{aligned}=&ie^{imt}\langle\psi_{L0}|\frac{d}{dt}\left(e^{-imt}|\psi_{L0}\rangle\right)+ie^{imt}\langle\psi_{R0}|\frac{d}{dt}\left(e^{-imt}|\psi_{R0}\rangle\right)+\langle\psi_{L0}|(\sigma^a(\hat{p}_a-q\hat{A}_a)-q\hat{\phi})|\psi_{L0}\rangle\\&-\langle\psi_{R0}|(\sigma^a(\hat{p}_a-q\hat{A}_a)+q\hat{\phi})|\psi_{R0}\rangle-m\langle\psi_{L0}|\psi_{R0}\rangle-m\langle\psi_{R0}|\psi_{L0}\rangle\end{aligned} \tag{3.5.307}$$

$$\begin{aligned}=&i\langle\psi_{L0}|\dot{\psi_{L0}}\rangle+i\langle\psi_{R0}|\dot{\psi_{R0}}\rangle+\langle\psi_{L0}|(\sigma^a(\hat{p}_a-q\hat{A}_a)-q\hat{\phi}+m)|\psi_{L0}\rangle\\&-\langle\psi_{R0}|(\sigma^a(\hat{p}_a-q\hat{A}_a)+q\hat{\phi}-m)|\psi_{R0}\rangle-m\langle\psi_{L0}|\psi_{R0}\rangle-m\langle\psi_{R0}|\psi_{L0}\rangle\end{aligned} \tag{3.5.308}$$

$$\begin{aligned}=&i\langle\psi_G|\dot{\psi_G}\rangle-\langle\psi_G|q\hat{\phi}|\psi_G\rangle+i\langle\psi_A|\dot{\psi_A}\rangle-\langle\psi_A|q\hat{\phi}|\psi_A\rangle+2m\langle\psi_A|\psi_A\rangle\\&+\langle\psi_A|\sigma^a(\hat{p}_a-q\hat{A}_a)|\psi_G\rangle+\langle\psi_G|\sigma^a(\hat{p}_a-q\hat{A}_a)|\psi_A\rangle.\end{aligned} \tag{3.5.309}$$

This expression for the Lagrangian so far is exact. Now using the non-relativistic limit:

$$i\langle\psi_A|\dot{\psi_A}\rangle-\langle\psi_A|q\hat{\phi}|\psi_A\rangle+2m\langle\psi_A|\psi_A\rangle\approx 2m\langle\psi_A|\psi_A\rangle \tag{3.5.310}$$

and so:

$$\begin{aligned}\mathcal{L}\approx&i\langle\psi_G|\dot{\psi_G}\rangle-\langle\psi_G|q\hat{\phi}|\psi_G\rangle+2m\langle\psi_G|\left(-\frac{\sigma^a(\hat{p}_a-q\hat{A}_a)}{2m}\right)\left(-\frac{\sigma^b(\hat{p}_b-q\hat{A}_b)}{2m}\right)\psi_G\rangle\\&+\langle\psi_G|\left(-\frac{\sigma^b(\hat{p}_b-q\hat{A}_b)}{2m}\right)\sigma^a(\hat{p}_a-q\hat{A}_a)|\psi_G\rangle+\langle\psi_G|\sigma^a(\hat{p}_a-q\hat{A}_a)\left(-\frac{\sigma^b(\hat{p}_b-q\hat{A}_b)}{2m}\right)|\psi_G\rangle\end{aligned} \tag{3.5.311}$$

$$=i\langle\psi_G|\dot{\psi_G}\rangle-\langle\psi_G|\left(\frac{(\sigma^a(\hat{p}_a-q\hat{A}_a))^2}{2m}+q\hat{\phi}\right)|\psi_G\rangle \tag{3.5.312}$$

$$=i\langle\psi_G|\dot{\psi_G}\rangle-\langle\psi_G|\hat{H}_G|\psi_G\rangle. \tag{3.5.313}$$

Hence, in the non-relativistic limit $|\psi_G\rangle$ enters into the Lagrangian exactly like the non-relativistic wavefunction and thus we confirm our identification with $\psi_G$ as the non-relativistic wavefunction and $\hat{H}_G$ as the non-relativistic Hamiltonian operator. So, the Dirac equation specifically predicts a gyromagnetic ratio of $g = 2$ (before including higher order quantum effects).

We have established how the Dirac spinor transforms under parity and charge conjugation but we have not yet discussed our final discrete transformation of interest: time reversal. Our time reversed $\psi$ must be some linear combination of $\psi$ and $\psi^*$ in order to preserve the bilinear form of the Lagrangian density, just as with parity. And, just as with parity, time reversal does not reverse the U(1) transformation by definition and so it must actually only depend on $\psi$ and not $\psi^*$ (as $\psi^*$ has inverted U(1) transformation). Consequently, for some $4 \times 4$ matrix $\hat{\mathcal{T}}$:

$$\psi_{\mathcal{T}}(t, \vec{x}) = \hat{\mathcal{T}}\psi(-t, \vec{x}). \tag{3.5.314}$$

For a classical particular under time reversal, position by definition does not change:

$$\vec{x}_{\mathcal{T}}(t) = \vec{x}(-t). \tag{3.5.315}$$

As well, because the Lagrangian is the integrand of the action from an initial time to a later time, it is time reversal even (the minus sign of changing the direction of the time variable is canceled by the minus sign of swapping the integration limits due to their ordering changing). Consequently, velocity and linear momentum are inverted:

$$\dot{\vec{x}}_{\mathcal{T}} = \frac{d\vec{x}_{\mathcal{T}}}{dt} = \frac{d}{dt}\vec{x}(-t) = -\dot{\vec{x}}(-t), \qquad \vec{p}_{\mathcal{T}} = \frac{\partial \mathcal{L}}{\partial \dot{\vec{x}}_{\mathcal{T}}} = -\frac{\partial \mathcal{L}}{\partial \dot{\vec{x}}} = -\vec{p}, \tag{3.5.316}$$

the angular momentum is inverted:

$$\vec{L}_{\mathcal{T}} = \vec{x}_{\mathcal{T}} \times \vec{p}_{\mathcal{T}} = \vec{x} \times (-\vec{p}) = -\vec{x} \times \vec{p} = -\vec{L}, \tag{3.5.317}$$

the Hamiltonian is preserved:

$$\mathcal{H}_{\mathcal{T}} = \vec{p}_{\mathcal{T}} \cdot \dot{\vec{x}}_{\mathcal{T}} - \mathcal{L} = (-\vec{p}) \cdot (-\dot{\vec{x}}) - \mathcal{L} = \mathcal{H}, \tag{3.5.318}$$

and the generator of Lorentz boosts is preserved:

$$\vec{K}_{\mathcal{T}} = \vec{p}_{\mathcal{T}} t_{\mathcal{T}} - \mathcal{H}_{\mathcal{T}} \vec{x}_{\mathcal{T}} = (-\vec{p})(-t) - \mathcal{H}\vec{x} = \vec{K}. \tag{3.5.319}$$

Thus, because the angular momentum is inverted and the generator of boosts is preserved, we must have:

$$\hat{J}_{\mathcal{T}a} = -\hat{J}_a, \qquad \hat{K}_{\mathcal{T}a} = \hat{K}_a \tag{3.5.320}$$

so that the time reversed $\psi$ transforms as:

$$\psi'_{\mathcal{T}} = \hat{U}_{\mathcal{T}}\psi_{\mathcal{T}} = e^{-i\theta^a \hat{J}_{\mathcal{T}a} - i\varphi^a \hat{K}_{\mathcal{T}a}}\psi_{\mathcal{T}} = e^{i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}\hat{\mathcal{T}}\psi. \tag{3.5.321}$$

We can compute directly how $\psi_{\mathcal{T}}$ transforms as well (recalling that time reversal takes the complex conjugate of any coefficients present):

$$\psi'_{\mathcal{T}} = \mathcal{T}(\psi') = \mathcal{T}(U\psi) = \mathcal{T}\left(e^{-i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}\psi\right) = \hat{\mathcal{T}}e^{i\theta^a \hat{J}_a^* + i\varphi^a \hat{K}_a^*}\psi \tag{3.5.322}$$

Matching these two:

$$e^{i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}\hat{\mathcal{T}}\psi = \hat{\mathcal{T}}e^{i\theta^a \hat{J}_a^* + i\varphi^a \hat{K}_a^*}\psi, \qquad \forall\psi \tag{3.5.323}$$

$$\implies e^{i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}\hat{\mathcal{T}} = \hat{\mathcal{T}}e^{i\theta^a \hat{J}_a^* + i\varphi^a \hat{K}_a^*}. \tag{3.5.324}$$

Recall that $\sigma_1$ and $\sigma_3$ contain only real entries while $\sigma_2$ is purely imaginary. Therefore:

$$\sigma_2\sigma_1\sigma_2 = -\sigma_1\sigma_2\sigma_2 = -\sigma_1 = -\sigma_1^* \tag{3.5.325}$$

$$\sigma_2\sigma_2\sigma_2 = \sigma_2 = -\sigma_2^* \tag{3.5.326}$$

$$\sigma_2\sigma_3\sigma_2 = -\sigma_3\sigma_2\sigma_2 = -\sigma_3 = -\sigma_3^* \tag{3.5.327}$$

$$\implies \sigma_a^* = -\sigma_2\sigma_a\sigma_2. \tag{3.5.328}$$

Thus:

$$\hat{J}_a^* = \begin{pmatrix} \frac{\sigma_a^*}{2} & 0 \\ 0 & \frac{\sigma_a^*}{2} \end{pmatrix} = \begin{pmatrix} -\sigma_2\frac{\sigma_a}{2}\sigma_2 & 0 \\ 0 & -\sigma_2\frac{\sigma_a}{2}\sigma_2 \end{pmatrix} = \begin{pmatrix} i\sigma_2 & 0 \\ 0 & i\sigma_2 \end{pmatrix} \hat{J}_a \begin{pmatrix} -i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{pmatrix}. \tag{3.5.329}$$

Then, let:

$$A = -\gamma^1\gamma^3 = \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix} = \begin{pmatrix} -i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{pmatrix} \tag{3.5.330}$$

$$\implies A^\dagger A = \begin{pmatrix} i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{pmatrix}\begin{pmatrix} -i\sigma_2 & 0 \\ 0 & i\sigma_2 \end{pmatrix} = \mathbb{1}. \tag{3.5.331}$$

Consequently:

$$\hat{J}_a^* = A^\dagger \hat{J}_a A. \tag{3.5.332}$$

As well:

$$\hat{K}_a^* = \begin{pmatrix} i\frac{\sigma_a^*}{2} & 0 \\ 0 & -i\frac{\sigma_a^*}{2} \end{pmatrix} = \begin{pmatrix} -\frac{i}{2}\sigma_2\sigma_a\sigma_2 & 0 \\ 0 & \frac{i}{2}\sigma_2\sigma_a\sigma_2 \end{pmatrix} = -\begin{pmatrix} i\sigma_2 & 0 \\ 0 & i\sigma_2 \end{pmatrix}\hat{K}_a\begin{pmatrix} -i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{pmatrix} = -A^\dagger \hat{K}_a A. \tag{3.5.333}$$

For any matrix $M$:

$$(A^\dagger M A)^n = A^\dagger M A A^\dagger M A ... A^\dagger M A = A^\dagger M \mathbb{1} M \mathbb{1} ... M A = A^\dagger M^n A. \tag{3.5.334}$$

Consequently, for any $f(M)$ analytic near 0:

$$f(A^\dagger M A) = \sum_{n=0}^{\infty} c_n (A^\dagger M A)^n = \sum_{n=0}^{\infty} c_n A^\dagger M^n A = A^\dagger f(M) A. \tag{3.5.335}$$

Then we have:

$$e^{i\theta^a \hat{J}_a^* + i\varphi^a \hat{K}_a^*} = e^{A^\dagger (i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a) A} = A^\dagger e^{i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a} A. \tag{3.5.336}$$

With this equation (3.5.324) becomes:

$$e^{i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}\hat{\mathcal{T}} = \hat{\mathcal{T}} A^\dagger e^{i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a} A \tag{3.5.337}$$

$$\implies e^{i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}\hat{\mathcal{T}} A^\dagger = \hat{\mathcal{T}} A^\dagger e^{i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}. \tag{3.5.338}$$

This is true for all $\theta$ and $\varphi$ and so we can take $\theta$ to $-\theta$:

$$e^{-i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a}\hat{\mathcal{T}} A^\dagger = \hat{\mathcal{T}} A^\dagger e^{-i\theta^a \hat{J}_a - i\varphi^a \hat{K}_a} \tag{3.5.339}$$

$$\implies U\hat{\mathcal{T}} A^\dagger = \hat{\mathcal{T}} A^\dagger U \tag{3.5.340}$$

$$\implies [U, \hat{\mathcal{T}} A^\dagger] = 0. \tag{3.5.341}$$

Then, by equations (3.5.68) and (3.5.77) for some coefficients $a_1$ and $a_2$:

$$\hat{\mathcal{T}} A^\dagger = \begin{pmatrix} (a_1 - a_2)\mathbb{1} & 0 \\ 0 & (a_1 + a_2)\mathbb{1} \end{pmatrix} = a_1 \mathbb{1} + a_2 \gamma^5 \tag{3.5.342}$$

$$\implies \hat{\mathcal{T}} = (a_1 \mathbb{1} + a_2 \gamma^5) A = -a_1 \gamma^1 \gamma^3 - a_2 \gamma^5 \gamma^1 \gamma^3. \tag{3.5.343}$$

Just as parity$^2$ is always a symmetry, time reversal squared is always a symmetry as doubly inverting the direction of motion leaves no inversion at all. So, our Lagrangian density must have symmetry under time reversal squared. Under time reversal squared, the two complex conjugations cancel out on any complex number and the time derivative automatically returns to its original self. So, under time reversal squared we have:

$$\mathfrak{L}_{\mathcal{T}^2} = i\bar{\psi}_{\mathcal{T}^2}\gamma^\mu \partial_\mu \psi_{\mathcal{T}^2} - m\bar{\psi}_{\mathcal{T}^2}\psi_{\mathcal{T}^2} \overset{!}{=} i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi. \tag{3.5.344}$$

Matching the $\dot{\psi}$ term, this requires:

$$\bar{\psi}_{\mathcal{T}^2}\gamma^0 \dot{\psi}_{\mathcal{T}^2} = \bar{\psi}\gamma^0\dot{\psi} \implies \psi^\dagger_{\mathcal{T}^2}\dot{\psi}_{\mathcal{T}^2} = \psi^\dagger \dot{\psi}. \tag{3.5.345}$$

From our expression for $\mathcal{T}$:

$$\psi_{\mathcal{T}^2} = \mathcal{T}(\mathcal{T}(\psi)) = (-a_1\mathbb{1} - a_2\gamma^5)\gamma^1\gamma^3(-a_1^*\mathbb{1} - a_2^*\gamma^5)\gamma^1\gamma^3\psi \tag{3.5.346}$$

$$= (a_1\mathbb{1} + a_2\gamma^5)(a_1^*\mathbb{1} + a_2^*\gamma^5)\gamma^1\gamma^3\gamma^1\gamma^3\psi = ((|a_1|^2 + |a_2|^2)\mathbb{1} + (a_1^*a_2 + a_2^*a_1)\gamma^5)\gamma^1(-\gamma^1\gamma^3)\gamma^3\psi \tag{3.5.347}$$

$$= -((|a_1|^2 + |a_2|^2)\mathbb{1} + (a_1^*a_2 + a_2^*a_1)\gamma^5)\psi. \tag{3.5.348}$$

Then, (3.5.345) requires:

$$\psi^\dagger_{\mathcal{T}^2}\dot{\psi}_{\mathcal{T}^2} = \psi^\dagger((|a_1|^2 + |a_2|^2)\mathbb{1} + (a_1^*a_2 + a_2^*a_1)\gamma^5)((|a_1|^2 + |a_2|^2)\mathbb{1} + (a_1^*a_2 + a_2^*a_1)\gamma^5)\dot{\psi} \tag{3.5.349}$$

$$= \psi^\dagger(((|a_1|^2 + |a_2|^2)^2 + (a_1^*a_2 + a_2^*a_1)^2)\mathbb{1} + (|a_1|^2 + |a_2|^2)(a_1^*a_2 + a_2^*a_1)\gamma^5)\dot{\psi} \tag{3.5.350}$$

$$\overset{!}{=} \psi^\dagger\dot{\psi} \tag{3.5.351}$$

$$\implies (|a_1|^2 + |a_2|^2)(a_1^*a_2 + a_2^*a_1) = 0, \qquad (|a_1|^2 + |a_2|^2)^2 + (a_1^*a_2 + a_2^*a_1)^2 = 1. \tag{3.5.352}$$

Thus, we need either $|a_1|^2 + |a_2|^2$ or $a_1^* a_2 + a_2^* a_1$ to vanish. If the first of these is zero then immediately $a_1 = a_2 = 0$ and the second condition cannot be satisfied. So, we must have:

$$a_1^* a_2 + a_2^* a_1 = 0. \tag{3.5.353}$$

Writing $a_1$ and $a_2$ in terms of magnitude and phase:

$$a_1 = r_1 e^{i\alpha_1}, \qquad a_2 = r_2 e^{i\alpha_2} \tag{3.5.354}$$

$$a_1^* a_2 + a_2^* a_1 = r_1 r_2 \left( e^{-i(\alpha_1 - \alpha_2)} + e^{i(\alpha_1 - \alpha_2)} \right) = 2 r_1 r_2 \cos(\alpha_1 - \alpha_2) = 0 \tag{3.5.355}$$

$$\implies \alpha_2 = \alpha_1 \pm \frac{\pi}{2}. \tag{3.5.356}$$

One may worry about the alternate possibilities or $r_1 = 0$ or $r_2 = 0$ but if either of these is true then the corresponding $\alpha_1$ or $\alpha_2$ is arbitrary and so can be set so that (3.5.355) is true. Thus, we can take (3.5.355) as guaranteed. Then, we have:

$$a_2 = r_2 e^{i(\alpha_1 \pm \frac{\pi}{2})} = \pm i r_2 e^{i\alpha_1}. \tag{3.5.357}$$

The second condition in (3.5.352) then becomes:

$$|a_1|^2 + |a_2|^2 = r_1^2 + r_2^2 = 1 \tag{3.5.358}$$

and so we can identify a $\beta$ so that:

$$r_1 = \cos\beta, \qquad r_2 = \pm \sin\beta \tag{3.5.359}$$

where we fix the sign so that:

$$a_1 = e^{i\alpha_1} \cos\beta, \qquad a_2 = i e^{i\alpha_1} \sin\beta. \tag{3.5.360}$$

The factor of $e^{\alpha_1}$ on both coefficients then gives in (3.5.343) an overall phase on the definition of the time reversal transformation. Just as we did with parity, we can use the U(1) symmetry of our theory to adjust the definition of time reversal by this phase so that by definition $\alpha_1 = 0$. So, for some parameter $\beta$ we have:

$$\psi_{\mathcal{T}} = -(\cos\beta \mathbb{1} + i \sin\beta \gamma^5) \gamma^1 \gamma^3 \psi. \tag{3.5.361}$$

With this, we simply have:

$$\psi_{\mathcal{T}^2} = -\psi \tag{3.5.362}$$

and so the remaining conditions in (3.5.344) are automatically satisfied. With our given expression for time reversal:

$$\bar{\psi}_{\mathcal{T}} = \psi_{\mathcal{T}}^\dagger \gamma^0 = -\psi^\dagger \gamma^3 \gamma^1 (\cos\beta \mathbb{1} - i \sin\beta \gamma^5) \gamma^0 = -\bar{\psi} \gamma^3 \gamma^1 (\cos\beta \mathbb{1} + i \sin\beta \gamma^5) \tag{3.5.363}$$

and so the Lagrangian density becomes:

$$\begin{aligned}
\mathfrak{L}_{\mathcal{T}} =& - i \bar{\psi}_{\mathcal{T}} (\gamma^\mu)^* \partial_{\mathcal{T}\mu} \psi_{\mathcal{T}} - m \bar{\psi}_{\mathcal{T}} \psi_{\mathcal{T}} && (3.5.364)\\
=& - i \bar{\psi} \gamma^3 \gamma^1 (\cos\beta \mathbb{1} + i \sin\beta \gamma^5) \gamma^2 \gamma^\mu \gamma^2 (\cos\beta \mathbb{1} + i \sin\beta \gamma^5) \gamma^1 \gamma^3 \partial_{\mathcal{T}\mu} \psi \\
& - m \bar{\psi} \gamma^3 \gamma^1 (\cos\beta \mathbb{1} + i \sin\beta \gamma^5)(\cos\beta \mathbb{1} + i \sin\beta \gamma^5) \gamma^1 \gamma^3 \psi && (3.5.365)\\
=& - i \bar{\psi} \gamma^3 \gamma^1 \gamma^2 \gamma^\mu \gamma^2 (\cos\beta \mathbb{1} - i \sin\beta \gamma^5)(\cos\beta \mathbb{1} + i \sin\beta \gamma^5) \gamma^1 \gamma^3 \partial_{\mathcal{T}\mu} \psi \\
& - m \bar{\psi} \gamma^3 \gamma^1 ((\cos^2\beta - \sin^2\beta) \mathbb{1} + 2 i \sin\beta \cos\beta \gamma^5) \gamma^1 \gamma^3 \psi && (3.5.366)\\
=& - i \bar{\psi} \gamma^3 \gamma^1 \gamma^2 \gamma^\mu \gamma^2 \gamma^1 \gamma^3 \partial_{\mathcal{T}\mu} \psi - m \bar{\psi} \gamma^3 \gamma^1 \gamma^1 \gamma^3 (\cos(2\beta) \mathbb{1} + i \sin(2\beta) \gamma^5) \psi && (3.5.367)\\
=& i \bar{\psi} \gamma^1 \gamma^2 \gamma^3 \gamma^\mu \gamma^1 \gamma^2 \gamma^3 \partial_{\mathcal{T}\mu} \psi - m \bar{\psi} (\cos(2\beta) \mathbb{1} + i \sin(2\beta) \gamma^5) \psi && (3.5.368)\\
=& i \bar{\psi} \frac{1}{i} \gamma^0 i \gamma^0 \gamma^1 \gamma^2 \gamma^3 \gamma^\mu \frac{1}{i} \gamma^0 i \gamma^0 \gamma^1 \gamma^2 \gamma^3 \partial_{\mathcal{T}\mu} \psi - m \bar{\psi} (\cos(2\beta) \mathbb{1} + i \sin(2\beta) \gamma^5) \psi && (3.5.369)\\
=& - i \bar{\psi} \gamma^0 \gamma^5 \gamma^\mu \gamma^0 \gamma^5 \partial_{\mathcal{T}\mu} \psi - m \bar{\psi} (\cos(2\beta) \mathbb{1} + i \sin(2\beta) \gamma^5) \psi && (3.5.370)\\
=& - i \bar{\psi} \gamma^0 \gamma^\mu \gamma^0 \gamma^5 \gamma^5 \partial_{\mathcal{T}\mu} \psi - m \bar{\psi} (\cos(2\beta) \mathbb{1} + i \sin(2\beta) \gamma^5) \psi && (3.5.371)\\
=& - i \bar{\psi} \left(\gamma^\mu\right)^\dagger \partial_{\mathcal{T}\mu} \psi - m \bar{\psi} (\cos(2\beta) \mathbb{1} + i \sin(2\beta) \gamma^5) \psi && (3.5.372)\\
=& - i \bar{\psi} \left(\gamma^0\right)^\dagger \partial_{\mathcal{T}0} \psi - i \bar{\psi} \left(\gamma^a\right)^\dagger \partial_{\mathcal{T}a} \psi - m \bar{\psi} (\cos(2\beta) \mathbb{1} + i \sin(2\beta) \gamma^5) \psi && (3.5.373)\\
=& - i \bar{\psi} \gamma^0 (-\partial_0) \psi - i \bar{\psi} \left(-\gamma^a\right) \partial_a \psi - m \bar{\psi} (\cos(2\beta) \mathbb{1} + i \sin(2\beta) \gamma^5) \psi && (3.5.374)\\
=& i \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} (\cos(2\beta) \mathbb{1} + i \sin(2\beta) \gamma^5) \psi. && (3.5.375)
\end{aligned}$$

We have no further requirements for our time reversal transformation to satisfy which are not already satisfied for general $\beta$. So, we can choose $\beta$ at will. We see here that if we choose $\beta = 0$ then our Lagrangian density will have symmetry under time reversal. Therefore, we take by definition:

$$\psi_{\mathcal{T}}(t, \vec{x}) = \gamma^1\gamma^3\psi^*(-t, \vec{x}). \tag{3.5.376}$$

Under this time reversal transformation essentially repeating the steps above one finds that the current density in fact has the correct time reversal property described.

The only remaining discrete transformation of interest to study is charge conjugation. We identified the expression for the charge conjugation transformation of $\psi$ (and one can check by similar analysis to what we have done thus far that what is given is up to a phase the unique U(1) reversing transformation which preserves the Lorentz transformation and thus the only candidate for the charge conjugation transformation). Under charge conjugation we expect the current density to become its own negative, as we found in the scalar case. Directly computing:

$$J^\mu_{\mathcal{C}} = q\bar{\psi}_{\mathcal{C}}\gamma^\mu\psi_{\mathcal{C}} = q(-i\psi^T)\gamma^2\gamma^0\gamma^\mu(-i\gamma^2)\psi^* = -q\psi^T\gamma^2\gamma^0\gamma^\mu\gamma^0\gamma^0\gamma^2\psi^* \tag{3.5.377}$$

$$= q\psi^T\gamma^2\left(\gamma^\mu\right)^\dagger\gamma^2\gamma^0\psi^* = q\psi^T\left(-\gamma^2\right)^\dagger\left(\gamma^\mu\right)^\dagger\left(-\gamma^2\right)^\dagger\bar{\psi}^T = q\psi^T\left(\gamma^2\gamma^\mu\gamma^2\right)^\dagger\bar{\psi}^T \tag{3.5.378}$$

$$= q\psi^T\left(\left(\gamma^\mu\right)^*\right)^\dagger\bar{\psi}^T = q\psi^T\left(\gamma^\mu\right)^T\bar{\psi}^T. \tag{3.5.379}$$

The transpose of a complex number is simply itself and so transposing this whole result we would find:

$$J^\mu_{\mathcal{C}} = q\bar{\psi}\gamma^\mu\psi = J^\mu \tag{3.5.380}$$

which is **NOT** the result we expected. This is because so far this section we have not paid attention to the commuting/anticommuting property of the Dirac field. As a fermion, it needs to be anticommuting. This minus sign will give the current density the correct charge conjugation property.

Not only can the Dirac theory be coupled to electromagnetism, it can also be coupled easily to a real scalar field by what is called a Yukawa interaction:

$$\mathfrak{L}_{\text{Yukawa}} = -\lambda\phi\bar{\psi}\psi \tag{3.5.381}$$

for some constant $\lambda$. Because $\phi$ and $\bar{\psi}\psi$ are already Lorentz invariant this term is completely allowed. This leads to the Lagrangian density:

$$\mathfrak{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{M^2}{2}\phi^2 - \lambda\phi\bar{\psi}\psi. \tag{3.5.382}$$

Of course, this is only the lowest order possible coupling between these two theories and many higher order terms are possible.

## Exercises

**Problem 3.10: helicity, chirality, weyl spinor homework question.**

(a)

**Problem 3.11: show the problems with not using anti-unitarity of time reversal for scalar and dirac equation.**

(a)

**Problem 3.12: show correct non-relativistic expression for the current density as a homework question for dirac equation.**

(a)

**Problem 3.13: explanation of why psi and phi are not conjugated under time reversal using real and imaginary parts as coordinates and momenta.**

(a)

**Problem 3.14: creation annihilation operators.**

(a)

**Problem 3.15: hamiltonian for this theory.**

(a)

**Problem 3.16: solve equations of motion/diagonalize hamiltonian.**

(a)

**Problem 3.17: energy momentum tensor for dirac.**

(a)

**Problem 3.18: charge conjugation, time reversal, and parity in the non-relativistic limit for psiG.**
(a)
**Problem 3.19: solution of the dirac equation and components.**
(a)

## 3.6 Scattering Theory

### 3.6.1 Schwinger Dyson Equation

### 3.6.2 LSZ Reduction Formula

### 3.6.3 Feynman Diagrams

### 3.6.4 Cross Sections and Decay Rates

## 3.7 Regularization

## 3.8 Renormalization

## 3.9 Symmetries in Quantum Field Theory

## 3.10 Summary and Additional Reading

# Chapter 4

# Gauge Theory

*The world is a dynamic mess of jiggling things.*
– Richard Feynman [19]

*From time immemorial, man has desired to comprehend the complexity of nature in terms of as few elementary concepts as possible. Among his quests - in Feynman's words - has been the one for "wheels within wheels" - the task of natural philosophy being to discover the innermost wheels if any such exist. A second quest has concerned itself with the fundamental forces which make the wheels go round and enmesh with one another. The greatness of gauge ideas - of gauge field theories - is that they reduce these two quests to just one; elementary particles (described by relativistic quantum fields) are representations of certain charge operators, corresponding to gravitational mass, spin, flavor, color, electric charge and the like, while the fundamental forces are the forces of attraction or repulsion between these same charges. A third quest seeks for a unification between the charges (and thus of the forces) by searching for a single entity, of which the various charges are components in the sense that they can be transformed one into the other.*
– Abdus Salam [52]

## 4.1 Quantum Electrodynamics in General

### 4.1.1 Classical Limit

### 4.1.2 Gauge Invariance

### 4.1.3 Path Integral and LSZ Reduction for Photons

### 4.1.4 Ward Identities

## 4.2 Scalar Quantum Electrodynamics

## 4.3 Spinor Quantum Electrodynamics

## 4.4 Geometry of Nonabelian Gauge Theory

### 4.4.1 Geometry of Gauge Transformations

We can extend and generalize our geometric description of the connection and gauge covariant derivative given for the electromagnetic vector potential quite significantly. Suppose we have some field $\Phi(X)$ which at each point in spacetime is an element of an $N$ dimensional internal vector space $\mathbb{V}(X)$. Then, let $\xi_a(X)$ be a

list of $N$ basis vectors for that space so that for some coefficients $\phi^a(X)$ we have:

$$\Phi(X) = \phi^a(X)\xi_a(X). \tag{4.4.1}$$

Suppose this internal space has some symmetry group $G$, meaning a group of change of basis transformations of $\mathbb{V}(X)$ so that all of the predictions of our theory are unchanged under them. The lowest order non-trivial Hamiltonian or Lagrangian we can make in a theory is quadratic in its variables. So, our Lagrangian density will generally depend at lowest order on some quadratic combination of the components $\phi^a(X)$ given above or in the case they are complex some combination of $\phi^a(X)$ and $(\phi^a(X))^*$ which contains a total of two powers between them. So, it will depend bilinearly on either $\phi$ and another copy of $\phi$ or $\phi$ and $\phi^*$ or $\phi^*$ and another copy of $\phi^*$. Then, to preserve this bilinearity under the symmetry transformation, the symmetry transformations must be linear and so the gauge group $G$ acting on the components is given by some matrices with components $U^a{}_b$ so that:

$$\phi'^a(X) = U^a{}_b(X)\phi^b(X). \tag{4.4.2}$$

As well, the bilinear form acting on two elements of $\mathbb{V}(X)$ with components $\phi_1^a$ and $\phi_2^a$ must be given by some coefficients $\mathrm{M}_{ab}$ as:

$$(\Phi_1, \Phi_2) = \mathrm{M}_{ab}(\phi_1^a)^*\phi_2^b \tag{4.4.3}$$

where $(\cdot,\cdot)$ indicates the bilinear form (the complex conjugation is introduced by definition into the first argument, if one would not like complex conjugation one can simply consider $(\Phi_1^*, \Phi_2)$ instead; technically with this definition it is conjugate linear in the first argument and so is said to be sesquilinear overall). The coefficients $\mathrm{M}_{ab}$ cannot depend on $X$ because as described we will want such a term to appear in a translationally invariant Lagrangian density. Using the sesquilinearity property:

$$(\Phi_1, \Phi_2) = (\phi_1^a\xi_a, \phi_2^b\xi_b) = (\phi_1^a)^*\phi_2^b(\xi_a, \xi_b). \tag{4.4.4}$$

Because this must match (4.4.3) for all $\Phi_1$ and $\Phi_2$, we have:

$$(\xi_a, \xi_b) = \mathrm{M}_{ab}. \tag{4.4.5}$$

This matrix M with components $\mathrm{M}_{ab}$ is called the semimetric of the internal space (if it is additionally self-adjoint then it is called a pseudometric and if it is positive definite it is just called a metric). As well, the sesquilinear form above is called the semi-inner product (if it is additionally conjugate symmetric in its arguments then it is called a pseudo-inner product and if it is positive definite it is just called an inner product). As with any matrix, we can decompose M into the 0 operator on its kernel and an invertible operator on its image. Then, we can identify a basis of the internal space which is the direct sum of basis vectors of the kernel and basis vectors of the image. The components of any $\Phi$ which correspond to basis vectors in the kernel then do not contribute to the sesquilinear form and thus do not contribute to the Lagrangian density or equations of motion. As such, those components essentially do not exist in the theory as they do not affect the dynamics. So, we can without consequence take the components of $\Phi$ on the kernel of M to be 0 and thus we are left with an effective M on the image of the original M which is invertible. Going forward, we will assume we have performed this construction so that M is has an empty kernel (aside from the 0 vector) and thus is automatically invertible.

From (4.4.2), we can see how to view the transformations in $G$ as a change of basis of $\mathbb{V}(X)$. The vector $\Phi(X)$ in $\mathbb{V}(X)$ can be expressed in whatever basis we choose and so let us use bases $\xi_a(X)$ so that it has components $\phi^a(X)$ and $\xi'_a(X)$ so that it has components $\phi'^a(X)$ given by (4.4.2):

$$\Phi(X) = \phi^a(X)\xi_a(X) = \phi'^a(X)\xi'_a(X) \tag{4.4.6}$$

$$\implies \phi^a\xi_a = U^a{}_b\phi^b\xi'_a \tag{4.4.7}$$

$$\implies \phi^a(\xi_a - U^b{}_a\xi'_b) = 0. \tag{4.4.8}$$

Because this must hold for any components $\phi^a$, we have:

$$\xi_a = U^b{}_a\xi'_b \tag{4.4.9}$$

$$\implies \xi'_a(X) = U^{-b}{}_a(X)\xi_b(X). \tag{4.4.10}$$

Thus, we can view these symmetry transformations as a change of basis. It will also be useful to sometimes

work with $\phi$, defined as the column vector with entries given by the $\phi^a$:

$$\phi \stackrel{\text{def}}{=} \begin{pmatrix} \phi^1 \\ \phi^2 \\ \dots \\ \phi^N \end{pmatrix}. \tag{4.4.11}$$

Then, with $U$ as the matrix with components ${U^a}_b$ (4.4.2) becomes:

$$\phi' = U\phi. \tag{4.4.12}$$

Any quantity transforming in this way under the gauge transformations is said to be in the "fundamental representation of the gauge group". Because these transformations are to be the symmetry transformations of our Lagrangian density and thus our semimetric, we must have that expressed in either the primed or the unprimed basis the result of the semi-inner product must be unchanged:

$$(\Phi_1, \Phi_2) = \mathrm{M}_{ab}\phi_1^{a*}\phi_2^b = \mathrm{M}_{ab}\phi_1'^{a*}\phi_2'^b \tag{4.4.13}$$

$$\implies \mathrm{M}_{ab}\phi_1^{a*}\phi_2^b = \mathrm{M}_{cd}\left(U^c{}_a\phi_1^a\right)^*\left(U^d{}_b\phi_2^b\right) \tag{4.4.14}$$

$$\implies \left(\mathrm{M}_{ab} - \mathrm{M}_{cd}\left(U^c{}_a\right)^* U^d{}_b\right)\phi_1^{a*}\phi_2^b = 0. \tag{4.4.15}$$

Because this must hold for all $\phi_1$ and $\phi_2$:

$$\mathrm{M}_{ab} = \mathrm{M}_{cd}\left(U^c{}_a\right)^* U^d{}_b. \tag{4.4.16}$$

Considering M as the matrix with components $\mathrm{M}_{ab}$, this is equivalent to:

$$\mathrm{M} = U^\dagger \mathrm{M} U. \tag{4.4.17}$$

Thus, the symmetry group $G$ is the group of matrices $U$ satisfying this condition for a given semimetric.

In order to define a derivative of the field $\Phi$, we must have a parallel transport transformation just like when we considered the electromagnetic connection. So, suppose there is some map $\Upsilon(\Phi, X, X')$ (for $X$ and $X'$ which are sufficiently near each other) which acts on $\Phi \in \mathbb{V}(X)$ and returns a vector in $\mathbb{V}(X')$ in a way which is both linear and preserves the semi-inner product:

$$\Upsilon(a\Phi_1 + b\Phi_2, X, X') = a\Upsilon(\Phi_1, X, X') + b\Upsilon(\Phi_2, X, X') \tag{4.4.18}$$

$$(\Upsilon(\Phi_1, X, X'), \Upsilon(\Phi_2, X, X')) = (\Phi_1, \Phi_2) \tag{4.4.19}$$

and which is continuous in its dependence on $X$ and $X'$ so that limits can be used. Then, we can define the covariant derivative exactly as before:

$$\mathscr{D}_\mu\Phi \stackrel{\text{def}}{=} \lim_{\varepsilon\to 0}\frac{\Upsilon(\Phi(X+\varepsilon\vec{\imath}_\mu), X+\varepsilon\vec{\imath}_\mu, X) - \Phi(X)}{\varepsilon}. \tag{4.4.20}$$

Using this definition of the covariant derivative in general, we can reduce it to being determined by the covariant derivatives of the basis vectors:

$$\mathscr{D}_\mu\Phi = \lim_{\varepsilon\to 0}\frac{\Upsilon(\phi^a(X+\varepsilon\vec{\imath}_\mu)\xi_a(X+\varepsilon\vec{\imath}_\mu), X+\varepsilon\vec{\imath}_\mu, X) - \phi^a(X)\xi_a(X)}{\varepsilon}. \tag{4.4.21}$$

$$= \lim_{\varepsilon\to 0}\left(\frac{\phi^a(X+\varepsilon\vec{\imath}_\mu) - \phi^a(X)}{\varepsilon}\Upsilon(\xi_a(X+\varepsilon\vec{\imath}_\mu), X+\varepsilon\vec{\imath}_\mu, X) + \phi^a(X)\frac{\Upsilon(\xi_a(X+\varepsilon\vec{\imath}_\mu), X+\varepsilon\vec{\imath}_\mu, X) - \xi_a(X)}{\varepsilon}\right) \tag{4.4.22}$$

$$= (\partial_\mu\phi^a)\xi_a(X) + \phi^a(X)\mathscr{D}_\mu\xi_a. \tag{4.4.23}$$

Because $\mathscr{D}_\mu\xi_a$ is defined as a difference quotient of vectors in the vector space at $X$, it is itself an element of $\mathbb{V}(X)$. Therefore, there exists coefficients so that $\mathscr{D}_\mu\xi_a$ can be written as a linear combination of the basis vector $\xi_b$. Call these coefficients ${\Gamma^b}_{a\mu}$ the connection coefficients:

$$\mathscr{D}_\mu\xi_a(X) = {\Gamma^b}_{a\mu}(X)\xi_b(X). \tag{4.4.24}$$

Then, the covariant derivative of $\Phi$ becomes:

$$\mathscr{D}_\mu\Phi = (\partial_\mu\phi^a)\xi_a + \phi^a{\Gamma^b}_{a\mu}\xi_b = (\partial_\mu\phi^a + {\Gamma^a}_{b\mu}\phi^b)\xi_a. \tag{4.4.25}$$

Consequently, we define $\mathscr{D}_\mu\phi^a$ as the $a$ component of $\mathscr{D}_\mu\Phi$ and see that it is given by:

$$\mathscr{D}_\mu\phi^a = \partial_\mu\phi^a + {\Gamma^a}_{b\mu}\phi^b. \tag{4.4.26}$$

If we consider the covariant derivative of $\phi$ then as having these components aligned in a column vector:

$$\mathscr{D}_\mu \phi = \begin{pmatrix} \mathscr{D}_\mu \phi^1 \\ \mathscr{D}_\mu \phi^2 \\ \dots \\ \mathscr{D}_\mu \phi^N \end{pmatrix} \tag{4.4.27}$$

and define $\Gamma_\mu$ as the matrix with components ${\Gamma^a}_{b\mu}$ then we have:

$$\mathscr{D}_\mu \phi = \partial_\mu \phi + \Gamma_\mu \phi. \tag{4.4.28}$$

This $\Gamma_\mu$ is simply called the connection.

Because $\mathscr{D}_\mu \Phi$ is a vector in $\mathbb{V}(X)$, when a change of basis is applied according to (4.4.10) we must have that its components change according to (4.4.2):

$$(\mathscr{D}_\mu \phi^a)' = {U^a}_b \mathscr{D}_\mu \phi^b. \tag{4.4.29}$$

Or, in terms of $\mathscr{D}_\mu \phi$:

$$\mathscr{D}_\mu \phi' = U \mathscr{D}_\mu \phi. \tag{4.4.30}$$

Using the definition of the connection coefficients, we can see how they transform under gauge transformation:

$$\mathscr{D}_\mu \xi'_a = {\Gamma^{b'}}_{a\mu} \xi'_b \tag{4.4.31}$$

$$\implies \mathscr{D}_\mu ({U^{-b}}_a \xi_b) = {\Gamma^{b'}}_{a\mu} {U^{-c}}_b \xi_c \tag{4.4.32}$$

$$\implies \partial_\mu {U^{-b}}_a \xi_b + {U^{-b}}_a \mathscr{D}_\mu \xi_b = {\Gamma^{c'}}_{a\mu} {U^{-b}}_c \xi_b \tag{4.4.33}$$

$$\implies \partial_\mu {U^{-b}}_a \xi_b + {U^{-b}}_a {\Gamma^c}_{b\mu} \xi_c = {\Gamma^{c'}}_{a\mu} {U^{-b}}_c \xi_b \tag{4.4.34}$$

$$\implies \partial_\mu {U^{-b}}_a + {U^{-c}}_a {\Gamma^b}_{c\mu} = {\Gamma^{c'}}_{a\mu} {U^{-b}}_c \tag{4.4.35}$$

$$\implies \partial_\mu U^{-1} + \Gamma_\mu U^{-1} = U^{-1} \Gamma'_\mu \tag{4.4.36}$$

$$\implies \Gamma'_\mu = U \Gamma_\mu U^{-1} + U \partial_\mu U^{-1}. \tag{4.4.37}$$

By differentiating $U^{-1}U = \mathbb{1}$ we can move the derivative from the $U^{-1}$ to $U$ which is sometimes convenient:

$$\partial_\mu (U^{-1} U) = \partial_\mu \mathbb{1} = 0 \tag{4.4.38}$$

$$\implies \partial_\mu U^{-1} U + U^{-1} \partial_\mu U = 0 \tag{4.4.39}$$

$$\implies \partial_\mu U^{-1} = -U^{-1} \partial_\mu U U^{-1} \tag{4.4.40}$$

$$\implies \Gamma'_\mu = U \Gamma_\mu U^{-1} - \partial_\mu U U^{-1}. \tag{4.4.41}$$

With this we can compute $\mathscr{D}_\mu \phi'$ explicitly:

$$\mathscr{D}_\mu \phi' = \partial_\mu \phi' + \Gamma'_\mu \phi' = \partial_\mu (U\phi) + (U \Gamma_\mu U^{-1} - \partial_\mu U U^{-1}) U \phi \tag{4.4.42}$$

$$= \partial_\mu U \phi + U \partial_\mu \phi + U \Gamma_\mu \phi - \partial_\mu \phi = U(\partial_\mu \phi + \Gamma_\mu \phi) \tag{4.4.43}$$

$$= U \mathscr{D}_\mu \phi \tag{4.4.44}$$

exactly as we had anticipated.

The connection coefficients ${\Gamma^a}_{b\mu}$ are called the connection coefficients because they "connect" the basis vectors together across different spacetime points. We can see that with a connection we can take any covariant derivative and in order too take derivatives of fields with internal degrees of freedom we had to introduce that connection. However, by introducing the connection we have introduced new fields in these connection coefficients and so now need to find a Lagrangian density for those fields. As such, we need to study the transformation properties of the connection. In particular, consider the definition of the connection coefficients as the covariant derivative of the basis vectors:

$${\Gamma^b}_{a\mu} \xi_b = \mathscr{D}_\mu \xi_a = \lim_{\varepsilon \to 0} \frac{\Upsilon(\xi_a(X + \varepsilon \vec{\iota}_\mu), X + \varepsilon \vec{\iota}_\mu, X) - \xi_a(X)}{\varepsilon}. \tag{4.4.45}$$

Then, under a Lorentz transformation ${L^\mu}_\nu$, we now the coordinate basis vectors transform as:

$$\vec{\iota}'_\mu = {L^{-\nu}}_\mu \vec{\iota}_\nu. \tag{4.4.46}$$

Then, using the chain rule we see that the connection transforms as:

$$\mathscr{D}_\mu \xi'_a = \lim_{\varepsilon \to 0} \frac{\Upsilon(\xi_a(X + \varepsilon \vec{\iota}'_\mu), X + \varepsilon \vec{\iota}'_\mu, X) - \xi_a(X)}{\varepsilon} \tag{4.4.47}$$

$$= \lim_{\varepsilon\to 0} \frac{\Upsilon(\xi_a(X+\varepsilon L^{-\nu}{}_{\mu}\vec{\imath}_\nu), X+\varepsilon L^{-\nu}{}_{\mu}\vec{\imath}_\nu, X) - \xi_a(X)}{\varepsilon} \tag{4.4.48}$$

$$= L^{-\nu}{}_{\mu} \lim_{\varepsilon\to 0} \frac{\Upsilon(\xi_a(X+\varepsilon\vec{\imath}_\nu), X+\varepsilon\vec{\imath}_\nu, X) - \xi_a(X)}{\varepsilon} = L^{-\nu}{}_{\mu}\mathscr{D}_\nu\xi_a \tag{4.4.49}$$

$$\implies \Gamma'^{b}{}_{a\mu}\xi_b = L^{-\nu}{}_{\mu}\Gamma^{b}{}_{a\nu}\xi_b \implies \Gamma'^{b}{}_{a\mu} = L^{-\nu}{}_{\mu}\Gamma^{b}{}_{a\nu}. \tag{4.4.50}$$

Hence we can see that the components of the connection transform under a Lorentz boost just like the coordinate basis vectors or partial derivatives and so they form the components of a 4-vector field.

Now that we have established some properties of the covariant derivative, let us see how it can be used to reconstruct the parallel transport transformation. To define the covariant derivative we only needed the parallel transport map to exist in the immediate neighborhood of any point. That is, we only needed to parallel transport vectors at $X$ to vectors at $X+\varepsilon\vec{\imath}_\mu$ for $\varepsilon$ arbitrarily small. Suppose we have some such way of parallel transporting vectors for $\varepsilon$ sufficiently small. Then, we can define a parallel transport for vectors with greater spacetime separation than $\varepsilon$ by composing the transport of vectors by $\varepsilon$ over and over in the same direction until the final event is reached. Let us formalize this. Consider a particular field $\Phi(X)$ which is defined first at an event $X_0$ by some value $\Phi_0 \in \mathbb{V}(X_0)$. Then, consider the line of events $X(\tau) = X_0 + \tau u$ where $u$ is a 4-vector and $\tau$ is a real number which parameterizes the line and define $\Phi(X(\tau))$ along this line by requiring that if it is parallel transported to $X_0$ then it coincides with $\Phi_0$:

$$\Phi_0 = \Upsilon(\Phi(X(\tau)), X(\tau), X_0). \tag{4.4.51}$$

Then, by the definition of the covariant derivative:

$$u^\mu\mathscr{D}_\mu\Phi(X(\tau)) = \lim_{\varepsilon\to 0} \frac{\Upsilon(\Phi(X(\tau)+\varepsilon u), X(\tau)+\varepsilon u, X(\tau)) - \Phi(X(\tau))}{\varepsilon}. \tag{4.4.52}$$

Quite nicely we can write:

$$X(\tau) + \varepsilon u = X_0 + \tau u + \varepsilon u = X(\tau+\varepsilon). \tag{4.4.53}$$

Thus the covariant derivative in the $u$ direction becomes:

$$u^\mu\mathscr{D}_\mu\Phi(X(\tau)) = \lim_{\varepsilon\to 0} \frac{\Upsilon(\Phi(X(\tau+\varepsilon), X(\tau+\varepsilon), X(\tau)) - \Phi(X(\tau))}{\varepsilon}. \tag{4.4.54}$$

If we now parallel transport this vector to $X_0$ and use linearity:

$$\Upsilon(u^\mu\mathscr{D}_\mu\Phi(X(\tau)), X(\tau), X_0) = \lim_{\varepsilon\to 0} \frac{\Upsilon(\Upsilon(\Phi(X(\tau+\varepsilon)), X(\tau+\varepsilon), X(\tau)), X(\tau), X_0) - \Upsilon(\Phi(X(\tau)), X(\tau), X_0)}{\varepsilon}. \tag{4.4.55}$$

The first term in this difference quotient is a vector parallel transported from $X(\tau+\varepsilon)$ to $X(\tau)$ and then from $X(\tau)$ to $X_0$. This is simply the vector parallel transported directly from $X(\tau+\varepsilon)$ to $X_0$ by our definition of larger parallel transports by composition of smaller parallel transports. So, using (4.4.51) we have:

$$\Upsilon(u^\mu\mathscr{D}_\mu\Phi(X(\tau)), X(\tau), X_0) = \lim_{\varepsilon\to 0} \frac{\Upsilon(\Phi(X(\tau+\varepsilon)), X(\tau+\varepsilon), X_0) - \Upsilon(\Phi(X(\tau)), X(\tau), X_0)}{\varepsilon} \tag{4.4.56}$$

$$= \lim_{\varepsilon\to 0} \frac{\Phi_0 - \Phi_0}{\varepsilon} = 0. \tag{4.4.57}$$

Using the fact that the parallel transport preserves the semi-inner product and the sesquilinearity of that semi-inner product, for arbitrary $\Phi_2$:

$$(\Phi_2(X(\tau)), u^\mu\mathscr{D}_\mu\Phi(X(\tau))) = (\Upsilon(\Phi_2(X(\tau)), X(\tau), X_0), \Upsilon(u^\mu\mathscr{D}_\mu\Phi(X(\tau)), X(\tau), X_0)) \tag{4.4.58}$$

$$= (\Upsilon(\Phi_2(X(\tau)), X(\tau), X_0), 0) = 0. \tag{4.4.59}$$

Or, in components:

$$\mathrm{M}_{ab}\phi_2^{a*}u^\mu\mathscr{D}_\mu\phi^b = 0. \tag{4.4.60}$$

Because this is true for all $\Phi_2$:

$$\mathrm{M}_{ab}u^\mu\mathscr{D}_\mu\phi^b = 0 \tag{4.4.61}$$

$$\implies \mathrm{M}u^\mu\mathscr{D}_\mu\phi = 0. \tag{4.4.62}$$

Because M is invertible, we then must have:

$$u^\mu\mathscr{D}_\mu\phi = 0. \tag{4.4.63}$$

So, a field constructed as in (4.4.51) has vanishing covariant derivative in the $u$ direction and so is said to be

"covariantly constant" in that direction. Explicitly in terms of the derivatives this means:

$$u^\mu \mathscr{D}_\mu \phi = u^\mu \partial_\mu \phi(X(\tau)) + u^\mu \Gamma_\mu(X(\tau))\phi(X(\tau)) = 0. \tag{4.4.64}$$

The total $\tau$ derivative of $\phi(X(\tau))$ is given nicely using the chain rule as:

$$\frac{d\phi(X(\tau))}{d\tau} = \frac{\partial \phi}{\partial x^\mu}\frac{d}{d\tau}(x_0^\mu + \tau u^\mu) = u^\mu \partial_\mu \phi \tag{4.4.65}$$

and so covariant constancy becomes:

$$\frac{d\phi}{d\tau} + u^\mu \Gamma_\mu \phi = 0. \tag{4.4.66}$$

This can be viewed as a system linear differential equations for the components $\phi^a(X(\tau))$. Because they are linear, $\phi(X(\tau))$ will be related to $\phi(X(0))$ by some linear transformation $W(X_0, X(\tau))$:

$$\phi(X(\tau)) = W(X_0, X(\tau))\phi(X_0) \tag{4.4.67}$$

$$\implies \frac{dW}{d\tau}\phi(X_0) + u^\mu \Gamma_\mu W \phi(X_0) = 0. \tag{4.4.68}$$

Because this is true for any $\phi(X_0)$, we have:

$$\frac{dW}{d\tau} + u^\mu \Gamma_\mu W = 0 \tag{4.4.69}$$

and at $\tau = 0$:

$$\phi(X_0) = W(X_0, X_0)\phi(X_0) \quad \forall \phi(X_0) \implies W(X_0, X_0) = \mathbb{1}. \tag{4.4.70}$$

This differential equation and initial condition then completely determine $W(X_0, X(\tau))$. Now let us take (4.4.51) and parallel transport both sides from $X_0$ to $X(\tau)$ and again use our composition property that transporting from $X(\tau)$ to $X_0$ and then from $X_0$ to $X(\tau)$ is just transporting from $X(\tau)$ to $X(\tau)$ and so nothing is done:

$$\Upsilon(\Phi_0, X_0, X(\tau)) = \Upsilon(\Upsilon(\Phi(X(\tau)), X(\tau), X_0), X_0, X(\tau)) \tag{4.4.71}$$

$$\implies \Upsilon(\Phi_0, X_0, X(\tau)) = \Upsilon(\Phi(X(\tau)), X_0, X_0) = \Phi(X(\tau)). \tag{4.4.72}$$

In components then:

$$\Upsilon(\phi_0^a \xi_a, X_0, X(\tau)) = \phi^a(X(\tau))\xi_a(X(\tau)). \tag{4.4.73}$$

Using (4.4.61):

$$\phi(X(\tau)) = W(X_0, X(\tau))\phi(X_0) \implies \phi^a(X(\tau)) = W^a{}_b(X_0, X(\tau))\phi^b(X_0) \tag{4.4.74}$$

$$\implies \phi^a(X_0)\Upsilon(\xi_a, X_0, X(\tau)) = W^b{}_a(X_0, X(\tau))\phi^a(X_0)\xi_b(X(\tau)). \tag{4.4.75}$$

Because this is true for all $\phi^a(X_0)$, we have that the parallel transported basis vectors are related to the basis vectors at the new location by:

$$\Upsilon(\xi_a, X_0, X_0 + u\tau) = W^b{}_a(X_0, X(\tau))\xi_b(X_0 + u\tau). \tag{4.4.76}$$

Thus, this allows us to compute the parallel transport of any basis vector (and so any vector through linearity) along any straight line in spacetime simply by computing $W$. However, we can generalize further to develop how to compute parallel transport along any curved path, not simply a straight line. Before we do so, it will be helpful to understand better some structure of this parallel transport along a line.

Consider (4.4.76) and parallel transport both sides from $X(\tau)$ to $X_0$:

$$\Upsilon(\Upsilon(\xi_a, X_0, X(\tau)), X(\tau), X_0) = \Upsilon(W^b{}_a(X_0, X(\tau))\xi_b(X(\tau)), X(\tau), X_0) \tag{4.4.77}$$

$$\implies \Upsilon(\xi_a, X_0, X_0) = W^b{}_a(X_0, X(\tau))\Upsilon(\xi_b(X(\tau)), X(\tau), X_0) \tag{4.4.78}$$

$$\implies W^{-b}{}_a(X_0, X(\tau))\xi_b(X_0) = \Upsilon(\xi_a(X(\tau)), X(\tau), X_0). \tag{4.4.79}$$

Now let us call the point $X_0$ as $X$ and $X(\tau)$ as $X'$. After parallel transporting the basis vectors from $X'$ to $X$, they will be in the same vector space by construction. As well, because the parallel transport by definition preserves the semi-inner product the parallel transported basis vectors also satisfy (4.4.5) and thus they are linearly related to the basis vectors at $X$ by an element of $G$ (as these matrices are precisely those which preserve (4.4.5) by definition). So, it must be the case that $W^{-1}$ is an element of $G$. By (4.4.70) when $X' = X$ we have $W^{-1} = \mathbb{1}$. Because $W^{-1}$ will vary continuously in $X'$ (otherwise the covariant derivative does not exist) and is connected to the identity, $W^{-1}$ must not only be in $G$ but further must be in the component of the gauge group which is connected to the identity. This connected component is then a connected Lie group

and so has some basis of generators $T_a$ with parameters $\theta^a$ in terms of which any element $U$ can be written as:

$$U = e^{\theta^a T_a}. \tag{4.4.80}$$

By (4.4.17) we then have:

$$e^{\theta^a T_a^\dagger} \mathrm{M} e^{\theta^b T_b} = \mathrm{M}. \tag{4.4.81}$$

Differentiating both sides with respect to $\theta^a$ at $\theta = 0$ we have:

$$T_a^\dagger \mathrm{M} + \mathrm{M} T_a = 0. \tag{4.4.82}$$

Hence, the generators $T_a$ are precisely the matrices which satisfy this condition. Because $W^{-1}$ must be in this connected Lie group there must be some $\theta^a(X, X')$ so that:

$$W^{-1}(X, X') = e^{\theta^a(X,X')T_a}. \tag{4.4.83}$$

Then, (4.4.70) becomes that $\theta^a(X, X) = 0$. Now let us consider $X'$ which are near $X$ and perform a Taylor expansion:

$$\theta^a(X, X') = \theta^a(X, X) + \partial'_\mu \theta^a\big|_{X'=X} (x'^\mu - x^\mu) + \mathcal{O}(x' - x)^2 \tag{4.4.84}$$

$$= \partial'_\mu \theta^a\big|_{X'=X} (x'^\mu - x^\mu) + \mathcal{O}(x' - x)^2 \tag{4.4.85}$$

$$W^{-1}(X, X') = e^{\theta^a(X,X')T_a} = e^{\partial'_\mu \theta^a\big|_{X'=X}(x'^\mu - x^\mu)T_a + \mathcal{O}(x'-x)^2} \tag{4.4.86}$$

$$= \mathbb{1} + \partial'_\mu \theta^a\big|_{X'=X} (x'^\mu - x^\mu)T_a + \mathcal{O}(x' - x)^2. \tag{4.4.87}$$

In components then:

$$W^{-b}{}_a(X, X') = \delta^b_a + \partial'_\mu \theta^c\big|_{X'=X} (x'^\mu - x^\mu)(T_c)^b{}_a + \mathcal{O}(x' - x)^2. \tag{4.4.88}$$

$$\Upsilon(\xi_a(X'), X', X) = \left(\delta^b_a + \partial'_\mu \theta^c\big|_{X'=X} (x'^\mu - x^\mu)(T_c)^b{}_a + \mathcal{O}(x' - x)^2\right) \xi_b(X) \tag{4.4.89}$$

Using this to compute the covariant derivative of the basis vectors:

$$\mathscr{D}_\mu \xi_a(X) = \lim_{\varepsilon \to 0} \frac{\Upsilon(\xi_a(X + \varepsilon \vec{\iota}_\mu), X + \varepsilon \vec{\iota}_\mu, X) - \xi_a(X)}{\varepsilon} \tag{4.4.90}$$

$$= \lim_{\varepsilon \to 0} \frac{\xi_a(X) + \partial'_\mu \theta^c\big|_{X'=X} (T_c)^b{}_a \xi_b(X)\varepsilon + \mathcal{O}(\varepsilon^2) - \xi_a(X)}{\varepsilon} \tag{4.4.91}$$

$$= \partial'_\mu \theta^c\big|_{X'=X} (T_c)^b{}_a \xi_b(X). \tag{4.4.92}$$

By matching with (4.4.24), we see that the connection coefficients can be expressed as:

$$\Gamma^b{}_{a\mu}(X) = \partial'_\mu \theta^c\big|_{X'=X} (T_c)^b{}_a. \tag{4.4.93}$$

Thus, the connection can be written as a linear combination of the generators of the gauge group. Those coefficients of linear combination we write as $\Gamma^a_\mu(X)$ and call the connection fields:

$$\Gamma^a_\mu \overset{\text{def}}{=} \partial'_\mu \theta^a\big|_{X'=X} \implies \Gamma^b{}_{a\mu} = \Gamma^c_\mu (T_c)^b{}_a, \qquad \Gamma_\mu = \Gamma^c_\mu T_c.s \tag{4.4.94}$$

Note that this notation can be somewhat confusing at first but what is going on can be sorted out by looking at the number of indices on $\Gamma$. When $\Gamma$ is written with only 1 index like $\Gamma_\mu$ it is referring to the matrix form of the connection. When $\Gamma$ is written with 2 indices like $\Gamma^a_\mu$ it is referring to the components of the expansion of the connection in terms of the basis of gauge group generators $T_a$. When $\Gamma$ is written with 3 indices like $\Gamma^b{}_{a\mu}$, it is referring to the components of the matrix $\Gamma_\mu$. To compute the $\Gamma^b{}_{a\mu}$ from the $\Gamma^a_\mu$, one needs the explicit matrix forms of the gauge group generators $T_a$. Because the connection can be written as a linear combination of the group generators, the number of independent fields in the connection is equal to the number of generators of the group (that precise number will of course depend on the group, it is the number of linearly independent $T_a$ which satisfy (4.4.82)). The number of generators of the gauge group is called the dimension of the gauge group and we will write it as $N_g$. Recall that the $\Gamma_\mu$ are $N$x$N$ matrices. Because the generators of the gauge group are by definition linearly independent and the dimension of the vector space of $N$x$N$ matrices is $N^2$ the dimension of the gauge group $N_g$ is at most $N^2$.

We can now translate our results for the gauge transformation of the connection to the gauge transformation of the fields $\Gamma^a_\mu$. We found that for a gauge transformation $U(X)$ the connection is transformed as:

$$\Gamma'_\mu = U \Gamma_\mu U^{-1} + U \partial_\mu U^{-1} \tag{4.4.95}$$

If we consider gauge transformations which belong to the connected component of the gauge group, we can give a more detailed expression for this in components. In all currently experimentally tested gauge theories, the gauge group is connected and so all elements of it belong to the connected component. However, it is theoretically possible for the gauge group to be disconnected (and thus involve some discrete transformations as well). Entries of the connected component can be written as exponentials of the generators:

$$U(X) = e^{\theta^a(X)T_a} \implies U^{-1}(X) = e^{-\theta^a(X)T_a} \tag{4.4.96}$$

Then, in terms of the fields (4.4.95) reads:

$$\Gamma'^a_\mu T_a = e^{\theta^b T_b}\Gamma^a_\mu T_a e^{-\theta^b T_b} + e^{\theta^b T_b}\partial_\mu\left(e^{-\theta^b T_b}\right). \tag{4.4.97}$$

In order to simplify this we need to be able to compute $e^{\theta^b T_b}T_a e^{-\theta^b T_b}$ and $e^{\theta^b T_b}\partial_\mu\left(e^{-\theta^b T_b}\right)$. Let us start with the first one. Define:

$$G_a(s) \stackrel{\text{def}}{=} e^{s\theta^b T_b}T_a e^{-s\theta^b T_b} \implies G_a(0) = T_a, \qquad G_a(1) = e^{\theta^b T_b}T_a e^{-\theta^b T_b}. \tag{4.4.98}$$

It turns out that $G_a(s)$ satisfies a nice differential equation. Differentiating $G$:

$$\frac{dG_a}{ds} = \frac{d}{ds}\left(e^{s\theta^b T_b}\right)T_a e^{-s\theta^b T_b} + e^{s\theta^b T_b}T_a\frac{d}{ds}\left(e^{-s\theta^b T_b}\right). \tag{4.4.99}$$

The derivative of this matrix exponential can be computed by using its power series definition:

$$\frac{d}{ds}\left(e^{s\theta^b T_b}\right) = \frac{d}{ds}\sum_{n=0}^{\infty}\frac{(s\theta^a T_a)^n}{n!} = \sum_{n=1}^{\infty}\frac{(\theta^a T_a)^n}{n!}ns^{n-1} \tag{4.4.100}$$

$$= \theta^a T_a\sum_{n=0}^{\infty}\frac{(\theta^b T_b)^n}{n!}s^n = \theta^a T_a e^{s\theta^b T_b}. \tag{4.4.101}$$

Thus:

$$\frac{dG_a}{ds} = \theta^b T_b e^{s\theta^c T_c}T_a e^{-s\theta^c T_c} + e^{s\theta^c T_c}T_a\left(-e^{-s\theta^c T_c}\theta^b T_b\right) \tag{4.4.102}$$

$$= \theta^b T_b G_a(s) - G_a(s)\theta^b T_b = [\theta^b T_b, G_a(s)]. \tag{4.4.103}$$

By differentiating $G_a(s)$ again, we find:

$$\frac{d^2G_a}{ds^2} = [\theta^b T_b, \frac{dG_a}{ds}] = [\theta^b T_b, [\theta^c T_c, G_a]]. \tag{4.4.104}$$

Hence each time $G_a$ is differentiated we simply compose the commutator over once more:

$$\frac{d^nG_a}{ds^n} = [\theta^b T_b, \cdot]^n G_a \tag{4.4.105}$$

Reconstructing $G_a(s)$ from its series we have:

$$G_a(s) = \sum_{n=0}^{\infty}\frac{s^n}{n!}G_a^{(n)}(0) = \sum_{n=0}^{\infty}\frac{s^n}{n!}[\theta^b T_b, \cdot]^n G_a(0) = \sum_{n=0}^{\infty}\frac{s^n}{n!}[\theta^b T_b, \cdot]^n T_a. \tag{4.4.106}$$

Because the $T_a$ are the generators of a Lie group, there must exist structure constants ${f_{ab}}^c$ so that:

$$[T_a, T_b] = {f_{ab}}^c T_c. \tag{4.4.107}$$

Thus, the $n$ fold commutator of the generators can be written as a linear combination of the generators. Hence there exist coefficients ${C^b}_a(n)$ so that:

$$[\theta^b T_b, \cdot]^n T_a = {C^b}_a(n)T_b. \tag{4.4.108}$$

We can find a recurrence relation for these coefficients:

$${C^b}_a(n+1)T_b = [\theta^b T_b, \cdot]^{n+1}T_a = [\theta^b T_b, [\theta^c T_c, \cdot]^n T_a] = [\theta^b T_b, {C^c}_a(n)T_c] \tag{4.4.109}$$

$$= \theta^b {C^c}_a(n)[T_b, T_c] = \theta^b {C^c}_a(n){f_{bc}}^d T_d \tag{4.4.110}$$

$$\implies {C^b}_a(n+1) = {f_{dc}}^b\theta^d {C^c}_a(n). \tag{4.4.111}$$

If we define $f_a$ as the matrices with components ${f_{ac}}^b$ (that is, the matrices corresponding to the adjoint representation of the group) and $C(n)$ as the matrices with components ${C^b}_a(n)$, we have:

$$C(n+1) = \theta^a f_a C(n) \implies C(n) = (\theta^a f_a)^n C(0). \tag{4.4.112}$$

$C(0) = \mathbb{1}$ and so:

$$C(n) = (\theta^a f_a)^n \implies [\theta^b T_b, \cdot]^n T_a = ((\theta^c f_c)^n)^b{}_a T_b \tag{4.4.113}$$

$$\implies G_a(s) = \sum_{n=0}^{\infty} \frac{s^n}{n!} ((\theta^c f_c)^n)^b{}_a T_b = \left(e^{s\theta^c f_c}\right)^b{}_a T_b \tag{4.4.114}$$

$$\implies e^{\theta^b T_b} T_a e^{-\theta^b T_b} = \left(e^{\theta^c f_c}\right)^b{}_a T_b. \tag{4.4.115}$$

We can use a similar trick to compute the second term in (4.4.97). Define:

$$H_\mu(s) \stackrel{\text{def}}{=} e^{s\theta^b T_b} \partial_\mu \left(e^{-s\theta^b T_b}\right) \implies H_\mu(0) = \mathbb{1}\partial_\mu \mathbb{1} = 0, \qquad H_\mu(1) = e^{\theta^b T_b} \partial_\mu \left(e^{-\theta^b T_b}\right). \tag{4.4.116}$$

Differentiating $H_\mu$ with respect to $s$:

$$\frac{dH_\mu}{ds} = \frac{d}{ds}\left(e^{s\theta^b T_b}\right) \partial_\mu \left(e^{-s\theta^b T_b}\right) + e^{s\theta^b T_b} \partial_\mu \left(\frac{d}{ds}\left(e^{-s\theta^b T_b}\right)\right) \tag{4.4.117}$$

$$= e^{s\theta^b T_b} \theta^a T_a \partial_\mu \left(e^{-s\theta^b T_b}\right) + e^{s\theta^b T_b} \partial_\mu \left(-\theta^a T_a e^{-s\theta^b T_b}\right) \tag{4.4.118}$$

$$= e^{s\theta^b T_b} \theta^a T_a \partial_\mu \left(e^{-s\theta^b T_b}\right) - e^{s\theta^b T_b} \partial_\mu \theta^a T_a e^{-s\theta^b T_b} - e^{s\theta^b T_b} \theta^a T_a \partial_\mu \left(e^{-s\theta^b T_b}\right) \tag{4.4.119}$$

$$= -\partial_\mu \theta^a e^{s\theta^b T_b} T_a e^{-s\theta^b T_b} = -\partial_\mu \theta^a G_a(s). \tag{4.4.120}$$

Now integrating we get:

$$e^{\theta^b T_b} \partial_\mu \left(e^{-\theta^b T_b}\right) = H_\mu(1) = H_\mu(0) + \int_0^1 \frac{dH_\mu}{ds} ds \tag{4.4.121}$$

$$\implies e^{\theta^b T_b} \partial_\mu \left(e^{-\theta^b T_b}\right) = 0 - \int_0^1 \partial_\mu \theta^a G_a(s) ds = -\partial_\mu \theta^a \int_0^1 \left(e^{s\theta^c f_c}\right)^b{}_a T_b ds. \tag{4.4.122}$$

This integral is best computed in the series expansion of the exponential:

$$\int_0^1 e^{s\theta^c f_c} ds = \int_0^1 \sum_{n=0}^{\infty} \frac{s^n}{n!} (\theta^c f_c)^n \, ds = \sum_{n=0}^{\infty} \frac{1}{n!} (\theta^c f_c)^n \frac{1}{n+1} = \frac{e^{\theta^c f_c} - \mathbb{1}}{\theta^c f_c}. \tag{4.4.123}$$

This last equality is intended to mean that if the function on the right is series expanded then the series which is achieved is the one which appears to the left. Therefore, this function on the right should be interpreted as a shorthand for the series on the left. Then, we can write:

$$e^{\theta^b T_b} \partial_\mu \left(e^{-\theta^b T_b}\right) = -\partial_\mu \theta^a \left(\frac{e^{\theta^c f_c} - \mathbb{1}}{\theta^c f_c}\right)^b{}_a T_b. \tag{4.4.124}$$

Altogether, this means that (1.7.95) can be rewritten as:

$$\Gamma'^a_\mu T_a = \Gamma^a_\mu \left(e^{\theta^c f_c}\right)^b{}_a T_b - \partial_\mu \theta^a \left(\frac{e^{\theta^c f_c} - \mathbb{1}}{\theta^c f_c}\right)^b{}_a T_b \tag{4.4.125}$$

$$\implies \Gamma'^a_\mu = \left(e^{\theta^c f_c}\right)^a{}_b \Gamma^b_\mu - \left(\frac{e^{\theta^c f_c} - \mathbb{1}}{\theta^c f_c}\right)^a{}_b \partial_\mu \theta^b. \tag{4.4.126}$$

Hence the fields mix into each other in a very complicated way (which requires computing a matrix exponential) under the gauge transformation.

We have established how to parallel transport a vector along a straight line. Let us now understand how to parallel transport along a general path. First consider parallel transporting along a path which is a sequence of $K$ line segments. Then, along each line segment there is an associated $u_k$ and $\tau_k$ so that we can define:

$$X_k(\tau_k) = X_k(0) + u_k \tau_k \tag{4.4.127}$$

and each line segment trades off with the one of the next $k$ at the value of $\tau_k$ so that $X_k(\tau_k) = X_{k+1}(0)$. Call that value of $\tau_k$ as $\mathrm{T}_k$. We will not intend any implicit summation over $k$. Then, on each of these segments, we can define $W_k(\tau_k)$ so that:

$$W_k(0) = \mathbb{1}, \qquad \frac{dW_k}{d\tau_k} + u_k^\mu \Gamma_\mu(X_k) W_k = 0. \tag{4.4.128}$$

This is simply the parallel transport along each line segment. Then, the parallel transport on the sequence of line segments is given using the composition property by parallel transporting along each segment in order:

$$\phi = W\phi_0 = W_K ... W_2 W_1 \phi_0. \tag{4.4.129}$$

Thus, the effective parallel transport map $W$ is given simply as the product of the parallel transports along each of the line segments:

$$W = W_K ... W_2 W_1. \tag{4.4.130}$$

If we define $\tau$ to be $\tau_1$ for $\tau < \mathrm{T}_k$, as $\mathrm{T}_1 + \tau_2$ for $\mathrm{T}_1 < \tau < \mathrm{T}_1 + \mathrm{T}_2$, as $\mathrm{T}_1 + \mathrm{T}_2 + \tau_3$ for $\mathrm{T}_1 + \mathrm{T}_2 < \tau < \mathrm{T}_1 + \mathrm{T}_2 + \mathrm{T}_3$ and so on in this fashion then $X(\tau)$ is given by $X_k(\tau_k)$ on each of these respective regions. As well then, we can identify that on each of these regions:

$$\frac{dx^\mu}{d\tau}(\tau) = u_k \tag{4.4.131}$$

where which $k$ is to be selected is determined by which region $\tau$ falls into. Finally then, if we define $W(\tau)$ as the product of $W_k$ down to $W_1$ where the $k$ is selected by the region into which $\tau$ falls, then from the differential equations satisfied by the various $W_k$ we immediately have:

$$W(0) = \mathbb{1}, \qquad \frac{dW}{d\tau} + \frac{dx^\mu}{d\tau}\Gamma_\mu(X)W = 0. \tag{4.4.132}$$

This is true for any number of line segments. We can approximate a generic differentiable curve by a sequence of line segments and so we simply need to take the limit $K \to \infty$ to apply this to any differentiable curve. Conveniently, (4.4.132) is already written in a way which already does not explicitly depend on $K$ and so that limit is essentially trivial. Thus, we see that if we have a differentiable curve given by the points $X(\tau)$ where $\tau$ parameterizes the points of the curve then an initial vector with components $\phi_0$ at $X_0$ can be parallel transported to $X(\tau)$ by:

$$\phi(\tau) = W(\tau)\phi_0 \tag{4.4.133}$$

where $\phi(\tau)$ gives its components as a column vector at $X(\tau)$ and $W(\tau)$ satisfies:

$$W(0) = \mathbb{1}, \qquad \frac{dW}{d\tau} + \frac{dx^\mu}{d\tau}\Gamma_\mu(X(\tau))W = 0. \tag{4.4.134}$$

This matrix $W$ then completely characterizes the parallel transport operation; it is called the Wilson line matrix. It is difficult to find a useful explicit solution for the Wilson line matrix but we can make some progress. First, integrate both sides from 0 to $\tau$:

$$W(\tau) - W(0) + \int_0^\tau \frac{dx^\mu}{ds}\Gamma_\mu(X(s))W(s)ds = 0 \tag{4.4.135}$$

$$\implies W(\tau) = \mathbb{1} - \int_0^\tau \frac{dx^\mu}{dx}\Gamma_\mu(X(s))W(s)ds. \tag{4.4.136}$$

For convenience going forward, define:

$$G(\tau) = \frac{dx^\mu}{d\tau}(\tau)\Gamma_\mu(X(\tau)). \tag{4.4.137}$$

Then:

$$W(\tau) = \mathbb{1} - \int_0^\tau G(s)W(s)ds. \tag{4.4.138}$$

We can insert another copy of this formula into itself on the right hand side for $W(s)$ to get:

$$W(\tau) = \mathbb{1} - \int_0^\tau G(s)\left(\mathbb{1} - \int_0^s G(s')W(s')ds'\right)ds \tag{4.4.139}$$

$$\implies W(\tau) = \mathbb{1} - \int_0^\tau G(s)ds + \int_0^\tau \int_0^s G(s)G(s')W(s')ds'ds. \tag{4.4.140}$$

Iterating this formula $N$ times gives:

$$W(\tau) = \mathbb{1} - \int_0^\tau G(s_1)ds_1 + \int_0^\tau \int_0^{s_1} G(s_1)G(s_2)ds_2 ds_1 - ... + (-1)^N \int_0^\tau ... \int_0^{s_{N-1}} G(s_1)...G(s_N)W(s_N)ds_N ...ds_1. \tag{4.4.141}$$

In the limit that $N \to \infty$ we then have a series expression for $W(\tau)$:

$$W(\tau) = \sum_{n=0}^{\infty} (-1)^n \int_0^\tau \overset{n}{\cdots} \int_0^{s_{n-1}} G(s_1)...G(s_n) ds_n ... ds_1. \tag{4.4.142}$$

This series expansion for $W$ can at least be useful for approximations of the Wilson line matrix. Let us define the path ordering operation $\mathcal{P}:$ which acting on a product of matrices evaluated along the path reorders the matrices so that the later point of evaluation is to the left:

$$\mathcal{P} : A(\tau_1)B(\tau_2) = \begin{cases} A(\tau_1)B(\tau_2), & \tau_1 > \tau_2 \\ \frac{A(\tau_1)B(\tau_2)+B(\tau_2)A(\tau_1)}{2}, & \tau_1 = \tau_2 \\ B(\tau_2)A(\tau_1), & \tau_1 < \tau_2. \end{cases} \tag{4.4.143}$$

Acting on a product of 3 or more matrices, the path ordering operation is defined to equivalently reorder them from last to first along the path when writing the product from left to right. Then, because $s_n < ... < s_2 < s_1 < \tau$, it is already path ordered and so we can trivially apply the path ordering operation to it:

$$W(\tau) = \sum_{n=0}^{\infty} (-1)^n \mathcal{P} : \int_0^\tau \overset{n}{\cdots} \int_0^{s_{n-1}} G(s_1)...G(s_n) ds_n ... ds_1. \tag{4.4.144}$$

Because of the path ordering, we can reorder the factors of $G$ arbitrarily as the path ordering will just return them back to their original order anyway. There are $n!$ ways to order $n$ factors and so we can sum over all $n!$ orderings and divide by $n!$ because under the path ordering all of them will be equivalent. Let $\sigma$ be a list containing each of the numbers 1 to $n$ once each in some order (this is called a permutation of $n$). Then, let $\sigma_k$ be the $k^{\text{th}}$ entry in that list. Then:

$$W(\tau) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \mathcal{P} : \int_0^\tau \overset{n}{\cdots} \int_0^{s_{n-1}} \sum_{\text{All } \sigma} G(s_{\sigma_1})...G(s_{\sigma_n}) ds_n ... ds_1. \tag{4.4.145}$$

If we in each term in the sum over all $\sigma$ rename $s_{\sigma_k}$ to $s_k$ then the resulting integration region will be the set of all values of $s_1, ..., s_n$ from 0 to $\tau$ with the $s_k$ ordered so that the $k^{\text{th}}$ highest element is $\sigma_k$. The region of values of $s_1, ..., s_n$ each from 0 to $\tau$ can be divided into $n!$ regions in which the $s_k$ take a particular ordering on each region. Then, the sum over all such orderings of these integrals is equivalent to simply integrating each $s_k$ from 0 to $\tau$ without having to sum over orderings:

$$\sum_{\text{All } \sigma} \int_0^\tau \overset{n}{\cdots} \int_0^{s_{n-1}} G(s_{\sigma_1})...G(s_{\sigma_n}) ds_n ... ds_1 = \int_0^\tau \overset{n}{\cdots} \int_0^\tau G(s_{\sigma_1})...G(s_{\sigma_n}) ds_{\sigma_n} ... ds_{\sigma_1}. \tag{4.4.146}$$

Then, the Wilson line matrix becomes:

$$W(\tau) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \mathcal{P} : \int_0^\tau \overset{n}{\cdots} \int_0^\tau G(s_{\sigma_1})...G(s_{\sigma_n}) ds_{\sigma_n} ... ds_{\sigma_1} \tag{4.4.147}$$

$$= \mathcal{P} : \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left( \int_0^\tau G(s) ds \right)^n \tag{4.4.148}$$

$$= \mathcal{P} : e^{-\int_0^\tau G(s) ds}. \tag{4.4.149}$$

Using (4.4.137) this becomes:

$$W(\tau) = \mathcal{P} : e^{-\int_0^\tau \frac{dx^\mu}{ds}(s) \Gamma_\mu(X(s)) ds} = \mathcal{P} : e^{-\int_0^\tau \Gamma_\mu(X(s)) dx^\mu(s)}. \tag{4.4.150}$$

If we call $\gamma$ the path traced out by $X(\tau)$ then we see that the Wilson line matrix can be written:

$$W[\gamma] = \mathcal{P} : e^{-\int_\gamma \Gamma_\mu(X) dx^\mu} \tag{4.4.151}$$

where here $W[\gamma]$ is written to indicate that it is a functional of the path. This exponential form might seem compact and useful. However, matrix exponentials are already hard to compute and the path ordering makes the result quite difficult to compute analytically and so in fact (4.4.142) is often more amenable to computations. This exponential form though is sometimes useful for proving theorems.

Using (4.4.133) we can compute the gauge transformation of the Wilson line matrix:

$$\phi'(\tau) = W'(\tau)\phi'_0 \tag{4.4.152}$$

$$\implies U(X(\tau))\phi(\tau) = W'(\tau)U(X_0)\phi_0 \tag{4.4.153}$$

$$\implies U(X(\tau))W(\tau)\phi_0 = W'(\tau)U(X_0)\phi_0. \tag{4.4.154}$$

Because this is true for all $\phi_0$ we have:

$$U(X(\tau))W(\tau) = W'(\tau)U(X_0) \tag{4.4.155}$$

$$\implies W'(\tau) = U(X(\tau))W(\tau)U^{-1}(X_0). \tag{4.4.156}$$

If we let $X_f$ be the final point on the curve $\gamma$ this can equivalently be written:

$$W'[\gamma] = U(X_f)W[\gamma]U^{-1}(X_0). \tag{4.4.157}$$

If the curve $\gamma$ is actually a closed loop then $X_f = X_0$ by definition and we have:

$$W'[\gamma] = U(X_0)W[\gamma]U^{-1}(X_0) \qquad \text{(for a closed loop)}. \tag{4.4.158}$$

If we further take the trace of this matrix and use the cyclicity of the trace:

$$\mathrm{Tr}(W'[\gamma]) = \mathrm{Tr}(U(X_0)W[\gamma]U^{-1}(X_0)) = \mathrm{Tr}(U^{-1}(X_0)U(X_0)W[\gamma]) = \mathrm{Tr}(W[\gamma]) \tag{4.4.159}$$

for a closed loop. Hence, this quantity is actually gauge invariant. This is the first gauge invariant quantity involving only the connection we have identified. It is quite important and it is called the Wilson loop:

$$\mathcal{W}[\gamma] = \mathrm{Tr}(W[\gamma]) = \mathrm{Tr}\ \mathcal{P} : e^{-\oint_\gamma \Gamma_\mu dx^\mu}. \tag{4.4.160}$$

These Wilson loops are not only gauge invariant but in fact if one knows the Wilson loop for any closed path $\gamma$, it is actually sufficient information to reconstruct $\Gamma_\mu$ up to a gauge transformation. To establish such is no small effort but the details can be found in [21]. Thus, the Wilson loops contain all gauge invariant information about the connection.

### 4.4.2 Curvature of the Gauge Connection

Ultimately we will seek a Lagrangian density for our new connection fields and thus will need to make one which has both Lorentz and gauge symmetries. As such, we would like to understand what gauge invariants can be made out of our connection. The Wilson loops do form a complete set of gauge invariants in the sense that from them the connection can be reconstructed up to gauge transformation. However, the Wilson loop is also non-local because it contains evaluation of the connection at many different points. In order to make a local gauge invariant then we should study the Wilson loops in the limit of small loops as in the limit that the loop is shrunk down to a point the Wilson loop would then become local automatically. So, let us consider the Wilson line starting at $X$ along a path consisting of 4 straight segments: first a line segment of length $a$ in the direction $u$, second a line segment of length $b$ in the direction $v$, third a line segment of length $a$ in the direction $-u$, and finally a line segment of length $b$ in the direction $-v$. This path forms a closed rectangle with area $ab$ which we will take to be small. By (4.4.138) we immediately have that the Wilson line operator from this path is given by:

$$W[\gamma] = (\mathbb{1} + bv^\mu\Gamma_\mu(X+bv))(\mathbb{1} + au^\nu\Gamma_\nu(X+bv+au))(\mathbb{1} - bv^\rho\Gamma_\rho(X+au))(\mathbb{1} - au^\sigma\Gamma_\sigma(X)) + \mathcal{O}(a^2) + \mathcal{O}(b^2) \tag{4.4.161}$$

$$\begin{aligned}=&(\mathbb{1} + bv^\mu\Gamma_\mu(X))(\mathbb{1} + au^\nu\Gamma_\nu(X) + abu^\nu v^\alpha\partial_\alpha\Gamma_\nu(X))\\ &\cdot(\mathbb{1} - bv^\rho\Gamma_\rho(X) - abv^\rho u^\beta\partial_\beta\Gamma_\rho(X))(\mathbb{1} - au^\sigma\Gamma_\sigma(X)) + \mathcal{O}(\alpha^2) + \mathcal{O}(b^2)\end{aligned} \tag{4.4.162}$$

$$\begin{aligned}=&(\mathbb{1} + bv^\mu\Gamma_\mu)(\mathbb{1} + au^\nu\Gamma_\nu + abu^\nu v^\alpha\partial_\alpha\Gamma_\nu)\\ &\cdot(\mathbb{1} - bv^\rho\Gamma_\rho - abv^\rho u^\beta\partial_\beta\Gamma_\rho - au^\sigma\Gamma_\sigma + abv^\rho u^\sigma\Gamma_\rho\Gamma_\sigma) + \mathcal{O}(\alpha^2) + \mathcal{O}(b^2)\end{aligned} \tag{4.4.163}$$

$$\begin{aligned}=&(\mathbb{1} + bv^\mu\Gamma_\mu)(\mathbb{1} + au^\nu\Gamma_\nu + abu^\nu v^\alpha\partial_\alpha\Gamma_\nu - bv^\rho\Gamma_\rho - abu^\nu v^\rho\Gamma_\nu\Gamma_\rho - abv^\rho u^\beta\partial_\beta\Gamma_\rho\\ &- au^\sigma\Gamma_\sigma + abv^\rho u^\sigma\Gamma_\rho\Gamma_\sigma) + \mathcal{O}(\alpha^2) + \mathcal{O}(b^2)\end{aligned} \tag{4.4.164}$$

$$=(\mathbb{1} + bv^\mu\Gamma_\mu)(\mathbb{1} - bv^\rho\Gamma_\rho + abu^\nu v^\mu\partial_\mu\Gamma_\nu - abu^\nu v^\mu\Gamma_\nu\Gamma_\mu - abv^\mu u^\nu\partial_\nu\Gamma_\mu + abv^\mu u^\nu\Gamma_\mu\Gamma_\nu) + \mathcal{O}(\alpha^2) + \mathcal{O}(b^2) \tag{4.4.165}$$

$$=\mathbb{1} + bv^\mu\Gamma_\mu - bv^\rho\Gamma_\rho + abv^\mu u^\nu\left(\partial_\mu\Gamma_\nu - \partial_\nu\Gamma_\mu + \Gamma_\mu\Gamma_\nu - \Gamma_\nu\Gamma_\mu\right) + \mathcal{O}(\alpha^2) + \mathcal{O}(b^2) \tag{4.4.166}$$

$$=\mathbb{1} + abv^\mu u^\nu\left(\partial_\mu\Gamma_\nu - \partial_\nu\Gamma_\mu + \Gamma_\mu\Gamma_\nu - \Gamma_\nu\Gamma_\mu\right) + \mathcal{O}(\alpha^2) + \mathcal{O}(b^2). \tag{4.4.167}$$

Thus for small loops the first correction to the identity of the Wilson line is given by the quantity:

$$\Omega_{\mu\nu} = \partial_\mu\Gamma_\nu - \partial_\nu\Gamma_\mu + \Gamma_\mu\Gamma_\nu - \Gamma_\nu\Gamma_\mu \tag{4.4.168}$$

which is called the curvature tensor of the internal space. We can use the Lorentz transformation property of the connection to see that in fact the curvature tensor is a genuine tensor:

$$\begin{aligned}
\Omega'_{\mu\nu} =& \partial'_\mu \Gamma'_\nu - \partial'_\nu \Gamma'_\mu + \Gamma'_\mu \Gamma'_\nu - \Gamma'_\nu \Gamma'_\mu && (4.4.169)\\
=& L^{-\alpha}{}_\mu \partial_\alpha \left( L^{-\beta}{}_\nu \Gamma_\beta \right) - L^{-\beta}{}_\nu \partial_\beta \left( L^{-\alpha}{}_\mu \Gamma_\alpha \right) \\
&+ L^{-\alpha}{}_\mu L^{-\beta}{}_\nu \Gamma_\alpha \Gamma_\beta - L^{-\alpha}{}_\mu L^{-\beta}{}_\nu \Gamma_\beta \Gamma_\alpha && (4.4.170)\\
=& L^{-\alpha}{}_\mu L^{-\beta}{}_\nu \left( \partial_\alpha \Gamma_\beta - \partial_\beta \Gamma_\alpha + \Gamma_\alpha \Gamma_\beta - \Gamma_\beta \Gamma_\alpha \right) && (4.4.171)\\
=& L^{-\alpha}{}_\mu L^{-\beta}{}_\nu \Omega_{\alpha\beta} && (4.4.172)
\end{aligned}$$

and we can see that by its definition the curvature tensor is automatically anti-symmetric in its indices. For our small loop, it consisted of parallel transport along two different directions as well as their inverse directions but in the order so that the translations and their inverses don't automatically cancel. Thus, the small rectangular loop and thus the curvature tensor essentially measures the extent to which parallel translations along two given directions fail to commute. Because along the parallel transport of a vector in the $u$ direction (4.4.64) is satisfied, we see that the covariant derivative generates parallel transportations in the same way that the ordinary derivative generates constant translations. Thus, we suspect that there is some relationship between the commutator of covariant derivatives and the commutator of parallel transports (the curvature tensor). Let us explicitly compute the commutator of covariant derivatives:

$$\begin{aligned}
[\mathscr{D}_\mu, \mathscr{D}_\nu]\phi =& \mathscr{D}_\mu \left( \partial_\nu \phi + \Gamma_\nu \phi \right) - \mathscr{D}_\nu \left( \partial_\mu \phi + \Gamma_\mu \phi \right) && (4.4.173)\\
=& \left( \partial_\mu + \Gamma_\mu \right) \left( \partial_\nu \phi + \Gamma_\nu \phi \right) - \left( \partial_\nu + \Gamma_\nu \right) \left( \partial_\mu \phi + \Gamma_\mu \phi \right) && (4.4.174)\\
=& \partial^2_{\mu\nu} \phi + \Gamma_\mu \partial_\nu \phi + \partial_\mu \Gamma_\nu \phi + \Gamma_\nu \partial_\mu \phi + \Gamma_\mu \Gamma_\nu \phi - \partial^2_{\mu\nu} \phi - \Gamma_\nu \partial_\mu \phi - \partial_\nu \Gamma_\mu \phi - \Gamma_\mu \partial_\nu \phi - \Gamma_\nu \Gamma_\mu \phi && (4.4.175)\\
=& \left( \partial_\mu \Gamma_\nu + \Gamma_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu - \Gamma_\nu \Gamma_\mu \right) \phi = \Omega_{\mu\nu} \phi. && (4.4.176)
\end{aligned}$$

Thus, this gives another way of arriving at the curvature tensor. It is also useful to look at the covariant derivative acting on the basis vectors $\xi_a$:

$$\begin{aligned}
[\mathscr{D}_\mu, \mathscr{D}_\nu]\xi_a =& \mathscr{D}_\mu \mathscr{D}_\nu \xi_a - \mathscr{D}_\nu \mathscr{D}_\mu \xi_a = \mathscr{D}_\mu \left( \Gamma^b{}_{a\nu} \xi_b \right) - \mathscr{D}_\nu \left( \Gamma^b{}_{a\mu} \xi_b \right) && (4.4.177)\\
=& \partial_\mu \Gamma^b{}_{a\nu} \xi_b + \Gamma^b{}_{a\nu} \mathscr{D}_\mu \xi_b - \partial_\nu \Gamma^b{}_{a\mu} \xi_b - \Gamma^b{}_{a\mu} \mathscr{D}_\nu \xi_b && (4.4.178)\\
=& \partial_\mu \Gamma^b{}_{a\nu} \xi_b + \Gamma^b{}_{a\nu} \Gamma^c{}_{b\mu} \xi_c - \partial_\nu \Gamma^b{}_{a\mu} \xi_b - \Gamma^b{}_{a\mu} \Gamma^c{}_{b\nu} \xi_c && (4.4.179)\\
=& \left( \partial_\mu \Gamma^b{}_{a\nu} - \partial_\nu \Gamma^b{}_{a\mu} + \Gamma^b{}_{c\mu} \Gamma^c{}_{a\nu} - \Gamma^b{}_{c\nu} \Gamma^c{}_{a\mu} \right) \xi_b && (4.4.180)\\
=& \left( \Omega_{\mu\nu} \right)^b{}_a \xi_b. && (4.4.181)
\end{aligned}$$

This gives the commutator of covariant derivatives acting on $\Phi$ similarly as:

$$\begin{aligned}
[\mathscr{D}_\mu, \mathscr{D}_\nu]\Phi =& \mathscr{D}_\mu \mathscr{D}\nu\Phi - \mathscr{D}_\nu \mathscr{D}_\mu \Phi = \mathscr{D}_\mu \left( \partial_\nu \phi^a \xi_a + \phi^a \mathscr{D}_\nu \xi_a \right) - \mathscr{D}_\nu \left( \partial_\mu \phi^a \xi_a + \phi^a \mathscr{D}_\mu \xi_a \right) && (4.4.182)\\
=& \partial^2_{\mu\nu} \phi^a \xi_a + \partial_\nu \phi^a \mathscr{D}_\mu \xi_a + \partial_\mu \phi^a \mathscr{D}_\nu \xi_a + \phi^a \mathscr{D}_\mu \mathscr{D}_\nu \xi_a \\
&- \partial^2_{\mu\nu} \phi^a \xi_a - \partial_\mu \phi^a \mathscr{D}_\nu \xi_a - \partial_\nu \phi^a \mathscr{D}_\mu \xi_a - \phi^a \mathscr{D}_\nu \mathscr{D}_\mu \xi_a && (4.4.183)\\
=& \phi^a \left( \mathscr{D}_\mu \mathscr{D}_\nu - \mathscr{D}_\nu \mathscr{D}_\mu \right) \xi_a = \phi^a [\mathscr{D}_\mu, \mathscr{D}_\nu] \xi_a = \phi^a (\Omega_{\mu\nu})^b{}_a \xi_b && (4.4.184)
\end{aligned}$$

which is exactly consistent with (4.4.176). We can rewrite the right hand side explicitly in terms of $\Phi$ using the semi-inner product. Consider the semi-inner product of $\Phi$ and a basis vector and use (4.4.5) as well as the invertibility of M:

$$\begin{aligned}
(\xi_b, \Phi) = (\xi_b, \phi^a \xi_a) = \phi^a (\xi_b, \xi_a) = \mathrm{M}_{ba} \phi^a && (4.4.185)\\
\implies \phi^a = \mathrm{M}^{-ab} (\xi_b, \Phi). && (4.4.186)
\end{aligned}$$

Then, (4.4.184) can be rewritten:

$$[\mathscr{D}_\mu, \mathscr{D}_\nu]\Phi = (\Omega_{\mu\nu})^b{}_a \xi_b \mathrm{M}^{-ac} (\xi_c, \Phi). \tag{4.4.187}$$

Thus, we can define the curvature operator $\hat{\mathbf{\Omega}}_{\mu\nu}$ as a linear operator on the internal space (meaning it acts linearly on an element of $\mathbb{V}(X)$ and returns an element of $\mathbb{V}(X)$) by:

$$\begin{aligned}
\hat{\mathbf{\Omega}}_{\mu\nu} = (\Omega_{\mu\nu})^b{}_a \xi_b \mathrm{M}^{-ac} (\xi_c, \cdot) && (4.4.188)\\
\implies [\mathscr{D}_\mu, \mathscr{D}_\nu]\Phi = \hat{\mathbf{\Omega}}_{\mu\nu} \Phi. && (4.4.189)
\end{aligned}$$

Let us also see how the curvature tensor changes under gauge transformation. Directly calculating from the gauge transformation of the connection:

$$\Omega'_{\mu\nu} = \partial_\mu \Gamma'_\nu - \partial_\nu \Gamma'_\mu + \Gamma'_\mu \Gamma'_\nu - \Gamma'_\nu \Gamma'_\mu \tag{4.4.190}$$

$$= \partial_\mu \left( U\Gamma_\nu U^{-1} + U\partial_\nu U^{-1} \right) - \partial_\nu \left( U\Gamma_\mu U^{-1} + U\partial_\mu U^{-1} \right) + \left( U\Gamma_\mu U^{-1} + U\partial_\mu U^{-1} \right) \left( U\Gamma_\nu U^{-1} + U\partial_\nu U^{-1} \right) - \left( U\Gamma_\nu U^{-1} + U\partial_\nu U^{-1} \right) \left( U\Gamma_\mu U^{-1} + U\partial_\mu U^{-1} \right) \tag{4.4.191}$$

$$\begin{aligned} =& \partial_\mu U\Gamma_\nu U^{-1} + U\partial_\mu \Gamma_\nu U^{-1} + U\Gamma_\nu \partial_\mu U^{-1} + \partial_\mu U \partial_\nu U^{-1} + U\partial^2_{\mu\nu} U^{-1} \\ & - \partial_\nu U\Gamma_\mu U^{-1} - U\partial_\nu \Gamma_\mu U^{-1} - U\Gamma_\mu \partial_\nu U^{-1} - \partial_\nu U \partial_\mu U^{-1} - U\partial^2_{\mu\nu} U^{-1} \\ & + U\Gamma_\mu \Gamma_\nu U^{-1} + U\partial_\mu U^{-1} U\Gamma_\nu U^{-1} + U\Gamma_\mu \partial_\nu U^{-1} + U\partial_\mu U^{-1} U\partial_\nu U^{-1} \\ & - U\Gamma_\nu \Gamma_\mu U^{-1} - U\partial_\nu U^{-1} U\Gamma_\mu U^{-1} - U\Gamma_\nu \partial_\mu U^{-1} - U\partial_\nu U^{-1} U\partial_\mu U^{-1} \end{aligned} \tag{4.4.192}$$

$$\begin{aligned} =& \partial_\mu U\Gamma_\nu U^{-1} + U\partial_\mu \Gamma_\nu U^{-1} + \partial_\mu U \partial_\nu U^{-1} - \partial_\nu U\Gamma_\mu U^{-1} - U\partial_\nu \Gamma_\mu U^{-1} - \partial_\nu U \partial_\mu U^{-1} \\ & + U\Gamma_\mu \Gamma_\nu U^{-1} - \partial_\mu U\Gamma_\nu U^{-1} - \partial_\mu U \partial_\nu U^{-1} - U\Gamma_\nu \Gamma_\mu U^{-1} + \partial_\nu U\Gamma_\mu U^{-1} + \partial_\nu U \partial_\mu U^{-1} \end{aligned} \tag{4.4.193}$$

$$= U \left( \partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu + \Gamma_\mu \Gamma_\nu - \Gamma_\nu \Gamma_\mu \right) U^{-1} \tag{4.4.194}$$

$$= U\Omega_{\mu\nu} U^{-1}. \tag{4.4.195}$$

Any quantity transforming this way under the gauge transformations is said to be in the "adjoint representation of the gauge group". From this we can see that the curvature operator is gauge invariant (and recall (4.4.10) and (4.4.17)):

$$\hat{\mathbf{\Omega}}'_{\mu\nu} = \left( \Omega'_{\mu\nu} \right)^b{}_a \xi'_b \mathrm{M}^{-ac} (\xi'_c, \cdot) \tag{4.4.196}$$

$$= \left( U\Omega_{\mu\nu} U^{-1} \right)^b{}_a U^{-d}{}_b \xi_d \mathrm{M}^{-ac} (U^{-e}{}_c \xi_e, \cdot) \tag{4.4.197}$$

$$= U^b{}_f (\Omega_{\mu\nu})^f{}_g U^{-g}{}_a U^{-d}{}_b \xi_d \mathrm{M}^{-ac} \left( U^{-e}{}_c \right)^* (\xi_e, \cdot) \tag{4.4.198}$$

$$= (\Omega_{\mu\nu})^f{}_g U^{-1} U^d{}_f \xi_d \left( U^{-1} \mathrm{M}^{-1} U^{-\dagger} \right)^{ge} (\xi_e, \cdot) \tag{4.4.199}$$

$$= (\Omega_{\mu\nu})^f{}_g \delta^d_f \xi_d \left( U^\dagger \mathrm{M} U \right)^{-ge} (\xi_e, \cdot) \tag{4.4.200}$$

$$= (\Omega_{\mu\nu})^d{}_g \xi_d \mathrm{M}^{-ge} (\xi_e, \cdot) \tag{4.4.201}$$

$$= \hat{\mathbf{\Omega}}_{\mu\nu} \tag{4.4.202}$$

which one would expect given that everything else in (4.4.189) is already gauge invariant.

We can use the expression for the connection in terms of the connection fields to see that there are also curvature fields we can compute from them:

$$\Omega_{\mu\nu} = \partial_\mu \Gamma^a_\nu T_a - \partial_\nu \Gamma^a_\mu T_a + \Gamma^a_\mu T_a \Gamma^b T_b - \Gamma^b_\nu T_b \Gamma^a_\mu T_a \tag{4.4.203}$$

$$= \left( \partial_\mu \Gamma^a_\nu - \partial_\nu \Gamma^a_\mu \right) T_a + \Gamma^a_\mu \Gamma^b_\nu [T_a, T_b] \tag{4.4.204}$$

$$= \left( \partial_\mu \Gamma^a_\nu - \partial_\nu \Gamma^a_\mu \right) T_a + \Gamma^a_\mu \Gamma^b_\nu f_{ab}{}^c T_c \tag{4.4.205}$$

$$= \left( \partial_\mu \Gamma^a_\nu - \partial_\nu \Gamma^a_\mu + f_{bc}{}^a \Gamma^b_\mu \Gamma^c_\nu \right) T_a. \tag{4.4.206}$$

Hence just as it is possible to write the connection as a linear combination of the generators, it is also possible to write the curvature as one:

$$\Omega_{\mu\nu} = \Omega^a_{\mu\nu} T_a, \qquad \Omega^a_{\mu\nu} = \partial_\mu \Gamma^a_\nu - \partial_\nu \Gamma^a_\mu + f_{bc}{}^a \Gamma^b_\mu \Gamma^c_\nu. \tag{4.4.207}$$

Using (4.4.195) then we can find curvature fields change under gauge transformations connected to the identity:

$$\Omega'^a_{\mu\nu} T_a = \Omega'_{\mu\nu} = U\Omega_{\mu\nu} U^{-1} = \Omega^a_{\mu\nu} e^{\theta^b T_b} T_a e^{-\theta^b T_b} = \Omega^a_{\mu\nu} \left( e^{\theta^c f_c} \right)^b{}_a T_b \tag{4.4.208}$$

$$\implies \Omega'^a_{\mu\nu} = \left( e^{\theta^c f_c} \right)^a{}_b \Omega^b_{\mu\nu}. \tag{4.4.209}$$

There is some intuition as to why this tensor is called the curvature tensor. If the internal vector spaces at each point in spacetime happened to simply be exactly the same vector space so that we could use the same basis $\xi_a$ independent of $X$ we could simply take the parallel transport map to be the identity map. Consequently, all of the connection coefficients would vanish and the curvature tensor would vanish as well.

Then, the covariant derivative is just a partial derivative and so covariant derivatives commute. Under gauge transformation the connection could be made non-zero but still the curvature tensor would vanish as $\Gamma'^{\mu}$ does not remain 0 if $\Gamma^{\mu}=0$ from (4.4.37) but $\Omega'_{\mu\nu}=0$ if $\Omega_{\mu\nu}=0$ by (4.4.195). If the curvature tensor is non-zero somewhere then it is not possible to turn the covariant derivatives into partial derivatives everywhere or set the connection to 0 or trivialize the parallel transport and so it is not possible to use the same basis for the internal space across spacetime. Whatever basis vectors one has in that case must twist, turn, or as we say "curve" into being new basis vectors as one moves around and the extent by which those basis vectors change locally is given by the curvature tensor. Thus, $\Omega_{\mu\nu}$ measures the extent to which the internal space is "curved".

It will sometimes be useful to consider a covariant derivative of the curvature tensor. Because the curvature operator depends on the basis vectors $\xi_a$, in order to define a meaningful derivative for it, we must use the parallel transport map. Thus, the definition of the covariant derivative of the curvature operator is taken exactly as we would expect by parallel transporting in the same way as in (4.4.20):

$$\mathscr{D}_\rho\hat{\mathbf{\Omega}}_{\mu\nu}\overset{\text{def}}{=}\lim_{\varepsilon\to0}\frac{(\Omega_{\mu\nu})^b{}_a(X+\varepsilon\vec{\iota}_\rho)\mathrm{M}^{-ac}\Upsilon(\xi_b(X+\varepsilon\vec{\iota}_\rho),X+\varepsilon\vec{\iota}_\rho,X)\left(\Upsilon(\xi_c(X+\varepsilon\vec{\iota}_\rho),X+\varepsilon\vec{\iota}_\rho,X),\cdot\right)-\hat{\mathbf{\Omega}}_{\mu\nu}}{\varepsilon} \tag{4.4.210}$$

$$\begin{aligned}=\lim_{\varepsilon\to0}\Bigg(&\frac{(\Omega_{\mu\nu})^b{}_a(X+\varepsilon\vec{\iota}_\rho)-(\Omega_{\mu\nu})^b{}_a(X)}{\varepsilon}\mathrm{M}^{-ac}\Upsilon(\xi_b(X+\varepsilon\vec{\iota}_\rho),X+\varepsilon\vec{\iota}_\rho,X)\left(\Upsilon(\xi_c(X+\varepsilon\vec{\iota}_\rho),X+\varepsilon\vec{\iota}_\rho,X),\cdot\right)\\&+(\Omega_{\mu\nu})^b{}_a(X)\frac{\mathrm{M}^{-ac}\Upsilon(\xi_b(X+\varepsilon\vec{\iota}_\rho),X+\varepsilon\vec{\iota}_\rho,X)\left(\Upsilon(\xi_c(X+\varepsilon\vec{\iota}_\rho),X+\varepsilon\vec{\iota}_\rho,X),\cdot\right)-\mathrm{M}^{-ac}\xi_b(\xi_c,\cdot)}{\varepsilon}\Bigg)\end{aligned} \tag{4.4.211}$$

$$\begin{aligned}=&\partial_\rho(\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\xi_b\left(\xi_c,\cdot\right)\\&+(\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\lim_{\varepsilon\to0}\frac{\Upsilon(\xi_b(X+\varepsilon\vec{\iota}_\rho),X+\varepsilon\vec{\iota}_\rho,X)\left(\Upsilon(\xi_c(X+\varepsilon\vec{\iota}_\rho),X+\varepsilon\vec{\iota}_\rho,X),\cdot\right)-\xi_b(\xi_c,\cdot)}{\varepsilon}\end{aligned} \tag{4.4.212}$$

$$\begin{aligned}=&(\partial_\rho\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\xi_b\left(\xi_c,\cdot\right)\\&+(\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\lim_{\varepsilon\to0}\frac{\Upsilon(\xi_b(X+\varepsilon\vec{\iota}_\rho),X+\varepsilon\vec{\iota}_\rho,X)-\xi_b}{\varepsilon}\left(\Upsilon(\xi_c(X+\varepsilon\vec{\iota}_\rho),X+\varepsilon\vec{\iota}_\rho,X),\cdot\right)\\&+(\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\xi_b\lim_{\varepsilon\to0}\frac{\left(\Upsilon(\xi_c(X+\varepsilon\vec{\iota}_\rho),X+\varepsilon\vec{\iota}_\rho,X),\cdot\right)-(\xi_c,\cdot)}{\varepsilon}\end{aligned} \tag{4.4.213}$$

$$=(\partial_\rho\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\xi_b\left(\xi_c,\cdot\right)+(\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\mathscr{D}_\rho\xi_b\left(\xi_c,\cdot\right)+(\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\xi_b\left(\mathscr{D}_\rho\xi_c,\cdot\right) \tag{4.4.214}$$

$$=(\partial_\rho\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\xi_b\left(\xi_c,\cdot\right)+(\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\Gamma^d{}_{b\rho}\xi_d\left(\xi_c,\cdot\right)+(\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\xi_b\left(\Gamma^d{}_{c\rho}\xi_d,\cdot\right) \tag{4.4.215}$$

$$=(\partial_\rho\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\xi_b\left(\xi_c,\cdot\right)+(\Omega_{\mu\nu})^d{}_a\mathrm{M}^{-ac}\Gamma^b{}_{d\rho}\xi_b\left(\xi_c,\cdot\right)+(\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ad}\xi_b\left(\Gamma^c{}_{d\rho}\right)^*\left(\xi_c,\cdot\right) \tag{4.4.216}$$

$$=(\partial_\rho\Omega_{\mu\nu}+\Gamma_\rho\Omega_{\mu\nu})^b{}_a\mathrm{M}^{-ac}\xi_b\left(\xi_c,\cdot\right)+(\Omega_{\mu\nu})^b{}_e\mathrm{M}^{-ed}\xi_b\left(\Gamma^f{}_{d\rho}\right)^*\mathrm{M}_{fa}\mathrm{M}^{-ac}\left(\xi_c,\cdot\right) \tag{4.4.217}$$

$$=\left(\partial_\rho\Omega_{\mu\nu}+\Gamma_\rho\Omega_{\mu\nu}+\Omega_{\mu\nu}\mathrm{M}^{-1}\Gamma_\rho^\dagger\mathrm{M}\right)^b{}_a\mathrm{M}^{-ac}\xi_b\left(\xi_c,\cdot\right) \tag{4.4.218}$$

From (4.4.94) we can see that the connection fields are real and therefore by (4.4.82):

$$\Gamma^a_\rho T^\dagger\mathrm{M}+\mathrm{M}\Gamma^a_\rho T_a=0 \tag{4.4.219}$$

$$\implies\Gamma_\rho^\dagger\mathrm{M}+\mathrm{M}\Gamma_\rho=0 \tag{4.4.220}$$

$$\implies\mathrm{M}^{-1}\Gamma_\rho^\dagger\mathrm{M}=-\Gamma_\rho. \tag{4.4.221}$$

With this, (4.4.218) becomes:

$$\mathscr{D}_\rho\hat{\mathbf{\Omega}}_{\mu\nu}=(\partial_\rho\Omega_{\mu\nu}+\Gamma_\rho\Omega_{\mu\nu}-\Omega_{\mu\nu}\Gamma_\rho)^b{}_a\mathrm{M}^{-ac}\xi_b\left(\xi_c,\cdot\right) \tag{4.4.222}$$

Equivalently then, the components of the covariant derivative are:

$$\mathscr{D}_\rho(\Omega_{\mu\nu})^b{}_a=(\partial_\rho\Omega_{\mu\nu}+\Gamma_\rho\Omega_{\mu\nu}-\Omega_{\mu\nu}\Gamma_\rho)^b{}_a. \tag{4.4.223}$$

Or, even more nicely in terms of the curvature tensor matrix:

$$\mathscr{D}_\rho\Omega_{\mu\nu}=\partial_\rho\Omega_{\mu\nu}+\Gamma_\rho\Omega_{\mu\nu}-\Omega_{\mu\nu}\Gamma_\rho \tag{4.4.224}$$

Using our expression in terms of the curvature fields we also have:

$$\mathscr{D}_\rho \Omega_{\mu\nu} = \partial_\rho \Omega^a_{\mu\nu} T_a + \Gamma^b_\rho \Omega^c_{\mu\nu}(T_b T_c - T_c T_b) = \partial_\rho \Omega^a T_a + f_{bc}{}^a \Gamma^b_\rho \Omega^c_{\mu\nu} T_a \tag{4.4.225}$$

$$\implies \mathscr{D}_\rho \Omega^a_{\mu\nu} = \partial_\rho \Omega^a_{\mu\nu} + f_{bc}{}^a \Gamma^b_\rho \Omega^c_{\mu\nu}. \tag{4.4.226}$$

Because the covariant derivative is gauge invariant by definition and the curvature operator $\hat{\mathbf{\Omega}}_{\mu\nu}$ is gauge invariant, $\mathscr{D}_\rho \hat{\mathbf{\Omega}}_{\mu\nu}$ is also gauge invariant and so because of equations (4.4.195) and (4.4.202) we expect:

$$\mathscr{D}_\rho \Omega'_{\mu\nu} = U \mathscr{D}_\rho \Omega_{\mu\nu} U^{-1}. \tag{4.4.227}$$

We can verify this by explicit computation:

$$\mathscr{D}_\rho \Omega'_{\mu\nu} = \partial_\rho \Omega'_{\mu\nu} + \Gamma'_\rho \Omega'_{\mu\nu} - \Omega'_{\mu\nu} \Gamma'_\rho \tag{4.4.228}$$

$$= \partial_\rho (U \Omega_{\mu\nu} U^{-1}) + (U \Gamma_\rho U^{-1} - \partial_\rho U U^{-1}) U \Omega_{\mu\nu} U^{-1} - U \Omega_{\mu\nu} U^{-1} (U \Gamma_\rho U^{-1} + U \partial_\rho U^{-1}) \tag{4.4.229}$$

$$= \partial_\rho U \Omega_{\mu\nu} U^{-1} + U \partial_\rho \Omega_{\mu\nu} U^{-1} + U \Omega_{\mu\nu} \partial_\rho U^{-1} + U \Gamma_\rho \Omega_{\mu\nu} U^{-1} - \partial_\rho U \Omega_{\mu\nu} U^{-1} - U \Omega_{\mu\nu} \Gamma_\rho U^{-1} - U \Omega_{\mu\nu} \partial_\rho U^{-1} \tag{4.4.230}$$

$$= U(\partial_\rho \Omega_{\mu\nu} + \Gamma_\rho \Omega_{\mu\nu} - \Omega_{\mu\nu} \Gamma_\rho) U^{-1} \tag{4.4.231}$$

$$= U \mathscr{D}_\rho \Omega_{\mu\nu} U^{-1}. \tag{4.4.232}$$

Using the covariant derivative of the curvature tensor we can see it immediately satisfies the property:

$$\epsilon^{\mu\nu\rho\sigma} \mathscr{D}_\nu \Omega_{\rho\sigma} = \epsilon^{\mu\nu\rho\sigma} (\partial_\nu \Omega_{\rho\sigma} + \Gamma_\nu \Omega_{\rho\sigma} - \Omega_{\rho\sigma} \Gamma_\nu) \tag{4.4.233}$$

$$= \epsilon^{\mu\nu\rho\sigma} (\partial_\nu (\partial_\rho \Gamma_\sigma - \partial_\sigma \Gamma_\rho + \Gamma_\rho \Gamma_\sigma - \Gamma_\sigma \Gamma_\rho) + [\Gamma_\nu, \partial_\rho \Gamma_\sigma - \partial_\sigma \Gamma_\rho + \Gamma_\rho \Gamma_\sigma - \Gamma_\sigma \Gamma_\rho]) \tag{4.4.234}$$

$$= 2\epsilon^{\mu\nu\rho\sigma} (\partial_\nu (\partial_\rho \Gamma_\sigma + \Gamma_\rho \Gamma_\sigma) + [\Gamma_\nu, \partial_\rho \Gamma_\sigma + \Gamma_\rho \Gamma_\sigma]) \tag{4.4.235}$$

$$= 2\epsilon^{\mu\nu\rho\sigma} \left(\partial^2_{\nu\rho} \Gamma_\sigma + \partial_\nu \Gamma_\rho \Gamma_\sigma + \Gamma_\rho \partial_\nu \Gamma_\sigma + \Gamma_\nu \partial_\rho \Gamma_\sigma - \partial_\rho \Gamma_\sigma \Gamma_\nu + \Gamma_\nu \Gamma_\rho \Gamma_\sigma - \Gamma_\rho \Gamma_\sigma \Gamma_\nu\right) \tag{4.4.236}$$

$$= 2\epsilon^{\mu\nu\rho\sigma} (0 + \partial_\nu \Gamma_\rho \Gamma_\sigma + \Gamma_\rho \partial_\nu \Gamma_\sigma - \Gamma_\rho \partial_\nu \Gamma_\sigma - \partial_\nu \Gamma_\rho \Gamma_\sigma + \Gamma_\nu \Gamma_\rho \Gamma_\sigma - \Gamma_\nu \Gamma_\rho \Gamma_\sigma) \tag{4.4.237}$$

$$= 0. \tag{4.4.238}$$

This is called the Bianchi identity and it is equivalent to the fact that there exists a $\Gamma_\mu$ so that $\Omega_{\mu\nu}$ can be written as in (4.4.168).

In order to develop a Lagrangian density for our theory with gauge symmetry, we will need to be able to list all of the local gauge invariants which can be made out of our connection. We know that all of the gauge invariants can be made out of the Wilson loops (because the connection itself can be recovered from the Wilson loops up to gauge transformation) however the Wilson loop is non-local. We found above that the curvature tensor has quite a nice gauge transformation. Moreover, any product of factors of the curvature tensor shares that same gauge transformation property:

$$(\Omega_{\mu_1\nu_1} ... \Omega_{\mu_N\nu_N})' = U \Omega_{\mu_1\nu_1} U^{-1} ... U \Omega_{\mu_N\nu_N} U^{-1} = U \Omega_{\mu_1\nu_1} ... \Omega_{\mu_N\nu_N} U^{-1}. \tag{4.4.239}$$

Consequently, the trace of any such quantity is automatically both local and gauge invariant:

$$\mathrm{Tr}(\Omega'_{\mu_1\nu_1} ... \Omega'_{\mu_N\nu_N}) = \mathrm{Tr}(U \Omega_{\mu_1\nu_1} ... \Omega_{\mu_N\nu_N} U^{-1}) = \mathrm{Tr}(\Omega_{\mu_1\nu_1} ... \Omega_{\mu_N\nu_N}). \tag{4.4.240}$$

Therefore, such a term is a candidate (at least with regards to gauge invariance) for a term we may add to our Lagrangian density. Because we are looking for the lowest order Lagrangian density for our theory, we would like to consider one which depends on the fields $\Gamma^a_\mu$ and their first derivatives $\partial_\nu \Gamma^a_\mu$. Suppose we have some gauge invariant function made out of the connection and its first derivatives $f(\Gamma, \partial\Gamma)$. Then, because it is gauge invariant it is by definition the same in any gauge:

$$f(\Gamma, \partial\Gamma) = f(\Gamma', \partial\Gamma'). \tag{4.4.241}$$

Now consider a gauge transformation $U(X)$ so that at a particular point $X_0$ the gauge transformation parameters $\theta^a(X)$ vanish. Writing $U$ in terms of transformation parameters $\theta^a(X)$ and expanding in the neighborhood of our point:

$$U(X) = e^{\theta^a(X) T_a} \tag{4.4.242}$$

$$\theta^a(X) = \theta^a(X_0) + \partial_\mu \theta^a|_{X_0} (x^\mu - x^\mu_0) + \mathcal{O}((x - x_0)^2). \tag{4.4.243}$$

The requirement that the gauge transformation parameters vanish at $X_0$ then implies:

$$U(0) = e^{\theta^a(X_0) T_a} = \mathbb{1}, \qquad \theta^a(X) = \partial_\mu \theta^a|_{X_0} (x^\mu - x^\mu_0) + \mathcal{O}((x - x_0)^2). \tag{4.4.244}$$

Thus, under this transformation, the connection fields at the point change by:

$$\Gamma'^a_\mu(X_0) = \left(e^{\theta^c f_c}\right)^a_{\ b}\Big|_{X_0} \Gamma^b_\mu(X_0) - \left(\frac{e^{\theta^c f_c} - \mathbb{1}}{\theta^c f_c}\right)^a_{\ b}\Bigg|_{X_0} \partial_\mu \theta^b\big|_{X_0} \tag{4.4.245}$$

$$= \delta^a_b \Gamma^b_\mu(X_0) - \delta^a_b\ \partial_\mu\theta^b\big|_{X_0} = \Gamma^a_\mu(X_0) - \partial_\mu\theta^a\big|_{X_0}\,. \tag{4.4.246}$$

Because we can choose the functions $\theta^a(X)$ at will, we can always choose them so that at $X = 0$:

$$\partial_\mu\theta^a\big|_0 = \Gamma^a_\mu(0). \tag{4.4.247}$$

(One simple choice is just $\theta^a(X) = \Gamma^a_\mu(0)x^\mu$.) By making such a choice, the transformed connection is automatically 0 at the point. Thus, it is always possible to transform the connection to be 0 at any particular point. As such, there are no local gauge invariants which can be made algebraically (meaning without derivatives) out of the connection alone. After applying such a gauge transformation we have $\Gamma'^a_\mu(X_0) = 0$ and so by (4.4.241) our gauge invariant function $f$ becomes (at the point $X_0$):

$$f(\Gamma, \partial\Gamma) = f(0, \partial\Gamma'). \tag{4.4.248}$$

Now let us apply a second gauge transformation and see what it can do to the derivatives of the connection at $X_0$. In order to not spoil the condition $\Gamma^a_\mu(X_0) = 0$, we must take that for this transformation $\partial_\mu\theta^a\big|_{X_0} = 0$. Thus, expanding the transformation parameters to lowest order we have:

$$\theta^a(X) = \theta^a(X_0) + \partial_\mu\theta^a\big|_{X_0}(x^\mu - x^\mu_0) + \frac{1}{2}\,\partial^2_{\mu\nu}\theta^a\big|_{X_0}(x^\mu - x^\mu_0)(x^\nu - x^\nu_0) + \mathcal{O}((x-x_0)^3) \tag{4.4.249}$$

$$= 0 + 0 + \frac{1}{2}\,\partial^2_{\mu\nu}\theta^a\big|_{X_0}(x^\mu - x^\mu_0)(x^\nu - x^\nu_0) + \mathcal{O}((x-x_0)^3) \tag{4.4.250}$$

Thus in the neighborhood of $X_0$, the connection is:

$$\Gamma'^a_\mu(X) = \left(e^{\theta^c(X) f_c}\right)^a_{\ b}\Gamma^b_\mu - \left(\frac{e^{\theta^c(X) f_c} - \mathbb{1}}{\theta^c(X) f_c}\right)^a_{\ b}\partial_\mu\theta^b(X) \tag{4.4.251}$$

$$= \left(\delta^a_b + \mathcal{O}((x-x_0)^3)\right)\left(\Gamma^b_\mu(X_0) + \partial_\nu\Gamma^b_\mu\big|_{X_0}(x^\nu - x^\nu_0) + \mathcal{O}((x-x_0)^2)\right)$$
$$- \left(\delta^a_b + \mathcal{O}((x-x_0))\right)\partial_\mu\left(\frac{1}{2}\,\partial^2_{\alpha\beta}\theta^b\big|_0(x^\alpha - x^\alpha_0)(x^\beta - x^\beta_0) + \mathcal{O}(x^3)\right) \tag{4.4.252}$$

$$= \delta^a_b\ \partial_\nu\Gamma^b_\mu\big|_{X_0}(x^\nu - x^\nu_0) - \delta^a_b\ \partial^2_{\mu\alpha}\theta^b\big|_{X_0}(x^\alpha - x^\alpha_0) + \mathcal{O}((x-x_0)^2) \tag{4.4.253}$$

$$= \partial_\nu\Gamma^a_\mu\big|_{X_0}(x^\nu - x^\nu_0) - \partial^2_{\mu\nu}\theta^a\big|_{X_0}(x^\nu - x^\nu_0) + \mathcal{O}((x-x_0)^2). \tag{4.4.254}$$

This gives the derivative in the neighborhood of $X_0$ as:

$$\partial_\nu\Gamma'^a_\mu(X) = \partial_\nu\Gamma^a_\mu\big|_{X_0} - \partial^2_{\mu\nu}\theta^a\big|_{X_0} + \mathcal{O}((x-x_0)). \tag{4.4.255}$$

Finally then the derivative of the connection at the point is given by:

$$\partial_\nu\Gamma'^a_\mu\big|_{X_0} = \partial_\nu\Gamma^a_\mu\big|_{X_0} - \partial^2_{\mu\nu}\theta^a\big|_{X_0}\,. \tag{4.4.256}$$

Because $\partial^2_{\mu\nu}\theta^a$ is symmetric with respect to $\mu$ and $\nu$, there are not enough independent variables here to transform all of the first derivatives of the connection to 0. We can see that the anti-symmetric combination of derivatives is unaffected by this transformation:

$$\partial_\mu\Gamma'^a_\nu(X_0) - \partial_\nu\Gamma'^a_\mu(X_0) = \partial_\mu\Gamma^a_\nu\big|_{X_0} - \partial^2_{\mu\nu}\theta^a\big|_{X_0} - \partial_\nu\Gamma^a_\mu\big|_{X_0} + \partial^2_{\mu\nu}\theta^a\big|_{X_0} \tag{4.4.257}$$

$$= \partial_\mu\Gamma^a_\nu\big|_{X_0} - \partial_\nu\Gamma^a_\mu\big|_{X_0}\,. \tag{4.4.258}$$

This is not surprising as because $\Gamma^a_\mu(X_0) = \Gamma'^a_\mu(X_0) = 0$, we have that the components of the curvature at $X_0$ are given by:

$$\Omega^a_{\mu\nu}(X_0) = \partial_\mu\Gamma^a_\nu\big|_{X_0} - \partial_\nu\Gamma^a_\mu\big|_{X_0} + f_{bc}{}^a\ \Gamma^b_\mu\big|_{X_0}\Gamma^c_\nu\big|_{X_0} = \partial_\mu\Gamma^a_\nu\big|_{X_0} - \partial_\nu\Gamma^a_\mu\big|_{X_0} \tag{4.4.259}$$

which just amounts to showing that the curvature is unchanged by this transformation. On the other hand, the symmetric combination of derivatives is affected by the transformation:

$$\partial_\mu\Gamma'^a_\nu(X_0) + \partial_\nu\Gamma'^a_\mu(X_0) = \partial_\mu\Gamma^a_\nu\big|_{X_0} - \partial^2_{\mu\nu}\theta^a\big|_{X_0} + \partial_\nu\Gamma^a_\mu\big|_{X_0} - \partial^2_{\mu\nu}\theta^a\big|_{X_0} \tag{4.4.260}$$

Hence, by choosing:

$$\partial^2_{\mu\nu}\theta^a\big|_{X_0} = \frac{1}{2}\left(\partial_\mu\Gamma^a_\nu\big|_{X_0} + \partial_\nu\Gamma^a_\mu\big|_{X_0}\right) \tag{4.4.261}$$

we can transform the symmetric combination of first derivatives to 0. Then, once we have made such a gauge transformation, we have that the derivatives of the connection are simply given by:

$$\partial_\mu \Gamma'_\nu(X_0) = \frac{\partial_\mu \Gamma'_\nu(X_0) + \partial_\nu \Gamma'_\mu(X_0)}{2} + \frac{\partial_\mu \Gamma'_\nu(X_0) - \partial_\nu \Gamma'_\mu(X_0)}{2} = 0 + \frac{1}{2}\Omega_{\mu\nu}(X_0). \tag{4.4.262}$$

So, our gauge invariant function $f$ at $X_0$ then becomes:

$$f(\Gamma, \partial\Gamma) = f(0, \partial\Gamma') = f\left(0, \frac{1}{2}\Omega\right). \tag{4.4.263}$$

The expression on the right hand side depends only on the curvature tensor and under any further gauge transformations any expression of only the curvature tensor will remain a function of only the curvature tensor. Thus, any gauge invariant function of the connection and its first derivatives is necessarily actually just a function of the curvature tensor alone. When constructing our Lagrangian density then we will simply need to construct gauge invariant and Lorentz invariant functions of the curvature tensor. So, the curvature tensor is quite important. However, complete knowledge of the curvature tensor everywhere is not enough information to be able to reconstruct the connection [34].

### 4.4.3 Possible Lagrangians

Because any gauge invariant function of the connection and its first derivative is necessarily just a function of the curvature tensor, we know then that the Lagrangian density for the connection is simply then a function of the curvature tensor: $\mathfrak{L} = \mathfrak{L}(\Omega)$. As we generally do, we now consider an expansion of the Lagrangian density as a function of the curvature tensor valid for "weak fields" (meaning low curvature). Then, we have:

$$\mathfrak{L} = A + B_a^{\mu\nu}\Omega_{\mu\nu}^a + C_{ab}^{\mu\nu\rho\sigma}\Omega_{\mu\nu}^a\Omega_{\rho\sigma}^b + \mathcal{O}(\Omega^3) \tag{4.4.264}$$

for some constants $A$, $B_a^{\mu\nu}$, and $C_{ab}^{\mu\nu\rho\sigma}$. The coefficient $A$ can be removed as it is simply a surface term: $A = \partial_0(At)$. Because the curvature tensor is in fact a rank 2 tensor, in order for the second term to have Lorentz symmetry it must be the case that under a Lorentz transformation:

$$B_a^{\mu\nu}\Omega'^a_{\mu\nu} = B_a^{\mu\nu}\Omega_{\mu\nu}^a \tag{4.4.265}$$

$$\implies L^\mu{}_\rho L^\nu{}_\sigma B_a^{\rho\sigma} = B_a^{\mu\nu}. \tag{4.4.266}$$

Thus, $B_a^{\mu\nu}$ must be Lorentz invariant. The all Lorentz invariant rank-2 tensors are proportional to the Minkowski matrix. Therefore, for some coefficients $B_a$:

$$B_a^{\mu\nu} = B_a\eta^{\mu\nu}. \tag{4.4.267}$$

However, then the $B$ term in the Lagrangian density vanishes due to the anti-symmetry of the curvature tensor:

$$B_a^{\mu\nu}\Omega_{\mu\nu}^a = B_a\eta^{\mu\nu}\Omega_{\mu\nu}^a = -B_a\eta^{\mu\nu}\Omega_{\mu\nu}^a = -B_a^{\mu\nu}\Omega_{\mu\nu}^a \tag{4.4.268}$$

$$\implies B_a^{\mu\nu}\Omega_{\mu\nu}^a = 0. \tag{4.4.269}$$

Thus, our Lagrangian density is reduced to simply the $C$ term. Then, again owing to the Lorentz transformation properties of the curvature tensor we must have:

$$C_{ab}^{\mu\nu\rho\sigma}\Omega'^a_{\mu\nu}\Omega'^b_{\rho\sigma} = C_{ab}^{\mu\nu\rho\sigma}\Omega_{\mu\nu}^a\Omega_{\rho\sigma}^b \tag{4.4.270}$$

$$\implies L^\mu{}_\alpha L^\nu{}_\beta L^\rho{}_\gamma L^\sigma{}_\lambda C_{ab}^{\alpha\beta\gamma\lambda} = C_{ab}^{\mu\nu\rho\sigma}. \tag{4.4.271}$$

And so, the $C$ coefficients must form the components of a Lorentz invariant rank-4 tensor. Such a tensor is necessarily a linear combination of products of the Minkowski matrix and the Levi-Civita tensor. That is, for some coefficients:

$$C_{ab}^{\mu\nu\rho\sigma} = C_{ab}^1\eta^{\mu\nu}\eta^{\rho\sigma} + C_{ab}^2\eta^{\mu\rho}\eta^{\nu\sigma} + C_{ab}^3\eta^{\mu\sigma}\eta^{\nu\rho} + C_{ab}^4\epsilon^{\mu\nu\rho\sigma}. \tag{4.4.272}$$

Then, the resulting Lagrangian density is (using the anti-symmetry of the curvature tensor):

$$\mathfrak{L} = C_{ab}^1\eta^{\mu\nu}\eta^{\rho\sigma}\Omega_{\mu\nu}^a\Omega_{\rho\sigma}^b + C_{ab}^2\eta^{\mu\rho}\eta^{\nu\sigma}\Omega_{\mu\nu}^a\Omega_{\rho\sigma}^b + C_{ab}^3\eta^{\mu\sigma}\eta^{\nu\rho}\Omega_{\mu\nu}^a\Omega_{\rho\sigma}^b + C_{ab}^4\epsilon^{\mu\nu\rho\sigma}\Omega_{\mu\nu}^a\Omega_{\rho\sigma}^b \tag{4.4.273}$$

$$= 0 + (C_{ab}^2 - C_{ab}^3)\eta^{\mu\rho}\eta^{\nu\sigma}\Omega_{\mu\nu}^a\Omega_{\rho\sigma}^b + C_{ab}^4\epsilon^{\mu\nu\rho\sigma}\Omega_{\mu\nu}^a\Omega_{\rho\sigma}^b. \tag{4.4.274}$$

So, renaming the coefficients which appear in the Lagrangian density we can see that it is of the form:

$$\mathfrak{L} = A_{ab}\eta^{\mu\rho}\eta^{\nu\sigma}\Omega_{\mu\nu}^a\Omega_{\rho\sigma}^b + B_{ab}\epsilon^{\mu\nu\rho\sigma}\Omega_{\mu\nu}^a\Omega_{\rho\sigma}^b. \tag{4.4.275}$$

These coefficients $A_{ab}$ and $B_{ab}$ still have to be selected so that the resulting Lagrangian density is gauge invariant. Using (4.4.209) then we must have:

$$A_{ab}\eta^{\mu\rho}\eta^{\nu\sigma}\Omega'^{a}_{\mu\nu}\Omega'^{b}_{\rho\sigma} = A_{ab}\eta^{\mu\rho}\eta^{\nu\sigma}\Omega^{a}_{\mu\nu}\Omega^{b}_{\rho\sigma}, \qquad B_{ab}\epsilon^{\mu\nu\rho\sigma}\Omega'^{a}_{\mu\nu}\Omega'^{b}_{\rho\sigma} = B_{ab}\epsilon^{\mu\nu\rho\sigma}\Omega^{a}_{\mu\nu}\Omega^{b}_{\rho\sigma} \tag{4.4.276}$$

$$\implies A_{ab}\eta^{\mu\rho}\eta^{\nu\sigma}\left(e^{\theta^c f_c}\right)^{a}{}_{d}\left(e^{\theta^e f_e}\right)^{b}{}_{f}\Omega^{d}_{\mu\nu}\Omega^{f}_{\rho\sigma} = A_{ab}\eta^{\mu\rho}\eta^{\nu\sigma}\Omega^{a}_{\mu\nu}\Omega^{b}_{\rho\sigma} \tag{4.4.277}$$

$$\implies A_{cd}\left(e^{\theta^e f_e}\right)^{c}{}_{a}\left(e^{\theta^f f_f}\right)^{d}{}_{b} = A_{ab} \tag{4.4.278}$$

$$B_{ab}\epsilon^{\mu\nu\rho\sigma}\left(e^{\theta^e f_e}\right)^{a}{}_{c}\left(e^{\theta^f f_f}\right)^{b}{}_{d}\Omega^{c}_{\mu\nu}\Omega^{d}_{\rho\sigma} = B_{ab}\epsilon^{\mu\nu\rho\sigma}\Omega^{a}_{\mu\nu}\Omega^{b}_{\rho\sigma} \tag{4.4.279}$$

$$\implies B_{cd}\left(e^{\theta^e f_e}\right)^{c}{}_{a}\left(e^{\theta^f f_f}\right)^{d}{}_{b} = B_{ab}. \tag{4.4.280}$$

By (4.4.278) and (4.4.280) we see that $A_{ab}$ and $B_{ab}$ must satisfy the same conditions in order to have gauge invariance. Taking the derivative of these equations with respect to $\theta$ at $\theta = 0$ we see that gauge invariance requires:

$$f_{da}{}^{c}A_{cb} + A_{ac}f_{db}{}^{c} = 0, \qquad f_{da}{}^{c}B_{cb} + B_{ac}f_{db}{}^{c} = 0. \tag{4.4.281}$$

The set of all $A_{ab}$ (or $B_{ab}$) satisfying this equation will depend on the gauge group as it depends on the structure coefficients. As well, one can see that any anti-symmetric part of $A_{ab}$ or $B_{ab}$ fails to contribute to the Lagrangian density and so they are symmetric in exchanging their indices automatically.

It turns out the $B$ term in the Lagrangian density can actually be simplified. Expanding out the $B$ term we see:

$$B_{ab}\epsilon^{\mu\nu\rho\sigma}\Omega^{a}_{\mu\nu}\Omega^{b}_{\rho\sigma} = B_{ab}\epsilon^{\mu\nu\rho\sigma}(\partial_\mu\Gamma^{a}_{\nu} - \partial_\nu\Gamma^{a}_{\mu} + f_{cd}{}^{a}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu})(\partial_\rho\Gamma^{b}_{\sigma} - \partial_\sigma\Gamma^{b}_{\rho} + f_{ef}{}^{b}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma}) \tag{4.4.282}$$

$$= B_{ab}\epsilon^{\mu\nu\rho\sigma}(2\partial_\mu\Gamma^{a}_{\nu} + f_{cd}{}^{a}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu})(2\partial_\rho\Gamma^{b}_{\sigma} + f_{ef}{}^{b}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma}) \tag{4.4.283}$$

$$= B_{ab}\epsilon^{\mu\nu\rho\sigma}(4\partial_\mu\Gamma^{a}_{\nu}\partial_\rho\Gamma^{b}_{\sigma} + 4f_{cd}{}^{a}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\partial_\rho\Gamma^{b}_{\sigma} + f_{cd}{}^{a}f_{ef}{}^{b}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma}). \tag{4.4.284}$$

Looking at the first of these terms:

$$B_{ab}\epsilon^{\mu\nu\rho\sigma}\partial_\mu\Gamma^{a}_{\nu}\partial_\rho\Gamma^{b}_{\sigma} = \partial_\mu\left(B_{ab}\epsilon^{\mu\nu\rho\sigma}\Gamma^{a}_{\nu}\partial_\rho\Gamma^{b}_{\sigma}\right) - B_{ab}\epsilon^{\mu\nu\rho\sigma}\Gamma^{a}_{\nu}\partial^{2}_{\mu\rho}\Gamma^{b}_{\sigma}. \tag{4.4.285}$$

The second of these is zero because the Levi-Civita symbol is anti-symmetric under interchanging $\mu$ and $\rho$ while the mixed partial $\partial^{2}_{\mu\rho}\Gamma^{b}_{\sigma}$ is symmetric under that same interchange. Therefore:

$$B_{ab}\epsilon^{\mu\nu\rho\sigma}\partial_\mu\Gamma^{a}_{\nu}\partial_\rho\Gamma^{b}_{\sigma} = \partial_\mu\left(B_{ab}\epsilon^{\mu\nu\rho\sigma}\Gamma^{a}_{\nu}\partial_\rho\Gamma^{b}_{\sigma}\right). \tag{4.4.286}$$

So, this piece is a total divergence. Now let us look at the second term in (4.4.284):

$$B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\partial_\rho\Gamma^{b}_{\sigma} = \partial_\rho\left(B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{b}_{\sigma}\right) - B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}\Gamma^{b}_{\sigma}\partial_\rho(\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}) \tag{4.4.287}$$

$$= \partial_\rho\left(B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{b}_{\sigma}\right) + f_{cb}{}^{a}B_{ad}\epsilon^{\mu\nu\rho\sigma}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\partial_\rho\Gamma^{b}_{\sigma} + f_{bd}{}^{a}B_{ac}\epsilon^{\mu\nu\rho\sigma}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\partial_\rho\Gamma^{b}_{\sigma}. \tag{4.4.288}$$

Now using (4.4.281):

$$B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\partial_\rho\Gamma^{b}_{\sigma} = \partial_\rho\left(B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{b}_{\sigma}\right) - f_{cd}{}^{a}B_{ab}\epsilon^{\mu\nu\rho\sigma}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\partial_\rho\Gamma^{b}_{\sigma} - f_{cd}{}^{a}B_{ab}\epsilon^{\mu\nu\rho\sigma}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\partial_\rho\Gamma^{b}_{\sigma} \tag{4.4.289}$$

$$\implies 3B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\partial_\rho\Gamma^{b}_{\sigma} = \partial_\rho\left(B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{b}_{\sigma}\right) \tag{4.4.290}$$

$$\implies B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\partial_\rho\Gamma^{b}_{\sigma} = \partial_\mu\left(\frac{1}{3}B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}\Gamma^{c}_{\rho}\Gamma^{d}_{\sigma}\Gamma^{b}_{\nu}\right) \tag{4.4.291}$$

Thus, this piece is also a total divergence. Finally looking at the last term in (4.4.284) and using (4.4.281):

$$B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}f_{ef}{}^{b}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma} = -B_{bf}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}f_{ea}{}^{b}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma}. \tag{4.4.292}$$

Now using the Jacobi-identity for the structure constants (as derived in Orig. problem 1.31):

$$f_{ad}{}^{e}f_{bc}{}^{d} + f_{bd}{}^{e}f_{ca}{}^{d} + f_{cd}{}^{e}f_{ab}{}^{d} = 0 \tag{4.4.293}$$

we arrive at:

$$B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}f_{ef}{}^{b}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma} = B_{bf}\epsilon^{\mu\nu\rho\sigma}(f_{ca}{}^{b}f_{de}{}^{a} + f_{da}{}^{b}f_{ec}{}^{a})\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma} \tag{4.4.294}$$

$$= -f_{cf}{}^{b}f_{de}{}^{a}B_{ab}\epsilon^{\mu\nu\rho\sigma}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma} - f_{ec}{}^{a}f_{df}{}^{b}B_{ab}\epsilon^{\mu\nu\rho\sigma}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma} \tag{4.4.295}$$

$$= -B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}f_{ef}{}^{b}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma} - B_{ab}\epsilon^{\mu\nu\rho\sigma}f_{cd}{}^{a}f_{ef}{}^{b}\Gamma^{d}_{\nu}\Gamma^{e}_{\rho}\Gamma^{c}_{\mu}\Gamma^{f}_{\sigma}. \tag{4.4.296}$$

$$\implies 3B_{ab}\epsilon^{\mu\nu\rho\sigma} f_{cd}{}^{a} f_{ef}{}^{b}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma} = 0 \tag{4.4.297}$$

$$\implies B_{ab}\epsilon^{\mu\nu\rho\sigma} f_{cd}{}^{a} f_{ef}{}^{b}\Gamma^{c}_{\mu}\Gamma^{d}_{\nu}\Gamma^{e}_{\rho}\Gamma^{f}_{\sigma} = 0. \tag{4.4.298}$$

Thus, the $B$ term in the Lagrangian density can be rewritten as:

$$B_{ab}\epsilon^{\mu\nu\rho\sigma}\Omega^{a}_{\mu\nu}\Omega^{b}_{\rho\sigma} = \partial_{\mu}\left(4B_{ab}\epsilon^{\mu\nu\rho\sigma}\Gamma^{a}_{\nu}\partial_{\rho}\Gamma^{b}_{\sigma} + \frac{4}{3}B_{ab}\epsilon^{\mu\nu\rho\sigma} f_{cd}{}^{a}\Gamma^{c}_{\rho}\Gamma^{d}_{\sigma}\Gamma^{b}_{\nu}\right). \tag{4.4.299}$$

Hence, because this term is a total derivative it does not affect the equations of motion and can be dropped from the Lagrangian density. So, our final lowest order Lagrangian density for our theory of the connection with all of the necessary symmetries is:

$$\mathfrak{L} = A_{ab}\eta^{\mu\rho}\eta^{\nu\sigma}\Omega^{a}_{\mu\nu}\Omega^{b}_{\rho\sigma} \tag{4.4.300}$$

where these $A_{ab}$ satisfy $A_{ab} = A_{ba}$ and (4.4.281).

### 4.4.4 The Simplest Gauge Theories

As the simplest example of a space with internal degrees of freedom, consider a real two dimensional internal space with semi-inner product given by the standard dot product in two dimensions. That is, a field $\Phi(X)$ on this space has two components:

$$\Phi(X) = \phi^{1}(X)\xi_{1}(X) + \phi^{2}(X)\xi_{2}(X), \qquad \phi(X) = \begin{pmatrix}\phi^{1}(X)\\ \phi^{2}(X)\end{pmatrix} \tag{4.4.301}$$

and between two fields the inner product is given by:

$$(\Phi_{1}, \Phi_{2}) = (\phi_{1})^{1}(\phi_{2})^{1} + (\phi_{1})^{2}(\phi_{2})^{2}. \tag{4.4.302}$$

Then, the matrix M is given as:

$$\mathrm{M} = \begin{pmatrix}\mathrm{M}_{11} & \mathrm{M}_{12}\\ \mathrm{M}_{21} & \mathrm{M}_{22}\end{pmatrix} = \begin{pmatrix}(\xi_{1},\xi_{1}) & (\xi_{1},\xi_{2})\\ (\xi_{2},\xi_{1}) & (\xi_{2},\xi_{2})\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \mathbb{1}. \tag{4.4.303}$$

As well, a basis generator $T$ must then satisfy (4.4.82) as:

$$T^{\dagger} + T = 0. \tag{4.4.304}$$

Because $\phi$ is real, $U$ must be in order for it to remain so under gauge transformation and thus $T$ must be as well. Giving $T$ an explicit matrix representation then:

$$T = \begin{pmatrix}a & b\\ c & d\end{pmatrix} \tag{4.4.305}$$

$$T^{\dagger} + T = \begin{pmatrix}2a & b+c\\ b+c & 2d\end{pmatrix} \overset{!}{=} 0 \tag{4.4.306}$$

$$\implies a = d = 0, \qquad c = -b \tag{4.4.307}$$

$$\implies T = b\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}. \tag{4.4.308}$$

Thus, there is only one such generator which is linearly independent. By convention we choose $b = 1$ and so there is one group parameter $\theta(X)$ and a general gauge transformation is given by:

$$U(X) = e^{\theta(X)T}. \tag{4.4.309}$$

Immediately we have $T^{2} = -\mathbb{1}$ and so $U(X)$ can be simplified as:

$$U(X) = \sum_{n=0}^{\infty}\frac{\theta^{n}}{n!}T^{n} = \sum_{n=0}^{\infty}\frac{\theta^{2n}}{(2n)!}T^{2n} + \sum_{n=0}^{\infty}\frac{\theta^{2n+1}}{(2n+1)!}T^{2n+1} = \sum_{n=0}^{\infty}\frac{\theta^{2n}}{(2n)!}(-1)^{n}\mathbb{1} + \sum_{n=0}^{\infty}\frac{\theta^{2n+1}}{(2n+1)!}(-1)^{n}T \tag{4.4.310}$$

$$= \cos\theta\mathbb{1} + \sin\theta T = \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}. \tag{4.4.311}$$

Unsurprisingly, we can see that the transformations which leave the dot product unchanged are simply rotations. Then, because there is a single parameter, there is one connection field and one component of the curvature:

$$\Gamma_{\mu} = \Gamma^{1}_{\mu}T \tag{4.4.312}$$

$$\Omega_{\mu\nu} = \partial_\mu \Gamma^1_\nu T - \partial_\nu \Gamma^1_\mu T + f_{11}{}^1 \Gamma^1_\mu \Gamma^1_\nu T = \left(\partial_\mu \Gamma^1_\nu - \partial_\nu \Gamma^1_\mu\right) T. \tag{4.4.313}$$

Then, because there is only one generator the structure constants $f_{ab}{}^c$ are all 0 (and there is only one such constant, $f_{11}{}^1$ anyway). A group with all structure constants equal to 0 is called Abelian (otherwise it is non-Abelian). Then, the gauge transformation of the connection becomes (by (4.4.126)):

$$\Gamma'^1_\mu = \Gamma^1_\mu - \partial_\mu \theta. \tag{4.4.314}$$

By (4.4.300) there is only one $A_{ab}$ component $A_{11}$ and so it is automatically symmetric. As well, because all of the structure constants are 0 (4.4.281) is satisfied automatically. So, the Lagrangian density for $\Gamma^1_\mu$ is given by:

$$\mathfrak{L} = A_{11} \eta^{\mu\rho} \eta^{\nu\sigma} (\partial_\mu \Gamma^1_\nu - \partial_\nu \Gamma^1_\mu)(\partial_\rho \Gamma^1_\sigma - \partial_\sigma \Gamma^1_\rho) \tag{4.4.315}$$

$$= A_{11} (\partial^\mu \Gamma^{1\nu} - \partial^\nu \Gamma^{1\mu})(\partial_\mu \Gamma^1_\nu - \partial_\nu \Gamma^1_\mu). \tag{4.4.316}$$

This is looking very similar to the covariant Lagrangian density for Maxwell's equations. We can make this more transparent by defining the complex valued field:

$$\varphi \overset{\text{def}}{=} \frac{\phi^1 + i\phi^2}{\sqrt{2}}. \tag{4.4.317}$$

The $\sqrt{2}$ is conventional. Then, the semi-inner product becomes:

$$(\Phi_1, \Phi_2) = \varphi_1^* \varphi_2 + \varphi_2^* \varphi_1. \tag{4.4.318}$$

Computing the gauge transformation of $\phi$:

$$\phi' = U\phi = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix} = \begin{pmatrix} \phi^1 \cos\theta + \phi^2 \sin\theta \\ -\phi^1 \sin\theta + \phi^2 \cos\theta \end{pmatrix} \tag{4.4.319}$$

we can see that the gauge transformation of $\varphi$ is:

$$\varphi' = \frac{\phi'^1 + i\phi'^2}{\sqrt{2}} = \frac{1}{\sqrt{2}} \left(\phi^1 \cos\theta + \phi^2 \sin\theta - i\phi^1 \sin\theta + i\phi^2 \cos\theta\right) \tag{4.4.320}$$

$$= \frac{1}{\sqrt{2}} \left(\cos\theta - i\sin\theta\right) \left(\phi^1 + i\phi^2\right) = e^{-i\theta} \varphi. \tag{4.4.321}$$

Hence, the rotation of the internal degree of freedom space is equivalent to this phase transformation of the complex valued field which is exactly the U(1) symmetry we have previously associated with the electromagnetic field. Computing the covariant derivative of $\phi$:

$$\mathscr{D}_\mu \phi = \begin{pmatrix} \mathscr{D}_\mu \phi^1 \\ \mathscr{D}_\mu \phi^2 \end{pmatrix} = \partial_\mu \phi + \Gamma_\mu \phi = \begin{pmatrix} \partial_\mu \phi^1 + \Gamma^1_\mu \phi^2 \\ \partial_\mu \phi^2 - \Gamma^1_\mu \phi^1 \end{pmatrix} \tag{4.4.322}$$

we can see that the covariant derivative of $\varphi$ is:

$$\mathscr{D}_\mu \varphi = \frac{1}{\sqrt{2}} \left(\mathscr{D}_\mu \phi^1 + i \mathscr{D}_\mu \phi^2\right) = \frac{1}{\sqrt{2}} \left(\partial_\mu \phi^1 + \Gamma^1_\mu \phi^2 + i\partial_\mu \phi^2 - i\Gamma^1_\mu \phi^1\right) \tag{4.4.323}$$

$$= \frac{1}{\sqrt{2}} \left(\partial_\mu - i\Gamma^1_\mu\right) \left(\phi^1 + i\phi^2\right) = \left(\partial_\mu - i\Gamma^1_\mu\right) \varphi. \tag{4.4.324}$$

Hence, recalling how the covariant derivative acts on a complex scalar field, we can identify that the connection $\Gamma^1_\mu$ is related to the vector potential by:

$$\Gamma^1_\mu = qA_\mu \tag{4.4.325}$$

where $q$ is the electric charge of the field $\varphi$. Then, the curvature becomes:

$$\Omega^1_{\mu\nu} = \partial_\mu \Gamma^1_\nu - \partial_\nu \Gamma^1_\mu = q\partial_\mu A_\nu - q\partial_\nu A_\mu = q\mathcal{F}_{\mu\nu}. \tag{4.4.326}$$

As well, the Lagrangian density becomes:

$$\mathfrak{L} = A_{11} q^2 \mathcal{F}^{\mu\nu} \mathcal{F}_{\mu\nu}. \tag{4.4.327}$$

And, as we already know, we have defined our units of charge so that the coefficient out front of this Lagrangian density is $-\frac{1}{4}$ for electromagnetism:

$$A_{11} q^2 = -\frac{1}{4} \implies A_{11} = -\frac{1}{4q^2}. \tag{4.4.328}$$

Hence the theory of a connection controlling rotational symmetry on a two dimensional internal space is precisely Maxwell's electromagnetism. Two dimensional rotations are equivalent to complex phases and this

is why we identified electromagnetism as the theory of $U(1)$ gauge symmetry. We can see then that (4.4.314) becomes:

$$qA'_\mu = qA_\mu - \partial_\mu \theta \tag{4.4.329}$$

and so we can see that the rotation parameters $\theta$ are related to the typical gauge transformation $f(X)$ of electromagnetism by:

$$\theta = -qf. \tag{4.4.330}$$

Using the complex representation, we can consider a 1 dimensional complex internal space instead of a 2 dimensional real one. Then, $U(X)$ on this space by (4.4.321) becomes:

$$U(X) = e^{-i\theta} = e^{iqf}. \tag{4.4.331}$$

As well, from this we can identify that in the complex representation $T = -i$. In the complex representation the Wilson loop of electromagnetism is quite nice:

$$W[\gamma] = \text{Tr } \mathcal{P} : e^{-\oint_\gamma \Gamma^1_\mu T dx^\mu} = e^{i \oint_\gamma A_\mu dx^\mu}. \tag{4.4.332}$$

The trace is trivial in 1 dimension and the path ordering is trivial because all of the generators commute with each other (as there is only one). The integrand in the exponential can be written quite nicely using the generalized Stokes' theorem. Consider an area $\mathbb{A}$ bounded by the loop $\gamma$. Then, by the generalized Stokes' theorem:

$$\oint_\gamma A_\mu dx^\mu = \oint_{\partial \mathbb{A}} A_\mu dx^\mu = \int_{\mathbb{A}} d(A_\mu dx^\mu) = \int_{\mathbb{A}} \partial_\nu A_\mu dx^\nu \wedge dx^\mu \tag{4.4.333}$$

$$= \frac{1}{2} \int_{\mathbb{A}} (\partial_\mu A_\nu - \partial_\nu A_\mu) dx^\mu \wedge dx^\nu = \frac{1}{2} \int_{\mathbb{A}} \mathcal{F}_{\mu\nu} dx^\mu \wedge dx^\nu. \tag{4.4.334}$$

So, the Wilson loop becomes:

$$W[\gamma] = e^{\frac{i}{2} \int_{\mathbb{A}} \mathcal{F}_{\mu\nu} dx^\mu \wedge dx^\nu}. \tag{4.4.335}$$

Thus, for U(1) the Wilson loop can be very cleanly expressed in terms of the curvature (this is not true for non-Abelian groups).

It is helpful to pick a reference charge (traditionally the electron) and define another notion of electric charge relative to it. The electron has negative charge and so define $\mathtt{e}$ as the magnitude of the electron charge: $q_e = -\mathtt{e}$. Then, for any other field, its charge $q$ can be written as $Q\mathtt{e}$ for some $Q$. This $\mathtt{e}$ is called the electromagnetic coupling constant, electron charge (magnitude), or the "elementary charge" and $Q$ is then how strongly that particular field couples. This organization of terms is nice for many calculations and in making a similarity of equations between this U(1) gauge theory and more complicated gauge theories, as we will see in the next subsection. This $Q$ is also called the charge of the field (confusingly, since that is already the name of $q$). $Q$ is really the charge of the field in units of the (magnitude of the) electron charge. In the context of field theory, $Q$ is typically just called the electric charge and one has to infer from context whether $q$ or $Q$ is intended. If it is unclear, one should call $q$ the classical electric charge and $Q$ the charge quantum number.

The other extremely standard example of a space with internal degrees of freedom is an $N$ dimensional complex vector space with semi-inner product given by the standard complex inner product. That is, a field $\Phi(X)$ on this space has $N$ complex components:

$$\Phi(X) = \phi^a(X)\xi_a(X), \qquad \phi(X) = \begin{pmatrix} \phi^1(X) \\ ... \\ \phi^N(X) \end{pmatrix} \tag{4.4.336}$$

and between two fields the semi-inner product is given by:

$$(\Phi_1, \Phi_2) = \left((\phi_1)^1\right)^* (\phi_2)^1 + ... + \left((\phi_1)^N\right)^* (\phi_2)^N = \phi_1^\dagger \phi_2. \tag{4.4.337}$$

Then, the components of the matrix M are given as:

$$\mathrm{M}_{ab} = (\xi_a, \xi_b) = \sum_{\underline{n}=1}^{N} \delta^{\underline{n}}_a \delta^{\underline{n}}_b = \delta_{ab} \implies \mathrm{M} = \mathbb{1}. \tag{4.4.338}$$

Then, (4.4.82) becomes:

$$T_a^\dagger + T_a = 0. \tag{4.4.339}$$

This automatically means that $U$ is unitary:

$$U^\dagger = e^{\theta^a T_a^\dagger} = e^{-\theta^a T_a} = U^{-1}. \tag{4.4.340}$$

So, this group is the group U(N) of $N$x$N$ unitary matrices. There are $N^2$ linearly independent (over the reals) matrices $T_a$ satisfying (4.4.339) as any of the $N$ diagonal elements are required by this to be purely imaginary and any of the $N(N-1)/2$ lower diagonal elements are required to be the negative complex conjugate of the $N(N-1)/2$ upper diagonal complex elements; each upper diagonal complex element has 2 real numbers of degrees of freedom giving $N + 2N(N-1)/2 = N^2$ independent matrices. Of these $N^2$ matrices, we can select out the trace by subtracting the trace off of all of them giving $N^2 - 1$ independent such matrices (as the last diagonal element is no longer free) and considering the matrix $-i\mathbb{1}$ separately (which has a non-zero trace, completing the full $N^2$). The matrix $-i\mathbb{1}$ simply again gives U(1) transformations when exponentiated and commutes with the other $N^2 - 1$ linearly independent $T_a$ and so the group $U(N)$ factors into $U(1)$ and $SU(N)$, the group generated by the remaining $N^2 - 1$ generators. Because all of these $T_a$ are traceless, the determinant of $U(X)$ is automatically 1 and so $SU(N)$ is the group of $N$x$N$ unitary matrices with determinant 1. Because the group factors into two pieces which commute with each other, we can actually study the factors separately. We have already established that $U(1)$ simply reproduces electromagnetism. The theory produced by $SU(N)$ is called Yang-Mills. For SU(N) then there are $N^2 - 1$ independent components of the connection and equally many components of the curvature tensor. We can make (4.4.339) more familiar by defining a new set of generators $\tau_a$ by:

$$\tau_a = iT_a \implies T_a = -i\tau_a. \tag{4.4.341}$$

Then, (4.4.339) becomes:

$$\tau_a^\dagger = \tau_a. \tag{4.4.342}$$

Hence, the $\tau_a$ generators are Hermitian. In these generators, the Lie algebra is:

$$[\tau_a, \tau_b] = [iT_a, iT_b] = -f_{ab}{}^c T_c = i f_{ab}{}^c \tau_c. \tag{4.4.343}$$

In terms of these generators, the covariant derivative becomes:

$$\mathscr{D}_\mu \phi = \partial_\mu \phi - i\Gamma_\mu^a \tau_a \phi. \tag{4.4.344}$$

This motivates defining the gauge field components $A_\mu^a$ and coupling constant $g$ by:

$$\Gamma_\mu^a = g A_\mu^a \tag{4.4.345}$$

so that the covariant derivative is:

$$\mathscr{D}_\mu \phi = \partial_\mu \phi - ig A_\mu^a \tau_a \phi \tag{4.4.346}$$

which is precisely the same form as the electromagnetic covariant derivative. This definition only fixes the product $gA_\mu^a$ and not their individual scalings. Thus, $g$ can be set at will for now (but we will specify how to fix its value specifically later). Then, the gauge field is defined as:

$$A_\mu = A_\mu^a \tau_a \implies \Gamma_\mu = -igA_\mu. \tag{4.4.347}$$

Then the curvature tensor becomes:

$$\Omega_{\mu\nu} = \partial_\mu(-igA_\nu) - \partial_\nu(-igA_\mu) + [-igA_\mu, -igA_\nu] = -ig\left(\partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]\right). \tag{4.4.348}$$

This motivates defining the general field strength tensor:

$$\mathcal{F}_{\mu\nu} \stackrel{\text{def}}{=} \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu] \implies \Omega_{\mu\nu} = -ig\mathcal{F}_{\mu\nu}. \tag{4.4.349}$$

As well, from this we can compute the components of the field strength tensor:

$$\mathcal{F}_{\mu\nu} = \partial_\mu A_\nu^a \tau_a - \partial_\nu A_\mu^a \tau_a - igA_\mu^b A_\nu^c [\tau_b, \tau_c] \tag{4.4.350}$$

$$= \left(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f_{bc}{}^a A_\mu^b A_\nu^c\right) \tau_a \tag{4.4.351}$$

$$\implies \mathcal{F}_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f_{bc}{}^a A_\mu^b A_\nu^c. \tag{4.4.352}$$

We also rescale the group parameters $\theta$ as done in (4.4.330) by defining the new parameters $\alpha^a$ by:

$$\alpha^a = -\frac{1}{g}\theta^a. \tag{4.4.353}$$

Then, the gauge transformation operator $U$ becomes:

$$U(X) = e^{ig\alpha^a(X)\tau_a}. \tag{4.4.354}$$

As well, the gauge field has the transformation property:

$$A'_\mu = UA_\mu U^{-1} + \frac{i}{g} U \partial_\mu U^{-1} \tag{4.4.355}$$

and the components of the gauge field then have the transformation property:

$$A'^a_\mu = \left(e^{-g\alpha^c f_c}\right)^a{}_b A^b_\mu + \left(\frac{\mathbb{1} - e^{-g\alpha^c f_c}}{g\alpha^c f_c}\right)^a{}_b \partial_\mu \alpha^b. \tag{4.4.356}$$

The field strength tensor then has gauge transformation:

$$\mathcal{F}'_{\mu\nu} = U\mathcal{F}_{\mu\nu}U^{-1}. \tag{4.4.357}$$

or in components as:

$$\mathcal{F}'^a_{\mu\nu} = \left(e^{-g\alpha^c f_c}\right)^a{}_b \mathcal{F}^b_{\mu\nu}. \tag{4.4.358}$$

The covariant derivative of the field strength is given by:

$$\mathscr{D}_\rho \mathcal{F}_{\mu\nu} = \partial_\rho \mathcal{F}_{\mu\nu} - ig[A_\rho, \mathcal{F}_{\mu\nu}] \tag{4.4.359}$$

or in components as:

$$\mathscr{D}_\rho \mathcal{F}^a_{\mu\nu} = \partial_\rho \mathcal{F}^a_{\mu\nu} + g f_{bc}{}^a A^b_\rho \mathcal{F}^c_{\mu\nu}. \tag{4.4.360}$$

The covariant derivative of the field strength satisfies:

$$\epsilon^{\mu\nu\rho\sigma} \mathscr{D}_\nu \mathcal{F}_{\rho\sigma} = 0. \tag{4.4.361}$$

As well, the Wilson loop becomes:

$$W[\gamma] = \operatorname{Tr} \mathcal{P} : e^{ig \oint_\gamma A_\mu dx^\mu}. \tag{4.4.362}$$

Distinct from the Abelian case, the coupling constant $g$ appears in the field strength tensor whereas $q$ did not. So, while each $\phi$ can have its own $q$, any $\phi$ which transform in the same way under the gauge transformation $U$ must have the same $g$. Consequently, $g$ is a fundamental constant of the Yang-Mills theory. We can see by dimensional analysis that if we take $A^a_\mu$ to have the same units as electromagnetic $A_\mu$ (which we do, by definition), $g$ is in fact unitless and thus just a number.

Now we will identify the Lagrangian density for Yang-Mills theory. As established, it is given by (4.4.300) but with coefficients $A_{ab}$ satisfying (4.4.281). In order to find $A_{ab}$ satisfying such we first need the structure constants of SU($N$). We can compute them by giving an explicit expression to the generators $\tau_a$. Consider a $\tau_a$ with a single real non-zero entry in its upper triangle and zeros along the diagonal. Then, because it is self-adjoint, it must have a single equal corresponding entry on its lower triangle. There are $\frac{N(N-1)}{2}$ choices for where to put the non-zero entry on the upper triangle and so there are equally as many linearly independent $\tau_a$ of this form. We will call this class of $\tau_a$ "x" type because they resemble the Pauli matrix $\sigma_x$. We can identify an x-type generator uniquely by the row and column number of its upper triangle non-zero entry. So, let the matrix $X_{ab}$ have components $(X_{ab})^c{}_d$ so that it is non-zero only in the positions row $a$ column $b$ and row $b$ column $a$:

$$(X_{ab})^c{}_d = \frac{1}{2}\left(\delta^c_a \delta_{bd} + \delta^c_b \delta_{ad}\right). \tag{4.4.363}$$

This factor of $\frac{1}{2}$ is conventional. With this, we can take $\tau_1$ through $\tau_{\frac{N(N-1)}{2}}$ to be given in order by $X_{12}$, ... , $X_{1N}$, $X_{23}$, ... , $X_{(N-1)N}$. Now consider a $\tau_a$ with a single imaginary non-zero entry in its upper triangle and zeros along the diagonal. Then, because it is self-adjoint, it must have a single corresponding entry of equal magnitude and opposite sign in its lower triangle. There are again $\frac{N(N-1)}{2}$ choices for where to put the non-zero entry on the upper triangle and so there are equally as many linearly independent $\tau_a$ of this form. We will call this class of $\tau_a$ "y" type because they resemble the Pauli matrix $\sigma_y$. These are all linearly independent of each other and the x type generators as those were symmetric while these are all anti-symmetry. Let the matrix $Y_{ab}$ have components identified in the corresponding way to the $X_{ab}$:

$$(Y_{ab})^c{}_d = \frac{1}{2}\left(-i\delta^c_a \delta_{bd} + i\delta^c_b \delta_{ad}\right). \tag{4.4.364}$$

With this, we can take $\tau_{\frac{N(N-1)}{2}+1}$ through $\tau_{N(N-1)}$ to be given by the $Y_{ab}$ in the same enumeration scheme as was used for the $X_{ab}$. Finally, consider a $\tau_a$ with a single positive entry on its diagonal and an equal negative entry in the last entry of the diagonal so as to remain traceless. There are $N-1$ choices for where to put the positive entry and so there are as many linearly independent $\tau_a$ of this form. We will call

this class of $\tau_a$ "z" type because they resemble the Pauli matrix $\sigma_z$. These are all linearly independent of each other and the x and y type generators as those had only zero along the diagonal. Let the matrix $Z_a$ have a positive entry at diagonal element $a$ and negative entry in the last diagonal element:

$$(Z_a)^c{}_d = \frac{1}{2}\left(\delta^c_{\underline{a}}\delta_{\underline{a}d} - \delta^c_N\delta_{Nd}\right). \tag{4.4.365}$$

With this, we can take $\tau_{N(N-1)+1}$ through $\tau_{N^2-1}$ to be given by $Z_1$ , ... , $Z_{N-1}$. Then, because we have covered $N^2-1$ linearly independent generators, the x type, y type, and z type generators together form a complete basis of generators for SU(N). Going forward, we will commit an abuse of notation by identifying each generator $\tau_a$ with the corresponding x, y, or z type generator as enumerated. For example, (4.4.343) now splits into 6 cases (due to anti-symmetry of the commutator), depending on which classes the $\tau_a$ and $\tau_b$ belong to:

$$[X_{ab}, X_{cd}] = if_{X_{ab}X_{cd}}{}^{X_{ef}}X_{ef} + if_{X_{ab}X_{cd}}{}^{Y_{ef}}Y_{ef} + if_{X_{ab}X_{cd}}{}^{Z_e}Z_e \tag{4.4.366}$$

$$[X_{ab}, Y_{cd}] = if_{X_{ab}Y_{cd}}{}^{X_{ef}}X_{ef} + if_{X_{ab}Y_{cd}}{}^{Y_{ef}}Y_{ef} + if_{X_{ab}Y_{cd}}{}^{Z_e}Z_e \tag{4.4.367}$$

$$[X_{ab}, Z_c] = if_{X_{ab}Z_c}{}^{X_{de}}X_{de} + if_{X_{ab}Z_c}{}^{Y_{de}}Y_{de} + if_{X_{ab}Z_c}{}^{Z_d}Z_d \tag{4.4.368}$$

$$[Y_{ab}, Y_{cd}] = if_{Y_{ab}Y_{cd}}{}^{X_{ef}}X_{ef} + if_{Y_{ab}Y_{cd}}{}^{Y_{ef}}Y_{ef} + if_{Y_{ab}Y_{cd}}{}^{Z_e}Z_e \tag{4.4.369}$$

$$[Y_{ab}, Z_c] = if_{Y_{ab}Z_c}{}^{X_{de}}X_{de} + if_{Y_{ab}Z_c}{}^{Y_{de}}Y_{de} + if_{Y_{ab}Z_c}{}^{Z_d}Z_d \tag{4.4.370}$$

$$[Z_a, Z_b] = if_{Z_aZ_b}{}^{X_{cd}}X_{cd} + if_{Z_aZ_b}{}^{Y_{cd}}Y_{cd} + if_{Z_aZ_b}{}^{Z_c}Z_c. \tag{4.4.371}$$

These fully determine the structure coefficients of SU(N), though they are quite nasty in general. When $N=2$, this is literally our usual expression for the generators of SU(2). As well, we can see that our expressions for the generators satisfy:

$$\mathrm{Tr}(\tau_a\tau_b) = \frac{1}{2}\delta_{ab}. \tag{4.4.372}$$

With this we can define $2\mathrm{Tr}(\tau_a\tau_b)$ as defining a "metric" on the generators so that we can raise and lower indices on the internal space using it. Because of our particular basis of generators, the raised and lowered index versions are the same. With this, from (4.4.343) we have:

$$\mathrm{Tr}([\tau_a,\tau_b]\tau_c) = if_{ab}{}^d\mathrm{Tr}(\tau_d\tau_c) = \frac{i}{2}f_{ab}{}^d\delta_{cd} = \frac{i}{2}f_{abc} \tag{4.4.373}$$

$$\implies f_{abc} = -2i\mathrm{Tr}(\tau_a\tau_b\tau_c - \tau_b\tau_a\tau_c). \tag{4.4.374}$$

Using the cyclicity of the trace then, we can see that $f_{abc}$ is anti-symmetric in the interchange of any pair of its indices. Thus, if we define as usual $f_a$ as the matrix with components $f_a{}^b{}_c = f_{ac}{}^b$ then it these matrices are all real and anti-symmetric. Then raising and lowering indices in (4.4.281) and using that $A_{ab} = A_{ba}$ we have:

$$A^b{}_c f_d{}^c{}_a - f_d{}^b{}_c A^c{}_a = 0 \tag{4.4.375}$$

$$\implies [A, f_d] = 0 \tag{4.4.376}$$

where $A$ is the matrix with components $A_{ab}$. Thus, the matrix $A$ must commute with all of the generators $f_a$ in the adjoint representation. These $N^2-1$ commutators are quite a bit of effort to evaluate but after explicitly computing the $f_a$ from our x, y, and z type generators and then requiring these commutators, one can see that $A$ must in fact be proportional to the identity matrix. Thus, for some constant $k$:

$$A = k\mathbb{1} \implies A_{ab} = k\delta_{ab} = 2k\mathrm{Tr}(\tau_a\tau_b). \tag{4.4.377}$$

Consequently, our Lagrangian density becomes:

$$\mathfrak{L} = g^2 k\delta_{ab}\eta^{\mu\rho}\eta^{\nu\sigma}\mathcal{F}^a_{\mu\nu}\mathcal{F}^b_{\rho\sigma} = kg^2\mathcal{F}^{\mu\nu}_a\mathcal{F}^a_{\mu\nu}. \tag{4.4.378}$$

By convention, we again choose $g$ (which so far has been unspecified) so that:

$$kg^2 \stackrel{!}{=} -\frac{1}{4}. \tag{4.4.379}$$

Notice that this is only possible if $k$ is negative. A positive value of $k$ actually causes the weak field (meaning small field strength) limit to be unstable and so for our weak field expansion that we have done to produce this Lagrangian to be valid we need $k$ to be negative by stability. (This is explored in the homework.)

Then, the Lagrangian density for our connection over SU(N) is together with whatever other fields exist is:

$$\mathfrak{L} = -\frac{1}{4}\mathcal{F}_a^{\mu\nu}\mathcal{F}^a_{\mu\nu} + \mathfrak{L}_0 \tag{4.4.380}$$

where $\mathfrak{L}_0$ is the Lagrangian density due to whatever other fields are present. This specific expression depends on our basis of generators satisfying (4.4.372). However, we can write this in a basis independent way as:

$$\mathfrak{L} = -\frac{1}{4}2\mathrm{Tr}(\tau_a\tau_b)\mathcal{F}^{a\mu\nu}\mathcal{F}^b_{\mu\nu} + \mathfrak{L}_0 = -\frac{1}{2}\mathrm{Tr}(\mathcal{F}^{a\mu\nu}\tau_a\mathcal{F}^b_{\mu\nu}\tau_b) + \mathfrak{L}_0 \tag{4.4.381}$$

$$\implies \mathfrak{L} = -\frac{1}{2}\mathrm{Tr}(\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu}) + \mathfrak{L}_0. \tag{4.4.382}$$

The normalization of the Lagrangian density to be in this form fixed the value of the coupling constant $g$ (up to a sign, which is conventionally chosen to be positive.). So, Yang-Mills has a fundamental parameter $g$ which must be determined by experiment. Now that we have our Lagrangian density, let us compute the equation of motion for the gauge field. Because $\mathfrak{L}_0$ will depend on our other fields and those fields may have their covariant derivatives taken $\mathscr{D}_\mu$ may appear in $\mathfrak{L}_0$ and because the covariant derivative depends on $A_\mu$, $\mathfrak{L}_0$ is allowed to depend on $A_\mu$. Then:

$$\frac{\partial\mathfrak{L}}{\partial A^a_\mu} = -\frac{1}{2}\mathcal{F}_b^{\rho\sigma}\frac{\partial}{\partial A^a_\mu}\left(\partial_\rho A^b_\sigma - \partial_\sigma A^b_\rho + g f_{cd}{}^b A^c_\rho A^d_\sigma\right) + \frac{\partial\mathfrak{L}_0}{\partial A^a_\mu} \tag{4.4.383}$$

$$= -\frac{1}{2}\mathcal{F}_b^{\mu\sigma}2g f_{ac}{}^b A^c_\sigma + \frac{\partial\mathfrak{L}_0}{\partial A^a_\mu} \tag{4.4.384}$$

$$= \frac{\partial\mathfrak{L}_0}{\partial A^a_\mu} - g f_{abc}A^b_\nu\mathcal{F}^{c\mu\nu} \tag{4.4.385}$$

$$\frac{\partial\mathfrak{L}}{\partial\partial_\nu A^a_\mu} = -\frac{1}{2}\mathcal{F}_b^{\rho\sigma}\frac{\partial}{\partial\partial_\nu A^a_\mu}\left(\partial_\rho A^b_\sigma - \partial_\sigma A^b_\rho + g f_{cd}{}^b A^c_\rho A^d_\sigma\right) \tag{4.4.386}$$

$$= -\frac{1}{2}\mathcal{F}_a^{\nu\mu}2 = \mathcal{F}_a^{\mu\nu}. \tag{4.4.387}$$

So, the equation of motion becomes:

$$\partial_\nu\frac{\partial\mathfrak{L}}{\partial\partial_\nu A^a_\mu} = \frac{\partial\mathfrak{L}}{\partial A^a_\mu} \tag{4.4.388}$$

$$\implies \partial_\nu\mathcal{F}_a^{\mu\nu} = -g f_{abc}A^b_\nu\mathcal{F}^{c\mu\nu} + \frac{\partial\mathfrak{L}_0}{\partial A^a_\mu} \tag{4.4.389}$$

$$\implies \mathscr{D}_\nu\mathcal{F}_a^{\mu\nu} = \frac{\partial\mathfrak{L}_0}{\partial A^a_\mu}. \tag{4.4.390}$$

This is precisely the equivalent of Maxwell's equations because the covariant derivative of the electromagnetic field strength is just the partial derivative due to the vanishing commutators in that theory. Thus, we can identify that the Yang-Mills current density is:

$$J^\mu_a = \frac{\partial\mathfrak{L}_0}{\partial A^a_\mu}. \tag{4.4.391}$$

We generally call the property which electromagnetism interacts with electric charge. Similarly, we generally call the property which Yang-Mills interacts with color (even though it has nothing to do with the everyday life property known as color). So, this $J^\mu_a$ is called the color current density. With this the Yang-Mills equation of motion is:

$$\mathscr{D}_\nu\mathcal{F}_a^{\mu\nu} = J^\mu_a. \tag{4.4.392}$$

We know the gauge transformation property of the left hand side and so it determines the necessary gauge transformation property of the color current. Then, because the left hand side is in the adjoint representation, the color current is as well. Then, because the form of the covariant derivative of a quantity is determined by the representation, the expression for the covariant derivative of $J^\mu_a$ must be given analogously to (4.4.360) as:

$$\mathscr{D}_\nu J^\mu_a = \partial_\nu J^\mu_a + g f_{ab}{}^c A^b_\nu J^\mu_c. \tag{4.4.393}$$

Just like Maxwell's equations require the continuity equation for the electric current, the Yang-Mills equations require a covariant continuity equation for $J^\mu_a$. Computing the covariant divergence of the color

current:

$$\mathscr{D}_\mu J_a^\mu = \mathscr{D}_\mu \mathscr{D}_\nu \mathcal{F}_a^{\mu\nu} \tag{4.4.394}$$

$$= \frac{1}{2}(\mathscr{D}_\mu \mathscr{D}_\nu - \mathscr{D}_\nu \mathscr{D}_\mu)\mathcal{F}_a^{\mu\nu} + \frac{1}{2}(\mathscr{D}_\mu \mathscr{D}_\nu + \mathscr{D}_\nu \mathscr{D}_\mu)\mathcal{F}_a^{\mu\nu}. \tag{4.4.395}$$

The second term is the contraction of a tensor symmetric in $\mu$, $\nu$ with a tensor antisymmetric in $\mu$, $\nu$. Consequently, it can be sent to its own negative by interchanging the names of $\mu$ and $\nu$ and thus must be 0. The first term is the commutator of covariant derivatives on the field strength:

$$\mathscr{D}_\mu J_a^\mu = \frac{1}{2}[\mathscr{D}_\mu, \mathscr{D}_\nu]\mathcal{F}_a^{\mu\nu}. \tag{4.4.396}$$

We can find a nice formula for the commutator of covariant derivatives on any quantity $\varphi^a$ in the adjoint representation:

$$[\mathscr{D}_\mu, \mathscr{D}_\nu]\varphi^a = \partial_\mu \mathscr{D}_\nu \varphi^a + g f_{bc}{}^a A_\mu^b \mathscr{D}_\nu \varphi^a - \partial_\nu \mathscr{D}_\mu \varphi^a - g f_{bc}{}^a A_\nu^b \mathscr{D}_\mu \varphi^a \tag{4.4.397}$$

$$= \partial_\mu(\partial_\nu \varphi^a + g f_{bc}{}^a A_\nu^b \varphi^c) + g f_{bc}{}^a A_\mu^b (\partial_\nu \varphi^c + g f_{de}{}^c A_\nu^d \varphi^e) - \partial_\nu(\partial_\mu \varphi^a + g f_{bc}{}^a A_\mu^b \varphi^c)$$
$$- g f_{bc}{}^a A_\nu^b (\partial_\mu \varphi^c + g f_{de}{}^c A_\mu^d \varphi^e) \tag{4.4.398}$$

$$= g f_{bc}{}^a \partial_\mu A_\nu^b \varphi^c + g^2 f_{bc}{}^a f_{de}{}^c A_\mu^b A_\nu^d \varphi^e - g f_{bc}{}^a \partial_\nu A_\mu^b \varphi^c - g^2 f_{bc}{}^a f_{de}{}^c A_\nu^b A_\mu^d \varphi^e \tag{4.4.399}$$

$$= g f_{bc}{}^a (\partial_\mu A_\nu^b - \partial_\nu A_\mu^b)\varphi^c + g^2 (f_{bc}{}^a f_{de}{}^c + f_{dc}{}^a f_{eb}{}^c) A_\mu^b A_\nu^d \varphi^e. \tag{4.4.400}$$

Now using the Jacobi identity:

$$f_{bc}{}^a f_{de}{}^c + f_{dc}{}^a f_{eb}{}^c + f_{ec}{}^a f_{bd}{}^c = 0 \tag{4.4.401}$$

we have:

$$[\mathscr{D}_\mu, \mathscr{D}_\nu]\varphi^a = g f_{bc}{}^a (\partial_\mu A_\nu^b - \partial_\nu A_\mu^b)\varphi^c + g^2 f_{ce}{}^a f_{bd}{}^c A_\mu^b A_\nu^d \varphi^e \tag{4.4.402}$$

$$= g f_{bc}{}^a \mathcal{F}_{\mu\nu}^b \varphi^c. \tag{4.4.403}$$

Thus:

$$\mathscr{D}_\mu J_a^\mu = \frac{1}{2} g f_{bca} \mathcal{F}_{\mu\nu}^b \mathcal{F}^{c\mu\nu} = -\frac{1}{2} g f_{cba} \mathcal{F}_{\mu\nu}^b \mathcal{F}^{c\mu\nu} = -\frac{1}{2} g f_{bca} \mathcal{F}_{\mu\nu}^c \mathcal{F}^{b\mu\nu} = -\mathscr{D}_\mu J_a^\mu \tag{4.4.404}$$

$$\implies \mathscr{D}_\mu J_a^\mu = 0. \tag{4.4.405}$$

Thus, we have the covariant conservation of the color current in Yang-Mills.

### 4.4.5 Quantum Chromodynamics

Experimentally, it has been observed that nucleons (protons and neutrons) which comprise the nuclei of atoms are themselves comprised of "quarks" which are electrically charged Dirac spinors which in addition to electromagnetism interact via a Yang-Mills interaction with gauge group SU(3). This interaction is called the "nuclear strong force" or just "strong force" because it was discovered in the nucleus and in the nucleus it is quite strong. As well, it is called "chromodynamics" in the same way electromagnetism is sometimes called electrodynamics. The coupling constant for this force is called the "strong coupling constant "$g_S$". For the purposes of this subsection we will neglect the electromagnetic interaction. (Practically, when chromodynamics is relevant it is typically much more important than electromagnetism. To include electromagnetism later, we will simply have to add the electromagnetic contribution to the covariant derivative and the terms corresponding to the electromagnetic field strength to the Lagrangian density.) As SU(3) "colored" Dirac spinors, a quark is described by a Dirac spinor $\psi$ but instead of each component of $\psi$ just being a complex number it is 3 complex numbers which transform into each other under SU(3) transformations according to (4.4.2). Thus, $\psi$ is the $\phi$ appearing in (4.4.12). Our gauge field $G_\mu^a$ for this theory then has $a$ ranging from 1 to $N^2 - 1 = 8$. These 8 intermingling gauge fields are called the gluon potentials or gluon fields. Then, together the Lagrangian density for the interaction of a quark with the gluon is:

$$\mathfrak{L} = i\bar{\psi}\gamma^\mu \mathscr{D}_\mu \psi - m\bar{\psi}\psi - \frac{1}{2}\mathrm{Tr}\,(\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu}) \tag{4.4.406}$$

where the covariant derivative of the quark is given by (4.4.346) as:

$$\mathscr{D}_\mu \psi = \partial_\mu \psi - i g_S G_\mu^a \tau_a \psi \tag{4.4.407}$$

and $m$ is the mass of the quark. For $N = 3$, we have that there are $\frac{N(N-1)}{2} = 3$ x type generators, an equal number of y type generators, and $N - 1 = 2$ z type generators. In the $X, Y, Z$ basis, we have that the SU(3) generators can explicitly be taken to be:

$$X_{12} = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad X_{13} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad X_{23} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{4.4.408}$$

$$Y_{12} = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad Y_{13} = \frac{1}{2}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \qquad Y_{23} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \tag{4.4.409}$$

$$Z_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad Z_2 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{4.4.410}$$

### 4.4.6 The Weak Interaction

The Yang-Mills interaction appears in another way though for now this will appear to cause some quite significant challenges. Experimentally, the quarks interact via an SU(2) Yang-Mills interaction in addition to electromagnetism and the strong force. As well, there are particles called "leptons" which are also Dirac spinors and interact via electromagnetism and this same SU(2) interaction but not the strong force. This SU(2) Yang-Mills interaction is called the "nuclear weak force" or just "weak force" because it was discovered when studying atomic nuclei and it typically has much weaker effects than the strong force. The coupling constant for this force is called the weak coupling constant "$g_W$". Because we use color to name the interaction property of the strong interaction, we need a different name for the weak interaction. The "color charge" of the weak interaction is called "isospin" (some call it "weak isospin"). For $N = 2$, we have that there is $\frac{N(N-1)}{2} = 1$ $x$ type generator, an equal number of $y$ type generators, and $N - 1 = 1$ $z$ type generator. In fact, the the $X, Y, Z$ basis we have that these are simply half of the Pauli matrices:

$$\tau_1 = X_{12} = \frac{1}{2}\sigma_1, \qquad \tau_2 = Y_{12} = \frac{1}{2}\sigma_2, \qquad \tau_3 = Z_1 = \frac{1}{2}\sigma_3. \tag{4.4.411}$$

Consequently, we immediately know the structure constants for the weak force as we already know the commutation relations for the Pauli matrices:

$$[\sigma_a, \sigma_b] = 2i\epsilon_{ab}{}^c\sigma_c \implies [\tau_a, \tau_b] = i\epsilon_{ab}{}^c\tau_c \implies f_{ab}{}^c = \epsilon_{ab}{}^c. \tag{4.4.412}$$

The weak interaction works quite a bit differently from the strong interaction, surprisingly. In particular, the weak gauge fields don't appear to satisfy the equation of motion (4.4.392) at low energies. Instead they seem to have an additional term which in the classical limit corresponds to making the gauge fields massive. In pure Yang-Mills, the gauge fields are guaranteed to be massless. The quarks and leptons have some isospin up fields and each has a corresponding isospin down field to form a representation of SU(2). The electric charges of the isospin up fields do not match their isospin down counterparts. Further, only the left handed parts of the spinors actually couple to the weak interaction. That is, the left handed fields have isospin whereas the right handed fields do not and do not transform at all under the Weak SU(2) transformation. This is all very bizarre as it breaks the parity symmetry we have always had until this point and it doesn't seem to be consistent with all of the Yang-Mills theory as described. We will reconcile how all of this is possible in the next section.

## Exercises

**Problem 4.1: Wong's equations for a classical colored particle Lagrangian/Action/Hamiltonian.**
(a)

**Problem 4.2: CPT transformation of vector fields.**
(a)

**Problem 4.3: creation annihilation operators.**
(a)

**Problem 4.4: hamiltonian for this theory.**
(a)

**Problem 4.5: solve equations of motion/diagonalize hamiltonian.**
(a)

**Problem 4.6: energy momentum tensor for vector.**
(a)
**Problem 4.7: solve equations of motion/diagonalize hamiltonian.**
(a)
**Problem 4.8: homework question about trying spin 1 without using gauge theory show that it becomes gauge theory by requiring boundedness below of the Hamiltonian. Discuss why boundedness below of the Hamiltonian is a reasonable property to desire.**
(a)
**Problem 4.9: homework question establishing that 1/2 1/2 is the correct representation for gauge theory.**
(a)
**Problem 4.10: Lorentz invariant tensors of ranks up through 4.**
(a)
**Problem 4.11: recovering gauge field from field strength for electromagnetism, holonomy from action recovery argument.**
(a)
**Problem 4.12: gauge field which cannot be recovered from its curvature tensor.**
(a)
**Problem 4.13: derive yukawa potential.**
(a)
**Problem 4.14: Lorenz gauge term addition to the Lagrangian homework question (gauge fixing terms in the Lagrangian).**
(a)
**Problem 4.15: gauge fixing terms in the Yang mills Lagrangian.**
(a)
**Problem 4.16: large gauge transformations.**
(a)
**Problem 4.17: chern simons term, index theorem of the integral of it, gauge invariance for small gauge transformations but not large ones.**
(a)
**Problem 4.18: homework question about wrong sign Lagrangian for yang mills and instability**
(a)
**Problem 4.19: homework question explicitly compute the structure coefficients and commutators with $A$ to show that it has to be proportional to the identity.**
(a)
**Problem 4.20: question about why the color current is covariantly conserved rather than continuity conserved and what that means from Noether's theorem by computing the Noether generator corresponding to gauge transformations.**
(a)
**Problem 4.21: generator of u(1) transformations for electromagnetism, conservation law.**
(a)
**Problem 4.22: TODO why 3 colors for the strong force?**
(a)
**Problem 4.23: effective field strength tensor (meaning one which appears in particle equations of motion) for electromagnetism from Higgs ssb, show it can have monopoles (t'Hooft Polyakov)**
(a)

## 4.5 Dynamics of Nonabelian Gauge Theory

### 4.5.1 Path Integral and LSZ Reduction

### 4.5.2 Feynman Rules and Renormalization

### 4.5.3 Ward Identities

### 4.5.4 Becchi-Rouet-Stora-Tyutin Formalism

### 4.5.5 Wilson Loops and Confinement

## 4.6 Anomalies

### 4.6.1 Adler-Bell-Jackiw Anomaly

### 4.6.2 Chiral Symmetry Breaking

### 4.6.3 Gauge Anomalies

## 4.7 Summary and Additional Reading

# Chapter 5

# The Standard Model and Beyond

*The creation of Physics is the shared heritage of all mankind. East and West, North and South have equally participated in it.*
– Abdus Salam [53]

*Why is the muon, some dumb particle, 200 times heavier than the electron? Why is the proton about 2,000 times heavier than the electron? Why is the electric charge of the electron what it is? Why are there six quarks in nature? Why not seven or eleven or five? There are many, many "why" questions. Also a number of "how" questions. What is the mechanism that causes the weak interactions to be weak and the electromagnetic interactions not weak? All kinds of questions remain.*
– Sheldon Glashow [22]

*The years since the mid-1970s have been the most frustrating in the history of particle physics. We are paying the price of our own success: theory has advanced so far that further progress will require the study of processes at energies far beyond the reach of existing facilities.*
– Steven Weinberg [61]

## 5.1 Electroweak Symmetry Breaking and the Higgs Mechanism

The underlying reason behind why the weak interaction is so different is called the Higgs mechanism. Consider a complex scalar field $\Phi$ (called the Higgs field) which is charged under a U(1) interaction with gauge field $B_\mu$ and under an SU(2) interaction with gauge fields $W^a_\mu$. That is, $\Phi$ has two complex components which transform under U(1) and SU(2) transformations. We can take the generators of SU(2) transformations $\tau_a$ to just be $\sigma_a/2$ where $\sigma_a$ are the Pauli matrices. We will call the U(1) interaction here the "hypercharge" interaction (with coupling constant $g_Y$) and the SU(2) interaction the weak interaction (with coupling constant $g_W$). Then, the covariant derivative of $\Phi$ is given as:

$$\mathscr{D}_\mu\Phi = \partial_\mu\Phi - ig_Y B_\mu\Phi - ig_W\tau_a W^a_\mu\Phi. \tag{5.1.1}$$

We know that to write a Lagrangian density for such a field we are welcome to include the term $\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi$ as it has all of the necessary symmetries. As well, we can add any function of $\Phi^\dagger\Phi$ as it also has the necessary symmetries. Previously, we only considered adding a term linear in $\Phi^\dagger\Phi$ and we fixed its coefficient to be $-m^2$ where $m$ corresponds to the mass of $\Phi$ in the classical limit. This was motivated by assuming $\Phi = 0$ is a stable equilibrium and performing a lowest order Taylor expansion and matching our previous results about the point limit of a quadratic theory of waves. However, it need not necessarily be the case that $\Phi = 0$ is a stable equilibrium, in which case we could have higher order terms in $\Phi$ and perhaps identify a new stable equilibrium. One cannot simply shift the definition of $\Phi$ by adding some constant SU(2) vector to it so that the new field has equilibrium at 0 and has a Lagrangian density of the same form as such

a shift does not preserve the U(1) or SU(2) transformations. So, as the next possible term, let us consider adding a term $\Phi^\dagger\Phi$. That is, up to quartic order in the fields, the Lagrangian density for $\Phi$ is:

$$\mathfrak{L} = -\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi - a\Phi^\dagger\Phi - b(\Phi^\dagger\Phi)^2. \tag{5.1.2}$$

(The coefficient on the first term is fixed in magnitude by scaling $\Phi$ so that it has magnitude 1 by definition. The sign is fixed by the point limit of $\Phi$ and the previous matching, just as before. That logic is unchanged even when the equilibrium value of $\Phi$ is non-zero as it drops out in the differentiation.) We don't consider higher derivative terms because of renormalizability. For $a > 0$, this theory is again stable near $\Phi = 0$ and so essentially nothing we haven't already seen before happens. However, if $a < 0$ this theory is actually unstable near $\Phi = 0$. To see this, look at the equation of motion for $\Phi$:

$$\frac{\partial\mathfrak{L}}{\partial\Phi^\dagger} - \partial_\mu\frac{\partial\mathfrak{L}}{\partial\partial_\mu\Phi^\dagger} = \mathscr{D}^2\Phi - a\Phi - 2b(\Phi^\dagger\Phi)\Phi = 0. \tag{5.1.3}$$

If we consider $\Phi$ near 0, the last term is suppressed. As well, if we consider configurations which are absent any gauge fields and which for simplicity have $\Phi$ as spatially uniform then this becomes:

$$\frac{\partial^2\Phi}{\partial t^2} + a\Phi \approx 0. \tag{5.1.4}$$

For $a > 0$ this is just a harmonic oscillator and so is stable. For $a < 0$ this grows and decays exponentially and so is unstable. So, consider $a < 0$ to see what interesting possibilities arise and let $\mu^2 = -a$. We do want our theory to be stable overall; $a < 0$ simply makes the point $\Phi = 0$ unstable. If we now do not demand $\Phi$ to be small but still look at spatially uniform configurations in the absence of gauge fields we have:

$$\frac{\partial^2\Phi}{\partial t^2} - \mu^2\Phi + 2b(\Phi^\dagger\Phi)\Phi = 0. \tag{5.1.5}$$

In order for this theory to have an equilibrium we need there to be a constant $\Phi$ satisfying this equation of motion. Setting the time derivatives to 0 to find such a constant we see:

$$\Phi(-\mu^2 + 2b(\Phi^\dagger\Phi)) = 0. \tag{5.1.6}$$

So, $\Phi = 0$ is still an equilibrium but as discussed here it is an unstable equilibrium. The only other possibility for equilibrium then is a $\Phi$ which satisfies:

$$\Phi^\dagger\Phi = \frac{\mu^2}{2b}. \tag{5.1.7}$$

The left hand side is automatically positive and so the existence of a stable equilibrium in this theory requires the right hand side, and consequently $b$, to be positive as well. The more common name for $b$ is $\lambda$ and so this theory with $\mu > 0, \lambda > 0$ has Lagrangian density:

$$\mathfrak{L} = -\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi + \mu^2\Phi^\dagger\Phi - \lambda(\Phi^\dagger\Phi)^2. \tag{5.1.8}$$

If we put this together now with the pieces of the Lagrangian density corresponding to the gauge fields we have:

$$\mathfrak{L} = -\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi + \mu^2\Phi^\dagger\Phi - \lambda(\Phi^\dagger\Phi)^2 - \frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}) \tag{5.1.9}$$

where $\mathcal{F}_{Y\mu\nu}$ is the hyper charge gauge field defined by:

$$\mathcal{F}_{Y\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{5.1.10}$$

and $\mathcal{F}_{W\mu\nu}$ is the weak gauge field defined by:

$$\mathcal{F}_{W\mu\nu} = (\partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g_W\epsilon_{bc}{}^a W_\mu^b W_\nu^c)\tau_a. \tag{5.1.11}$$

From this Lagrangian density we have the exact equations of motion:

$$\frac{\partial\mathfrak{L}}{\partial\Phi^\dagger} - \partial_\mu\frac{\partial\mathfrak{L}}{\partial\partial_\mu\Phi^\dagger} = \mathscr{D}^2\Phi + \mu^2\Phi - 2\lambda(\Phi^\dagger\Phi)\Phi \stackrel{!}{=} 0 \tag{5.1.12}$$

$$\frac{\partial\mathfrak{L}}{\partial B_\mu} - \partial_\nu\frac{\partial\mathfrak{L}}{\partial\partial_\nu B_\mu} = 2g_Y\mathrm{Im}(\Phi^\dagger\mathscr{D}^\mu\Phi) - \partial_\nu\mathcal{F}_Y^{\mu\nu} \stackrel{!}{=} 0 \tag{5.1.13}$$

$$\frac{\partial\mathfrak{L}}{\partial W_\mu^a} - \partial_\nu\frac{\partial\mathfrak{L}}{\partial\partial_\nu W_\mu^a} = 2g_W\mathrm{Im}(\Phi^\dagger\tau_a\mathscr{D}^\mu\Phi) - \mathscr{D}_\nu\mathcal{F}_{Wa}^{\mu\nu} \stackrel{!}{=} 0. \tag{5.1.14}$$

Immediately then we can see that this theory does have an equilibrium whenever the gauge fields vanish and

$\Phi$ is a uniform constant $\Phi_0$ satisfying:

$$\Phi_0^\dagger \Phi_0 = \frac{\mu^2}{2\lambda}. \tag{5.1.15}$$

Now let us look at this theory near its equilibrium by linearizing the equations of motion:

$$\Phi(X) = \Phi_0 + \delta\Phi(X), \qquad B_\mu(X) = \delta B_\mu(X), \qquad W_\mu^a(X) = \delta W_\mu^a(X) \tag{5.1.16}$$

where we will keep only linear terms in the $\delta$-fields in approximating the equations of motion as these are our perturbations away from equilibrium. Linearizing then:

$$\begin{aligned}
&\mathscr{D}^2\Phi + \mu^2\Phi - 2\lambda(\Phi^\dagger\Phi)\Phi \\
&\quad = (\partial^\mu - ig_Y\delta B^\mu - ig_W\tau_b\delta W^{b\mu})(\partial_\mu - ig_Y\delta B_\mu - ig_W\tau_a\delta W_\mu^a)(\Phi_0 + \delta\Phi) \\
&\qquad + \mu^2(\Phi_0 + \delta\Phi) - 2\lambda(\Phi_0^\dagger + \delta\Phi^\dagger)(\Phi_0 + \delta\Phi)(\Phi_0 + \delta\Phi) && (5.1.17)\\
&\quad = \partial^2\delta\Phi - ig_Y(\partial^\mu\delta B_\mu)\Phi_0 - ig_W\tau_a(\partial^\mu\delta W_\mu^a)\Phi_0 - 2\lambda(\delta\Phi^\dagger\Phi_0 + \Phi_0^\dagger\delta\Phi)\Phi_0 + \text{H.O.T.} && (5.1.18)
\end{aligned}$$

$$\begin{aligned}
&2g_Y\mathrm{Im}(\Phi^\dagger\mathscr{D}^\mu\Phi) - \partial_\nu\mathcal{F}_Y^{\mu\nu} \\
&\quad = 2g_Y\mathrm{Im}((\Phi_0^\dagger + \delta\Phi^\dagger)(\partial^\mu - ig_Y\delta B^\mu - ig_W\tau_a\delta W_\mu^a)(\Phi_0 + \delta\Phi)) - \partial_\nu(\partial^\mu\delta B^\nu - \partial^\nu\delta B^\mu) && (5.1.19)\\
&\quad = \partial^2\delta B^\mu - \partial^\mu\partial^\nu\delta B_\nu + 2g_Y\mathrm{Im}(\Phi_0^\dagger\partial^\mu\delta\Phi) - \frac{g_Y^2\mu^2}{\lambda}\delta B^\mu - 2g_Y g_W\Phi_0^\dagger\tau_a\Phi_0\delta W^{a\mu} + \text{H.O.T.} && (5.1.20)
\end{aligned}$$

$$\begin{aligned}
&2g_W\mathrm{Im}(\Phi^\dagger\tau_a\mathscr{D}^\mu\Phi) - \mathscr{D}_\nu\mathcal{F}_{Wa}^{\mu\nu} \\
&\quad = 2g_W\mathrm{Im}((\Phi_0^\dagger + \delta\Phi^\dagger)\tau_a(\partial^\mu - g_Y\delta B^\mu - ig_W\tau_b\delta W^{b\mu})(\Phi_0 + \delta\Phi)) \\
&\qquad - (\delta_{ac}\partial_\nu + g_W\epsilon_{abc}\delta W_\nu^b)(\partial^\mu\delta W^{c\nu} - \partial^\nu\delta W^{c\mu} + g_W\epsilon_{de}{}^c\delta W^{d\mu}\delta W^{e\nu}) && (5.1.21)\\
&\quad = \partial^2\delta W_a^\mu - \partial^\mu\partial^\nu\delta W_{a\nu} + 2g_W\mathrm{Im}(\Phi_0^\dagger\tau_a\partial^\mu\delta\Phi) - 2g_W g_Y\Phi_0^\dagger\tau_a\Phi_0\delta B^\mu - \frac{g_W^2\mu^2}{4\lambda}\delta W_a^\mu + \text{H.O.T.} && (5.1.22)
\end{aligned}$$

So, we take as our linearized equations of motion:

$$\partial^2\delta\Phi - ig_Y(\partial^\mu\delta B_\mu)\Phi_0 - ig_W\tau_a(\partial^\mu\delta W_\mu^a)\Phi_0 - 2\lambda(\delta\Phi^\dagger\Phi_0 + \Phi_0^\dagger\delta\Phi)\Phi_0 = 0 \tag{5.1.23}$$

$$\partial^2\delta B^\mu - \partial^\mu\partial^\nu\delta B_\nu + 2g_Y\mathrm{Im}(\Phi_0^\dagger\partial^\mu\delta\Phi) - \frac{g_Y^2\mu^2}{\lambda}\delta B^\mu - 2g_Y g_W\Phi_0^\dagger\tau_a\Phi_0\delta W^{a\mu} = 0 \tag{5.1.24}$$

$$\partial^2\delta W_a^\mu - \partial^\mu\partial^\nu\delta W_{a\nu} + 2g_W\mathrm{Im}(\Phi_0^\dagger\tau_a\partial^\mu\delta\Phi) - 2g_W g_Y\Phi_0^\dagger\tau_a\Phi_0\delta B^\mu - \frac{g_W^2\mu^2}{4\lambda}\delta W_a^\mu = 0. \tag{5.1.25}$$

It is quite difficult to tell much about what is going on with these as they are written now. However, we have not yet used our ability to select a gauge. A generic SU(2) transformation is given as:

$$U(X) = e^{\frac{i}{2}g_W\alpha^a(X)\sigma_a} = \sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{ig_W}{2}\right)^n(\alpha^a\sigma_a)^n. \tag{5.1.26}$$

Define $|\alpha| = \sqrt{(\alpha^1)^2 + (\alpha^2)^2 + (\alpha^3)^2}$. Then, using:

$$(\alpha^a\sigma_a)^2 = \alpha^a\alpha^b\sigma_a\sigma_b = \alpha^a\alpha^b\delta_{ab}\mathbb{1} = |\alpha|^2\mathbb{1} \tag{5.1.27}$$

we have:

$$\begin{aligned}
U(X) &= \sum_{n=0}^{\infty}\frac{1}{(2n)!}\left(\frac{ig_W}{2}\right)^{2n}|\alpha|^{2n}\mathbb{1} + \sum_{n=0}^{\infty}\frac{1}{(2n+1)!}\left(\frac{ig_W}{2}\right)^{2n+1}|\alpha|^{2n}\alpha^a\sigma_a && (5.1.28)\\
&= \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}\left(\frac{g_W|\alpha|}{2}\right)^{2n}\mathbb{1} + i\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}\left(\frac{g_W|\alpha|}{2}\right)^{2n+1}\frac{\alpha^a\sigma_a}{|\alpha|} && (5.1.29)\\
&= \cos\left(\frac{g_W|\alpha|}{2}\right)\mathbb{1} + i\sin\left(\frac{g_W|\alpha|}{2}\right)\frac{\alpha^a\sigma_a}{|\alpha|} && (5.1.30)\\
&= \begin{pmatrix} \cos\left(\frac{g_W|\alpha|}{2}\right) + i\sin\left(\frac{g_W|\alpha|}{2}\right)\frac{\alpha^3}{|\alpha|} & \sin\left(\frac{g_W|\alpha|}{2}\right)\frac{\alpha^2+i\alpha^1}{|\alpha|} \\ \sin\left(\frac{g_W|\alpha|}{2}\right)\frac{-\alpha^2+i\alpha^1}{|\alpha|} & \cos\left(\frac{g_W|\alpha|}{2}\right) - i\sin\left(\frac{g_W|\alpha|}{2}\right)\frac{\alpha^3}{|\alpha|} \end{pmatrix}. && (5.1.31)
\end{aligned}$$

If we are not worried about the topological consequences of large gauge transformations (which we will discuss in detail later), we can use a constant and spatially uniform SU(2) transformation to guarantee that $\Phi_0$ is real and only has 1 non-zero component. In particular, if $\Phi_0$ before any gauge transformation has

components:

$$\Phi_0 = \begin{pmatrix} \phi_{03} + i\phi_{04} \\ \phi_{01} + i\phi_{02} \end{pmatrix} \tag{5.1.32}$$

and we select the gauge transformation parameters as:

$$\tan\left(\frac{g_W|\alpha|}{2}\right) = \frac{\sqrt{\phi_{02}^2 + \phi_{03}^2 + \phi_{04}^2}}{|\phi_{01}|} \tag{5.1.33}$$

$$\alpha^1 = -\frac{|\alpha|}{\sin\left(\frac{g_W|\alpha|}{2}\right)} \frac{\phi_{04}}{\sqrt{\phi_{01}^2 + \phi_{02}^2 + \phi_{03}^2 + \phi_{04}^2}} \tag{5.1.34}$$

$$\alpha^2 = -\frac{|\alpha|}{\sin\left(\frac{g_W|\alpha|}{2}\right)} \frac{\phi_{03}}{\sqrt{\phi_{01}^2 + \phi_{02}^2 + \phi_{03}^2 + \phi_{04}^2}} \tag{5.1.35}$$

$$\alpha^3 = \frac{|\alpha|}{\sin\left(\frac{g_W|\alpha|}{2}\right)} \frac{\phi_{02}}{\sqrt{\phi_{01}^2 + \phi_{02}^2 + \phi_{03}^2 + \phi_{04}^2}} \tag{5.1.36}$$

then after the gauge transformation we have:

$$\Phi_0' = U\Phi_0 = \begin{pmatrix} 0 \\ \sqrt{\phi_{01}^2 + \phi_{02}^2 + \phi_{03}^2 + \phi_{04}^2} \end{pmatrix}. \tag{5.1.37}$$

By (5.1.15) we must in particular have (dropping the $'$ going forward, assuming we have already made this transformation):

$$\Phi_0 = \begin{pmatrix} 0 \\ \sqrt{\frac{\mu^2}{2\lambda}} \end{pmatrix} \tag{5.1.38}$$

Going forward, it will be useful to work in components and so define $\phi_1, \phi_2, \phi_3, \phi_4$ as real fields satisfying:

$$\delta\Phi = \begin{pmatrix} \phi_3 + i\phi_4 \\ \phi_1 + i\phi_2 \end{pmatrix}. \tag{5.1.39}$$

Then, putting this expression together with our transformed $\Phi_0$ into (5.1.23) gives:

$$\begin{pmatrix} \partial^2\phi_3 + i\partial^2\phi_4 - i\frac{g_W}{2}\sqrt{\frac{\mu^2}{2\lambda}}\partial^\mu\delta W_\mu^1 - \frac{g_W}{2}\sqrt{\frac{\mu^2}{2\lambda}}\partial^\mu\delta W_\mu^2 \\ \partial^2\phi_1 + i\partial^2\phi_2 - ig_Y\sqrt{\frac{\mu^2}{2\lambda}}\partial^\mu\delta B_\mu + i\frac{g_W}{2}\sqrt{\frac{\mu^2}{2\lambda}}\partial^\mu\delta W_\mu^3 - 2\mu^2\phi_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{5.1.40}$$

Or, looking at the real and imaginary parts we have:

$$\partial^2\phi_1 - 2\mu^2\phi_1 = 0 \tag{5.1.41}$$

$$\partial^2\phi_2 + \sqrt{\frac{\mu^2}{2\lambda}}\partial^\mu\left(\frac{g_W}{2}\delta W_\mu^3 - g_Y\delta B_\mu\right) = 0 \tag{5.1.42}$$

$$\partial^2\phi_3 - \frac{g_W}{2}\sqrt{\frac{\mu^2}{2\lambda}}\partial^\mu\delta W_\mu^2 = 0 \tag{5.1.43}$$

$$\partial^2\phi_4 - \frac{g_W}{2}\sqrt{\frac{\mu^2}{2\lambda}}\partial^\mu\delta W_\mu^1 = 0. \tag{5.1.44}$$

So, we can see that the $\phi_1$ field evolves freely from the other fields and satisfies the Klein-Gordon equation with mass $m = \sqrt{2\mu^2}$. We will return to this shortly. The other components of $\Phi$ are not so simple. With our expression for $\Phi_0$ the $B_\mu$ equation of motion becomes:

$$\partial^2\delta B^\mu - \partial^\mu\partial^\nu\delta B_\nu + 2g_Y\sqrt{\frac{\mu^2}{2\lambda}}\partial^\mu\phi_2 - \frac{g_Y^2\mu^2}{\lambda}\delta B^\mu + g_Y g_W\frac{\mu^2}{2\lambda}\delta W^{3\mu} = 0. \tag{5.1.45}$$

As well, the equations of motion for the components of $W_\mu^a$ become:

$$\partial^2\delta W_1^\mu - \partial^\mu\partial^\nu\delta W_{1\nu} + g_W\sqrt{\frac{\mu^2}{2\lambda}}\partial^\mu\phi_4 - \frac{g_W^2\mu^2}{4\lambda}\delta W_1^\mu = 0 \tag{5.1.46}$$

$$\partial^2\delta W_2^\mu - \partial^\mu\partial^\nu\delta W_{2\nu} + g_W\sqrt{\frac{\mu^2}{2\lambda}}\partial^\mu\phi_3 - \frac{g_W^2\mu^2}{4\lambda}\delta W_2^\mu = 0 \tag{5.1.47}$$

$$\partial^2 \delta W_3^\mu - \partial^\mu \partial^\nu \delta W_{3\nu} - g_W \sqrt{\frac{\mu^2}{2\lambda}} \partial^\mu \phi_2 + g_W g_Y \frac{\mu^2}{2\lambda} \delta B^\mu - \frac{g_W^2 \mu^2}{4\lambda} \delta W_3^\mu = 0. \tag{5.1.48}$$

So, we see that $\phi_1$ satisfies an equation of motion all by itself, $\phi_2$, $\delta W_\mu^3$ and $\delta B_\mu$ all mix with each other in their equations of motion, $\phi_3$ and $W_\mu^2$ mix with each other, and $\phi_4$ and $W_\mu^1$ mix with each other. So far we have only applied a global gauge transformation to set $\Phi_0$ to be in the form described. However, we know we can select the gauge at each point in spacetime. Under a full gauge transformation:

$$\Phi' = U\Phi \implies \begin{pmatrix} \phi_3' + i\phi_4' \\ \sqrt{\frac{\mu^2}{2\lambda}} + \phi_1' + i\phi_2' \end{pmatrix} = U(X) \begin{pmatrix} \phi_3 + i\phi_4 \\ \sqrt{\frac{\mu^2}{2\lambda}} + \phi_1 + i\phi_2 \end{pmatrix}. \tag{5.1.49}$$

Using the exact same technique as before we can select $\alpha^a(X)$ as functions of position and time so that $\Phi'$ is real and only the bottom component is non-zero for all $X$. Then, after that gauge transformation:

$$\phi_2(X) = \phi_3(X) = \phi_4(X) = 0. \tag{5.1.50}$$

As well, it is conventional to define the reduced Higgs field $\phi(X)$ by:

$$\phi(X) = \sqrt{2}\phi_1 \implies \Phi = \begin{pmatrix} 0 \\ \sqrt{\frac{\mu^2}{2\lambda}} + \frac{\phi}{\sqrt{2}} \end{pmatrix}. \tag{5.1.51}$$

Consequently, in this gauge our set of 8 equations of motion simplify to:

$$\partial^2 \phi - 2\mu^2 \phi = 0 \tag{5.1.52}$$

$$\partial^\mu \left( \frac{g_W}{2} \delta W_\mu^3 - g_Y \delta B_\mu \right) = 0 \tag{5.1.53}$$

$$\partial^\mu \delta W_\mu^2 = 0 \tag{5.1.54}$$

$$\partial^\mu \delta W_\mu^1 = 0 \tag{5.1.55}$$

$$\partial^2 \delta B^\mu - \partial^\mu \partial^\nu \delta B_\nu - \frac{g_Y^2 \mu^2}{\lambda} \delta B^\mu + g_Y g_W \frac{\mu^2}{2\lambda} \delta W^{3\mu} = 0 \tag{5.1.56}$$

$$\partial^2 \delta W_1^\mu - \frac{g_W^2 \mu^2}{4\lambda} \delta W_1^\mu = 0 \tag{5.1.57}$$

$$\partial^2 \delta W_2^\mu - \frac{g_W^2 \mu^2}{4\lambda} \delta W_2^\mu = 0 \tag{5.1.58}$$

$$\partial^2 \delta W_3^\mu - \partial^\mu \partial^\nu \delta W_{3\nu} + g_W g_Y \frac{\mu^2}{2\lambda} \delta B^\mu - \frac{g_W^2 \mu^2}{4\lambda} \delta W_3^\mu = 0. \tag{5.1.59}$$

Thus, $\delta W_\mu^1$ and $\delta W_\mu^2$ both satisfy the Klein-Gordon equation with mass $m = \frac{g_W \mu}{2\lambda}$ and have vanishing 4-divergence. We can condense those 4 equations into 2 by defining the complex weak gauge field:

$$W_\mu = \frac{W_\mu^1 - iW_\mu^2}{\sqrt{2}} \tag{5.1.60}$$

$$\implies \partial^\mu \delta W_\mu = 0, \qquad \partial^2 \delta W_\mu - \frac{g_W^2 \mu^2}{4\lambda} \delta W_\mu = 0. \tag{5.1.61}$$

Even after gauge transformation, $\delta B_\mu$ and $\delta W_\mu^3$ are coupled. We can define a new set of variables to uncouple these equations of motion. To do so, first define the Weinberg angle (or weak angle) $\theta_W$ and electroweak coupling $g_{EW}$ by:

$$g_{EW} = \sqrt{g_W^2 + 4g_Y^2}, \qquad \sin\theta_W = \frac{2g_Y}{g_{EW}}, \qquad \cos\theta_W = \frac{g_W}{g_{EW}}. \tag{5.1.62}$$

Then, define the linear combinations $A_\mu$ and $Z_\mu$ by:

$$Z_\mu = W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W \tag{5.1.63}$$

$$A_\mu = B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W. \tag{5.1.64}$$

We can compute the equations of motion for these fields by making matching linear combinations of the equations of motion for $W_\mu^3$ and $B_\mu$. Doing so, we find that our 6 equations of motion now become:

$$\partial^2 \phi - 2\mu^2 \phi = 0 \tag{5.1.65}$$

$$\partial^\mu \delta Z_\mu = 0 \tag{5.1.66}$$

$$\partial^\mu \delta W_\mu = 0 \tag{5.1.67}$$

$$\partial_\nu(\partial^\mu \delta A^\nu - \partial^\nu \delta A^\mu) = 0 \tag{5.1.68}$$

$$\partial^2 \delta W_\mu - \frac{g_W^2 \mu^2}{4\lambda}\delta W_\mu = 0 \tag{5.1.69}$$

$$\partial^2 \delta Z_\mu - \frac{g_{EW}^2 \mu^2}{4\lambda}\delta Z_\mu = 0. \tag{5.1.70}$$

So, this theory has one scalar called the Higgs $\phi$ with mass:

$$m_H^2 = 2\mu^2 \tag{5.1.71}$$

two vectors (because $W_\mu$ has two independent components) called the W fields with mass:

$$m_W^2 = \frac{g_W^2 \mu^2}{4\lambda} \tag{5.1.72}$$

one vector called the Z field with mass:

$$m_Z^2 = \frac{g_{EW}^2 \mu^2}{4\lambda} \tag{5.1.73}$$

and one field $A_\mu$ which satisfies Maxwell's equations (with a current of 0). To get here, we have completely specified our SU(2) gauge but said nothing about our U(1) gauge. So, we expect to still have U(1) gauge symmetry. If we just apply the hypercharge U(1) transformation directly, we run into an issue. Because said transformation changes the complex phase of $\Phi$, it does not preserve the condition that $\Phi$ be real. This can be resolved if any addition to the U(1) hypercharge transformation we make an SU(2) transformation which restores $\Phi$ to being real while also keeping the upper entry of $\Phi$ as 0. Then, the conditions we have placed on our fields so far are preserved by this combined transformation. That is, we want to study the combined transformation:

$$\Phi' = e^{ig_Y \alpha_Y(X)} e^{ig_W \alpha^a \tau_a}\Phi \tag{5.1.74}$$

where $\alpha_Y$ and $\alpha^a$ are required to preserved the form of $\Phi$ in (5.1.51). Under this combined transformation, we immediately have:

$$\Phi' = e^{ig_Y\alpha_Y}\begin{pmatrix} \cos\left(\frac{g_W|\alpha|}{2}\right) + i\sin\left(\frac{g_W|\alpha|}{2}\right)\frac{\alpha^3}{|\alpha|} & \sin\left(\frac{g_W|\alpha|}{2}\right)\frac{\alpha^2+i\alpha^1}{|\alpha|} \\ \sin\left(\frac{g_W|\alpha|}{2}\right)\frac{-\alpha^2+i\alpha^1}{|\alpha|} & \cos\left(\frac{g_W|\alpha|}{2}\right) - i\sin\left(\frac{g_W|\alpha|}{2}\right)\frac{\alpha^3}{|\alpha|}\end{pmatrix}\begin{pmatrix} 0 \\ \sqrt{\frac{\mu^2}{2\lambda}} + \frac{\phi}{\sqrt{2}}\end{pmatrix} \tag{5.1.75}$$

$$= \begin{pmatrix} e^{ig_Y\alpha_Y}\sin\left(\frac{g_W|\alpha|}{2}\right)\frac{\alpha^2+i\alpha^1}{|\alpha|}\left(\sqrt{\frac{\mu^2}{2\lambda}} + \frac{\phi}{\sqrt{2}}\right) \\ e^{ig_Y\alpha_Y}\left(\cos\left(\frac{g_W|\alpha|}{2}\right) - i\sin\left(\frac{g_W|\alpha|}{2}\right)\frac{\alpha^3}{|\alpha|}\right)\left(\sqrt{\frac{\mu^2}{2\lambda}} + \frac{\phi}{\sqrt{2}}\right)\end{pmatrix}. \tag{5.1.76}$$

For the upper component to vanish we need $\alpha^1 = \alpha^2 = 0$. Then, after simplifying we have:

$$\Phi' = \begin{pmatrix} 0 \\ e^{ig_Y\alpha_Y - \frac{i}{2}g_W\alpha^3}\left(\sqrt{\frac{\mu^2}{2\lambda}} + \frac{\phi}{\sqrt{2}}\right)\end{pmatrix}. \tag{5.1.77}$$

In order for the lower component to remain real then we can simply take:

$$\alpha^3 = \frac{2g_Y}{g_W}\alpha_Y. \tag{5.1.78}$$

Then, under this transformation we actually see that the reduced Higgs field is unchanged:

$$\phi' = \phi. \tag{5.1.79}$$

As well, we can compute how the hypercharge gauge field changes using (4.4.355):

$$B'_\mu = B_\mu + \partial_\mu \alpha_Y \tag{5.1.80}$$

and we can compute how the weak gauge fields change as:

$$W'^a_\mu \tau_a = e^{ig_Y\alpha_Y\sigma_3}W^a_\mu\tau_a e^{-ig_Y\alpha_Y\sigma_3} + \frac{i}{g_W}e^{ig_Y\alpha_Y\sigma_3}\partial_\mu\left(e^{-ig_Y\alpha_Y\sigma_3}\right) \tag{5.1.81}$$

$$\begin{pmatrix} W'^3_\mu & W'^1_\mu - iW'^2_\mu \\ W'^1_\mu + iW'^2_\mu & -W'^3_\mu\end{pmatrix} = \begin{pmatrix} W^3_\mu & e^{2ig_Y\alpha_Y}(W^1_\mu - iW^2_\mu) \\ e^{-2ig_Y\alpha_Y}(W^1_\mu + iW^2_\mu) & -W^3_\mu\end{pmatrix} + \frac{2g_Y}{g_W}\partial_\mu\alpha_Y\begin{pmatrix}1 & 0 \\ 0 & -1\end{pmatrix} \tag{5.1.82}$$

$$\implies W'^1_\mu - iW'^2_\mu = e^{2ig_Y\alpha_Y}(W^1_\mu - iW^2_\mu), \qquad W'^3_\mu = W^3_\mu + \frac{2g_Y}{g_W}\partial_\mu\alpha_Y. \tag{5.1.83}$$

Together, these imply that the W, Z, and A fields change under the transformation as:

$$W'_\mu = e^{2ig_Y\alpha_Y}W_\mu \tag{5.1.84}$$

$$Z'_\mu = Z_\mu \tag{5.1.85}$$

$$A'_\mu = A_\mu + \frac{\partial_\mu\alpha_Y}{\cos\theta_W}. \tag{5.1.86}$$

This looks exactly like a U(1) transformation where now $A_\mu$ is interpreted as the gauge field and the gauge transformation function $f$ is given by:

$$f = \frac{\alpha_Y}{\cos\theta_W} \tag{5.1.87}$$

$$\implies A'_\mu = A_\mu + \partial_\mu f. \tag{5.1.88}$$

Then:

$$W'_\mu = e^{2ig_Y\cos\theta_W f}W_\mu. \tag{5.1.89}$$

This is a U(1) transformation where we identify the charge of the $W_\mu$ field (called $\mathtt{e}$, the fundamental charge) as:

$$\mathtt{e} = 2g_Y\cos\theta_W. \tag{5.1.90}$$

Thus, after fixing the SU(2) gauge this theory has a new U(1) symmetry given by this specific combination of the old U(1) symmetry and an appropriate SU(2) transformation. We call the old U(1) symmetry, the one corresponding to the $B_\mu$ field, the $\mathrm{U_Y}(1)$ or hypercharge U(1) transformation. We call the new U(1) symmetry, the one corresponding to the $A_\mu$ field, the $\mathrm{U_{EM}}(1)$ or electromagnetic U(1) transformation. The full Lagrangian in (5.1.9) has symmetry under $\mathrm{U_Y}(1)$ and SU(2) and after gauge fixing SU(2) it only has $\mathrm{U_{EM}}(1)$ symmetry. We express this as:

$$\mathrm{U_Y}(1)\times\mathrm{SU}(2)\to\mathrm{U_{EM}}(1). \tag{5.1.91}$$

The full Lagrangian density in (5.1.9) (called the electroweak Lagrangian density) has the full symmetry group on the left hand side here but when we consider the theory near stable equilibrium, that equilibrium state does not have that symmetry group. That is, symmetry transformations of the equilibrium state take us to a different equilibrium state rather than the same one. So, because they equilibrium state does not have the full symmetry group, when we expand near by the resulting effective theory (described by the equations of motion (5.1.65)-(5.1.70)) also does not have that full symmetry group. The effective theory only has the symmetries which are shared between the full Lagrangian density and the equilibrium state and that set is the symmetry group appearing on the right hand side above. This mechanism is called "spontaneous symmetry breaking". It happens whenever one studies a system near a stable equilibrium point which has fewer symmetries than the full theory of the system itself. This stable equilibrium is often called "the vacuum" and so the fact that in stable equilibrium the $\Phi$ field is non-zero is referred to as the $\Phi$ field acquiring a "vacuum expectation value" or vev (pronounced as a word, rather than as an acronym) for short. By definition, the vev v of the Higgs field is defined as:

$$\mathtt{v} = \frac{\mu^2}{\lambda} \tag{5.1.92}$$

The symmetry broken theory produces massive vector fields in the form of the W and Z fields. This is precisely what was meant earlier when we mentioned that the weak interaction somehow has "massive gauge fields". The full electroweak Lagrangian density just has the ordinary terms for a U(1) and an SU(2) gauge field. Some of the fields "acquire mass" during the spontaneous symmetry breaking that occurs when one expands the theory near stable equilibrium. In terms of the Higgs vev, Higgs mass, Weinberg angle, and fundamental charge we can express all of the parameters of the electroweak Lagrangian density as:

$$g_Y = \frac{\mathtt{e}}{2\cos\theta_W}, \qquad g_W = \frac{\mathtt{e}}{\sin\theta_W}, \qquad g_{EW} = \frac{\mathtt{e}}{\sin\theta_W\cos\theta_W} \tag{5.1.93}$$

$$\mu^2 = \frac{m_H^2}{2}, \qquad \lambda = \frac{m_H^2}{2\mathtt{v}^2} \tag{5.1.94}$$

$$m_W = \frac{\mathtt{ev}}{2\sin\theta_W}, \qquad m_Z = \frac{\mathtt{ev}}{2\sin\theta_W\cos\theta_W}. \tag{5.1.95}$$

We can complete the square in the electroweak Lagrangian density to put it into a form where it is more transparent to see the presence of the stable equilibrium:

$$\mathfrak{L} = -\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi + \frac{m_H^2}{2}\Phi^\dagger\Phi - \frac{m_H^2}{2\mathrm{v}^2}(\Phi^\dagger\Phi)^2 - \frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}) \tag{5.1.96}$$

$$= -\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi - \frac{m_H^2}{2\mathrm{v}^2}\left(\Phi^\dagger\Phi - \frac{\mathrm{v}^2}{2}\right)^2 + \frac{m_H^2\mathrm{v}^2}{8} - \frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}) \tag{5.1.97}$$

$$\simeq -\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi - \frac{m_H^2}{2\mathrm{v}^2}\left(\Phi^\dagger\Phi - \frac{\mathrm{v}^2}{2}\right)^2 - \frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}). \tag{5.1.98}$$

Thus, going forward we will actually take the electroweak Lagrangian density to be:

$$\mathfrak{L} = -\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi - \frac{m_H^2}{2\mathrm{v}^2}\left(\Phi^\dagger\Phi - \frac{\mathrm{v}^2}{2}\right)^2 - \frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}). \tag{5.1.99}$$

Because in experiment we study scenarios in which we perturb away from stable equilibrium, it is still most useful for many purposes to express this Lagrangian density in terms of the W, Z, and A fields as they are what are relevant in that perturbation. However, we will do so now exactly without assuming small fields. That is, we will only assume we have applied our SU(2) gauge fixing so that we have:

$$\Phi = \begin{pmatrix} 0 \\ \frac{\mathrm{v}+\phi}{\sqrt{2}} \end{pmatrix}, \qquad W_\mu^1 = \frac{W_\mu + W_\mu^*}{\sqrt{2}}, \qquad W_\mu^2 = \frac{i}{\sqrt{2}}(W_\mu - W_\mu^*), \tag{5.1.100}$$

$$W_\mu^3 = Z_\mu\cos\theta_W + A_\mu\sin\theta_W, \qquad B_\mu = A_\mu\cos\theta_W - Z_\mu\sin\theta_W. \tag{5.1.101}$$

With this gauge choice, $\phi_2 = \phi_3 = \phi_4 = 0$ which are holonomic constraints we can directly insert into the Lagrangian density. However, this will prevent us from being able to recover the information they gave us in their equations of motion from the Lagrangian (as all of their terms will just be set to 0 identically). So, while we can directly insert this form into the Lagrangian density to find correct equations of motion, it will not find us a complete set. In addition to the equations of motion gotten from the Lagrangian with this gauge choice, we must also have the equations gotten for general $\Phi$ evaluated at $\phi_2 = \phi_3 = \phi_4 = 0$. That is, take (5.1.12) and apply this gauge choice. Doing so gives that the $\phi_2, \phi_3, \phi_4$ equations of motion give:

$$\partial^\mu Z_\mu + \frac{2\partial^\mu\phi}{\mathrm{v}+\phi}Z_\mu = 0 \tag{5.1.102}$$

$$\mathscr{D}^\mu W_\mu + i\mathtt{e}\tan\theta_W Z^\mu W_\mu + \frac{2\partial^\mu\phi}{\mathrm{v}+\phi}W_\mu = 0 \tag{5.1.103}$$

where the covariant derivative of the W field is defined exactly as one would expect given its gauge transformation properties:

$$\mathscr{D}_\mu W_\nu = \partial_\mu W_\nu - i\mathtt{e}A_\mu W_\nu. \tag{5.1.104}$$

Thus, we can see that the (covariant) divergence of the W field and the divergence of the Z field are both constrained even when considering fields of arbitrary strength, not just small fields. However, for large fields those divergences are non-zero. These constraints we cannot simply insert into the Lagrangian density as they involve the time derivatives of the fields and so are non-holonomic. As well, if we in the future add more terms to the electroweak Lagrangian density which involve the Higgs field $\Phi$ then such terms will modify these non-holonomic constraints as they will affect the equations of motion for $\phi_2, \phi_3, \phi_4$. In order to apply the holonomic constraints from gauge fixing, we need to fully expand out the electroweak Lagrangian density, which is quite messy. Considering just the field strength part we have:

$$\begin{aligned}
-\frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}) \simeq\ & -\frac{1}{2}\partial^\mu B^\nu\partial_\mu B_\nu + \frac{1}{2}(\partial^\mu B_\mu)^2 - \frac{1}{2}\partial^\mu W^{1\nu}\partial_\mu W_\nu^1 + \frac{1}{2}(\partial^\mu W_\mu^1)^2 \\
& - \frac{1}{2}\partial^\mu W^{2\nu}\partial_\mu W_\nu^2 + \frac{1}{2}(\partial^\mu W_\mu^2)^2 - \frac{1}{2}\partial^\mu W^{3\nu}\partial_\mu W_\nu^3 + \frac{1}{2}(\partial^\mu W_\mu^3)^2 \\
& - g_W(\partial_\mu W_\nu^1 - \partial_\nu W_\mu^1)W^{2\mu}W^{3\nu} - g_W(\partial_\mu W_\nu^2 - \partial_\nu W_\mu^2)W^{3\mu}W^{1\nu} \\
& - g_W(\partial_\mu W_\nu^3 - \partial_\mu W_\nu^3)W^{1\mu}W^{2\nu} - \frac{g_W^2}{2}W^{2\mu}W_\mu^2W^{3\nu}W_\nu^3 + \frac{g_W^2}{2}(W^{2\mu}W_\mu^3)^2 \\
& - \frac{g_W^2}{2}W^{1\mu}W_\mu^1W^{3\nu}W_\nu^3 + \frac{g_W^2}{2}(W^{3\mu}W_{1\mu})^2 - \frac{g_W^2}{2}W^{1\mu}W_\mu^1W^{2\nu}W_\nu^2 + \frac{g_W^2}{2}(W^{1\mu}W_\mu^2)^2
\end{aligned} \tag{5.1.105}$$

$$\simeq -\frac{1}{4}\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu} - \frac{1}{2}\partial^\mu Z^\nu \partial_\mu Z_\nu + \frac{1}{2}\left(\partial^\mu Z_\mu\right)^2 - \mathscr{D}_\mu W_\nu^* \mathscr{D}^\mu W^\nu + \mathscr{D}^\mu W_\mu^* \mathscr{D}^\nu W_\nu$$
$$- 2i\mathtt{e} W_\mu^* \mathcal{F}^{\mu\nu} W_\nu - 2g_W \cos(\theta_W) Z^\nu \mathrm{Im}\left(2W_\mu^* \mathscr{D}^\mu W_\nu - W_\nu^* \mathscr{D}^\mu W_\mu - W_\mu^* \mathscr{D}_\nu W^\mu\right)$$
$$+ g_W \cos^2(\theta_W)(W_\mu^* Z^\mu W_\nu Z^\nu - W_\mu^* W^\mu Z_\nu Z^\nu) + \frac{g_W^2}{2}(W^\mu W_\mu W^{\nu *} W_\nu^* - (W_\mu^* W^\mu)^2) \quad (5.1.106)$$

where we have used the standard definition of the electromagnetic field strength:

$$\mathcal{F}_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (5.1.107)$$

The Higgs pieces of the electroweak Lagrangian density are much simpler:

$$-\mathscr{D}^\mu \Phi^\dagger \mathscr{D}_\mu \Phi - \frac{m_H^2}{2\mathtt{v}^2}\left(\Phi^\dagger\Phi - \frac{\mathtt{v}^2}{2}\right)^2 = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{m_H^2}{2}\phi^2\left(1+\frac{\phi}{2\mathtt{v}}\right)^2$$
$$- m_W^2 W_\mu^* W^\mu \left(1+\frac{\phi}{\mathtt{v}}\right)^2 - \frac{m_Z^2}{2} Z^\mu Z_\mu \left(1+\frac{\phi}{\mathtt{v}}\right)^2. \quad (5.1.108)$$

Thus, after gauge fixing the electroweak Lagrangian density becomes:

$$\mathfrak{L} = -\mathscr{D}^\mu \Phi^\dagger \mathscr{D}_\mu \Phi - \frac{m_H^2}{2\mathtt{v}^2}\left(\Phi^\dagger\Phi - \frac{\mathtt{v}^2}{2}\right)^2 - \frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}) \quad (5.1.109)$$

$$\simeq -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{m_H^2}{2}\phi^2\left(1+\frac{\phi}{2\mathtt{v}}\right)^2 - \frac{1}{4}\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu} - \frac{1}{2}\partial^\mu Z^\nu\partial_\mu Z_\nu + \frac{1}{2}\left(\partial^\mu Z_\mu\right)^2 - \frac{m_Z^2}{2}Z^\mu Z_\mu\left(1+\frac{\phi}{\mathtt{v}}\right)^2$$
$$- \mathscr{D}_\mu W_\nu^* \mathscr{D}^\mu W^\nu + \mathscr{D}^\mu W_\mu^* \mathscr{D}^\nu W_\nu - m_W^2 W_\mu^* W^\mu\left(1+\frac{\phi}{\mathtt{v}}\right)^2 - 2i\mathtt{e}W_\mu^*\mathcal{F}^{\mu\nu}W_\nu$$
$$- 2g_W\cos(\theta_W)Z^\nu\mathrm{Im}\left(2W_\mu^*\mathscr{D}^\mu W_\nu - W_\nu^*\mathscr{D}^\mu W_\mu - W_\mu^*\mathscr{D}_\nu W^\mu\right)$$
$$+ g_W\cos^2(\theta_W)(W_\mu^* Z^\mu W_\nu Z^\nu - W_\mu^* W^\mu Z_\nu Z^\nu) + \frac{g_W^2}{2}(W^\mu W_\mu W^{\nu *} W_\nu^* - (W_\mu^* W^\mu)^2) \quad (5.1.110)$$

where the $W$ and $Z$ fields additionally satisfy the constraints imposed by the $\phi_2, \phi_3, \phi_4$ equations of motion from the un-gauge-fixed Lagrangian density.

Now let us consider another field $\psi$ which is supposed to interact with this theory. In order to do so, $\psi$ will have to be in some representation of $\mathrm{U_Y}(1)$ and some representation of SU(2). For the hypercharge transformation, because it is just a U(1), we know that for any real number Y we could validly have $\psi$ transform as:

$$\psi' = e^{iYg_Y\alpha_Y}\psi \quad (5.1.111)$$

and so, this $Y$ is a property of the field $\psi$ called its hypercharge. As well, we already know from our study of spin how to enumerate all of the representations of SU(2). In particular, we know that an SU(2) representation is easily labeled by the size of the generator matrices in that representation. So, $\psi$ will be in some representation $d$ of SU(2) where the generators $\tau_a$ have dimensions $d \times d$ as matrices. This $d$ is called the isospin rep (short for representation) of $\psi$. For $d = 1$ $\psi$ is often called an isospin singlet, for $d = 2$ and isospin doublet, and so on. From our construction of SU(2) we know that $\tau_3$ is diagonal for arbitrary $d$. Let $I_n$ be the $n^{\text{th}}$ diagonal entry of $\tau_3$. This $I_n$ is called the isospin of $\psi^n$ (the $n^{\text{th}}$ isospin component of $\psi$). Then:

$$\tau_3\psi = \tau_3\begin{pmatrix}\psi^1\\ \dots \\ \psi^d\end{pmatrix} = \begin{pmatrix}I_1\psi^1\\ \dots \\ I_d\psi^d\end{pmatrix}. \quad (5.1.112)$$

After spontaneous symmetry breaking, we are left only with $\mathrm{U_{EM}}(1)$ symmetry. Then, under the electromagnetic U(1) transformation:

$$\psi' = e^{iYg_Y\alpha_Y}e^{ig_W\alpha^3\tau_3}\psi = e^{i\frac{\mathtt{e}Y}{2}f}e^{i\mathtt{e}\tau_3 f}\psi = \begin{pmatrix}e^{i(I_1+\frac{Y}{2})\mathtt{e}f}\psi^1\\ \dots \\ e^{i(I_d+\frac{Y}{2})\mathtt{e}f}\psi^d\end{pmatrix} \quad (5.1.113)$$

Thus $\psi^n$ has precisely the $\mathrm{U_{EM}}(1)$ transformation of a field with electric charge $I_n + \frac{Y}{2}$. So, the electric charge $Q$ of a field is related to its hypercharge $Y$ and isospin $I$ by:

$$Q = I + \frac{Y}{2}. \quad (5.1.114)$$

This is also reflected in the covariant derivative of $\psi$:

$$\mathscr{D}_\mu\psi = \partial_\mu\psi - iYg_YB_\mu\psi - ig_WW_\mu^a\tau_a\psi \tag{5.1.115}$$

$$= \partial_\mu\psi - i\mathsf{e}A_\mu\left(\tau_3 + \frac{Y}{2}\mathbb{1}\right)\psi + ieZ_\mu\left(\frac{Y}{2}\tan\theta_W\mathbb{1} - \frac{\tau_3}{\tan\theta_W}\right)\psi - ig_W\left(W_\mu\frac{\tau_1+i\tau_2}{\sqrt{2}} + W_\mu^*\frac{\tau_1-i\tau_2}{\sqrt{2}}\right)\psi. \tag{5.1.116}$$

Here we can see that the covariant derivative of $\psi^n$ involves $A_\mu$ as $-i\mathsf{e}\left(I_n + \frac{Y}{2}\right)A_\mu\psi^n$ and thus the identification of $Q$ is exactly what we found before.

The Higgs mechanism as we have seen so far is able to explain how there are massive vector fields arising from a gauge theory. So, it has fixed one problem. However, it has also caused another one. The weak interaction is observed to only interact with the left handed components of the electron's Dirac field and not the right handed. This of course violates parity symmetry and so under circumstances in which the weak interaction is noticeable, parity symmetry is broken. That is, the right handed $\psi_R$ is an isospin singlet and does not transform at all under SU(2) (and so has no terms in its covariant derivative to interact with the weak gauge field) but the left handed $\psi_L$ does have a coupling term to the weak gauge fields and so must not be an isospin singlet. Thus, the left handed electron must be one isospin component of a representation of some size $d$. All fields in this representation of size $d$ must have the same hypercharge in order to preserve the $\mathrm{U_Y}(1)$ symmetry. By (5.1.114) then their electric charges different only by their isospins $I$. But, for the generators of SU(2) in the representation of dimension $d$ the eigenvalues are $-\frac{d-1}{2}$ up to $\frac{d+1}{2}$ counting by adding one to the previous eigenvalue. So, the electric charges of these different fields all have to differ by integers. If the left handed electron were in an isospin representation with any fields of non-zero electric charge then we would experimentally see other electrically charged particles with charge different from the electron charge by integer amounts. However, we don't see such charges. Consequently, the left handed electron must be in a representation only with fields of electric charge 0. Because each value of $I$ is unique in any particular representation, there can only be one other such field and thus the left handed electron must be in an isospin doublet. The other field in the doublet is called the left handed electron neutrino. The left handed electron has negative charge (and thus lower than the left handed electron neutrino) and so it must be in the lower component of the doublet. Consequently, the left handed electron neutrino has isospin $I = \frac{1}{2}$ (due to being in the upper component of the doublet) and in order to have charge 0 the electron neutrino has hypercharge $Y = -1$. Then, because the the electron is in the lower component it has isospin $I = -\frac{1}{2}$ and the same hypercharge $Y = -1$ and thus it has charge $Q = -1$. So, the classical electric charge of the electron is $q = Q\mathsf{e} = -\mathsf{e}$ and this $\mathsf{e}$ is equivalent to the previously identified electric charge. We write the left and right handed electron fields (which themselves are two complex components due to being spin 1/2) as $e_L(X)$ and $e_R(X)$. We write the left and right handed electron neutrino fields (which again are two complex components due to being spin 1/2. Also, this is assuming right handed neutrino exists which are currently experimentally unconfirmed; to remove them simply set their field to 0) as $\nu_{eL}$ and $\nu_{eR}$. Because $\psi_L$ is now an isospin doublet:

$$\psi_L = \begin{pmatrix} e_L \\ \nu_{eL} \end{pmatrix}. \tag{5.1.117}$$

As well, $\psi_R$ can be split into two isospin singlets, $e_R$ and $\nu_{eR}$. But, it is not necessary or meaningful to organize them into a matrix as we have done with $\psi_L$ due to them being isospin singlets. Consequently, the traditional Dirac mass term is not SU(2) gauge invariant:

$$-m\bar{\psi}\psi = -m\psi_L^\dagger\psi_R - m\psi_R^\dagger\psi_L \tag{5.1.118}$$

and in fact is not a single complex number or even well defined as $\psi_R$ isn't specified. So, how do we get massive electrons? Just like with the massiveness of the W and Z, the weak interaction introduces a problem and the Higgs solves it. While $\psi_R^\dagger\psi_L$ is not gauge invariant, the combination $e_R^\dagger\Phi^\dagger\psi_L$ is because both $\Phi$ and $\psi_L$ are isospin doublets and so their inner product is an SU(2) invariant number. As well, we need to check $\mathrm{U_Y}(1)$ gauge invariant of this term. The $\mathrm{U_Y}(1)$ transformation of this term is:

$$e_R'^\dagger\Phi'^\dagger\psi_L' = \left(e^{iY_{e_R}g_Y\alpha_Y}e_R\right)\dagger\left(e^{iY_\Phi g_Y\alpha_Y}\Phi\right)^\dagger\left(e^{iY_{\psi_L}g_Y\alpha_Y}\psi_L\right) = e^{i(Y_{\psi_L}-Y_\Phi-Y_{e_R})g_Y\alpha_Y}e_R^\dagger\Phi^\dagger\psi_L. \tag{5.1.119}$$

To have gauge invariant then, we need:

$$Y_{e_R} = Y_{\psi_L} - Y_\Phi. \tag{5.1.120}$$

By the way we normalized the covariant derivative of the Higgs, we can see that the hypercharge of the Higgs is $Y = 1$ and earlier we found $\psi_L$ has hypercharge $-1$. Thus, $Y_{e_R}$ must have hypercharge $Y = -2$ in

order for this term to be gauge invariant. Because $I = 0$ for the right handed electron (as it is an isospin singlet), this predicts the right handed electron to have charge $Q = -1$, matching the left handed electron and thus reproducing the parity symmetry of the electromagnetic interaction. This explains why we cannot add a term $\nu_{eR}^\dagger \Phi^\dagger \psi_L$ as because $\nu_{eR}$ also must have $Q = 0$ (like the left handed neutrino) and it has $I = 0$ (due to being an isospin singlet) and so it has $Y = 0$ and this term is not hypercharge gauge invariant.

With the above term (and its equally allowed complex conjugate) we can see that after SU(2) gauge fixing, it produces the Dirac mass term for the electron (letting $\Gamma_e$ here be some complex coefficient):

$$-\Gamma_e e_R^\dagger \Phi^\dagger \psi_L - \Gamma_e^* \psi_L^\dagger \Phi e_R = -\Gamma_e e_R^\dagger \begin{pmatrix} 0 & \frac{\mathrm{v}+\phi}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix} - \Gamma_e^* \begin{pmatrix} \nu_{eL}^\dagger & e_L^\dagger \end{pmatrix} \begin{pmatrix} 0 \\ \frac{\mathrm{v}+\phi}{\sqrt{2}} \end{pmatrix} e_R \tag{5.1.121}$$

$$= -\frac{\Gamma_e \mathrm{v}}{\sqrt{2}} e_R^\dagger e_L - \frac{\Gamma_e^* \mathrm{v}}{\sqrt{2}} e_L^\dagger e_R - \frac{\Gamma_e}{\sqrt{2}} e_R^\dagger e_L \phi - \frac{\Gamma_e}{\sqrt{2}} e_L^\dagger e_R \phi. \tag{5.1.122}$$

For $\Gamma_e$ real and positive, these first two terms are the ordinary Dirac mass term with electron mass $m_e = \frac{\Gamma_e \mathrm{v}}{\sqrt{2}}$. The last two terms above just give an interaction between the right handed electron, left handed electron, and the Higgs. If $\Gamma_e$ is not both real and positive, we can actually make it so by adjusting the definitions of our fields very slightly. In particular, due to the global U(1) symmetry of our Dirac Lagrangian density (meaning symmetry under applying a uniform and constant phase shift) and the fact that $\Gamma_e \psi_L$ is the only combination of $\Gamma_e$ or $\psi_L$ in the above terms, we can simply redefine $\psi_L$ by an overall phase to make $\Gamma_e$ real and positive:

$$\psi_L' = e^{i\mathrm{Arg}(\Gamma_e)} \psi_L \implies \Gamma_e \psi_L = |\Gamma_e| e^{i\mathrm{Arg}(\Gamma_e)} \psi_L = |\Gamma_e| \psi_L'. \tag{5.1.123}$$

Thus, in terms of the transformed $\psi_L$ we have that $\Gamma_e$ is transformed to $|\Gamma_e|$. So, we can take $\Gamma_e \geq 0$ by definition.

What about mass terms for the neutrinos? Experimentally, they are necessary as well. However, if one simply adds the above type terms with $e_R$ replaced by $\nu_{eR}$, not only will the term not have the proper hypercharge symmetry but also it will just be zero after SU(2) gauge fixing. So, that is certainly not the way to do it. It turns out we can also add mass terms for the neutrinos, we just have to be slightly more clever and use a special property of SU(2). Recall that $\sigma_a^* = -\sigma_2 \sigma_a \sigma_2$. Then, it immediately follows that under SU(2) transformation:

$$(\sigma_2 \Phi^*)' = \sigma_2 U^* \Phi^* = \sigma_2 U^* \sigma_2 \sigma_2 \Phi^* = \sigma_2 e^{-\frac{i}{2} g_W \alpha_W^a \sigma_a^*} \sigma_2 \Phi^* = e^{\frac{i}{2} g_W \alpha_W^a \sigma_a} \sigma_2 \Phi^* = U(\sigma_2 \Phi^*). \tag{5.1.124}$$

Hence, the combination $\sigma_2 \Phi^*$ has the same SU(2) transformation property as $\Phi$ and we can use it to construct the term $i\Gamma_\nu \nu_{eR}^\dagger \Phi^T \sigma_2 \psi_L$ for our Lagrangian density as it has the correct SU(2) transformation property. As well, we need to check this term has the correct hypercharge transformation property. As we found, $\psi_L$ has a hypercharge of $-1$ and the Higgs has a hypercharge of 1. So, the combination $\Phi^T \psi_L$ or equivalently $\psi_L^\dagger \Phi^*$ has a hypercharge of 0 as the phases immediately cancel out. Because $\nu_{eR}$ has a hypercharge of 0, this means the above terms have the correct hypercharge transformation and so they are fully allowed terms in our Lagrangian density. After SU(2) gauge fixing, we can see that this term (and its equally allowed complex conjugate) produce mass terms for the neutrinos:

$$i\Gamma_\nu \nu_{eR}^\dagger \Phi^T \sigma_2 \psi_L - i\Gamma_\nu^* \psi_L^\dagger \sigma_2 \Phi^* \nu_{eR} = i\Gamma_\nu \nu_{eR}^\dagger \begin{pmatrix} 0 & \frac{\mathrm{v}+\phi}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix} - i\Gamma_\nu^* \begin{pmatrix} \nu_{eL}^\dagger & e_L^\dagger \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{\mathrm{v}+\phi}{\sqrt{2}} \end{pmatrix} \nu_{eR} \tag{5.1.125}$$

$$= -\frac{\Gamma_\nu \mathrm{v}}{\sqrt{2}} \nu_{eR}^\dagger \nu_{eL} - \frac{\Gamma_\nu^* \mathrm{v}}{\sqrt{2}} \nu_{eL}^\dagger \nu_{eR} - \frac{\Gamma_\nu}{\sqrt{2}} \nu_{eR}^\dagger \nu_{eL} \phi - \frac{\Gamma_\nu^*}{\sqrt{2}} \nu_{eL}^\dagger \nu_{eR} \phi. \tag{5.1.126}$$

For $\Gamma_\nu$ real and positive, these first two terms are the ordinary Dirac mass term with neutrino mass $m_\nu = \frac{\Gamma_\nu \mathrm{v}}{\sqrt{2}}$. The last two terms above just give an interaction between the right handed neutrino, left handed neutrino, and the Higgs. If $\Gamma_\nu$ is not both real and positive, we can make it so just like we did with $\Gamma_e$ by adjusting the definitions of our fields very slightly. In particular, due to the global U(1) symmetry of the electroweak Lagrangian density and the fact that $\Gamma_\nu \Phi^T$ is the only combination of $\Gamma_\nu$ or $\Phi$ in the above terms, we can simply redefine $\Phi$ by an overall phase to make $\Gamma_\nu$ real and positive:

$$\Phi' = e^{i\mathrm{Arg}(\Gamma_\nu)} \Phi \implies \Gamma_\nu \Phi^T = |\Gamma_\nu| e^{i\mathrm{Arg}(\Gamma_\nu)} \Phi^T = |\Gamma_\nu| \Phi'^T. \tag{5.1.127}$$

Thus, in terms of the transformed $\Phi$ we have that $\Gamma_\nu$ is transformed to $|\Gamma_\nu|$. So, we can take $\Gamma_\nu \geq 0$ by definition. The terms in (5.1.122) and (5.1.126) are called Yukawa couplings and we see that they fix the

problem of having massive electrons and neutrinos while also having SU(2) gauge symmetry.

One may worry that potentially the other terms in the Dirac Lagrangian encounter an issue now that the left handed parts are SU(2) doublets and the right handed parts are SU(2) singlets. But, expanding out the derivative terms or "kinetic" terms in the Dirac Lagrangian density we find in fact that in any specific term one finds only left handed components or right handed components and so there is no issue of SU(2) symmetry:

$$i\bar{\psi}\gamma^\mu\partial_\mu\psi = i\psi^\dagger\gamma^0\gamma^0\dot{\psi} + i\psi^\dagger\gamma^0\gamma^a\partial_a\psi = i\psi_L^\dagger\dot{\psi}_L - i\psi_L^\dagger\sigma^a\partial_a\psi_L + i\psi_R^\dagger\dot{\psi}_R + i\psi_R^\dagger\sigma^a\partial_a\psi_R. \tag{5.1.128}$$

In order to write all of the kinetic terms then for the left and right handed versions of the electron and electron neutrino we have:

$$\mathfrak{L}_{\text{kinetic}} = ie_L^\dagger\dot{e}_L - ie_L^\dagger\sigma^a\partial_a e_L + ie_R^\dagger\dot{e}_R + ie_R^\dagger\sigma^a\partial_a e_R + i\nu_{eL}^\dagger\dot{\nu}_{eL} - i\nu_{eL}^\dagger\sigma^a\partial_a\nu_{eL} + i\nu_{eR}^\dagger\dot{\nu}_{eR} + i\nu_{eR}^\dagger\sigma^a\partial_a\nu_{eR}. \tag{5.1.129}$$

The electron and the electron neutrino are called "leptons" and so it is common to define $\ell_L(X)$ as another name for $\psi_L(X)$ we have used before:

$$\ell_L(X) = \psi_L = \begin{pmatrix} e_L \\ \nu_{eL} \end{pmatrix}. \tag{5.1.130}$$

It is okay to define $\ell_R$ analogously as long as one remembers that the above $\ell_L$ is a genuine isospin doublet whereas $\ell_R$ would simply be two unrelated isospin singlets. With $\ell_L$ we can equivalently write our kinetic terms as:

$$\mathfrak{L}_{\text{kinetic}} = i\ell_L^\dagger\dot{\ell}_L - i\ell_L^\dagger\sigma^a\partial_a\ell_L + ie_R^\dagger\dot{e}_R + ie_R^\dagger\sigma^a\partial_a e_R + i\nu_{eR}^\dagger\dot{\nu}_{eR} + i\nu_{eR}^\dagger\sigma^a\partial_a\nu_{eR} \tag{5.1.131}$$

where now the SU(2) invariance of the Lagrangian density is manifest. (As written the above Lagrangian density is only invariant under global SU(2) transformations, one of course needs to modify the derivatives into covariant derivatives to make it invariant under local SU(2) transformations.) One can also define the lepton field $\ell(X)$ by:

$$\ell(X) = \begin{pmatrix} \ell_L \\ \nu_{eR} \\ e_R \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix} \\ \nu_{eR} \\ e_R \end{pmatrix} \tag{5.1.132}$$

so that the kinetic terms can be expressed simply as:

$$\mathfrak{L}_{\text{kinetic}} = i\bar{\ell}\gamma^\mu\partial_\mu\ell \tag{5.1.133}$$

where the $\gamma^\mu$ matrices are understood to act on the pairs $e_L$, $e_R$ and $\nu_{eL}$, $\nu_{eR}$ separately. With this and appropriate covariant derivatives the Lagrangian density for the leptons becomes:

$$\mathfrak{L} = i\bar{\ell}\gamma^\mu\mathscr{D}_\mu\ell - \Gamma_e e_R^\dagger\Phi^\dagger\ell_L - \Gamma_e\ell_L^\dagger\Phi e_R + i\Gamma_\nu\nu_{eR}^\dagger\Phi^T\sigma_2\ell_L - i\Gamma_\nu^*\ell_L^\dagger\sigma_2\Phi^*\nu_{eR} \tag{5.1.134}$$

$$= i\bar{\ell}\gamma^\mu\mathscr{D}_\mu\ell - 2\Gamma_e\text{Re}(e_R^\dagger\Phi^\dagger\ell_L) - 2\Gamma_\nu\text{Im}(\nu_{eR}^\dagger\Phi^T\sigma_2\ell_L). \tag{5.1.135}$$

For the right handed neutrinos there is actually another mass term we can add to the Lagrangian density. Previously we defined the charge conjugation property of a Dirac spinor and found that the charge conjugated spinor has the same Lorentz transformation properties as the unconjugated spinor. Consequently, because the term $\bar{\psi}\psi$ is a Lorentz scalar, the term $\bar{\psi}_\mathcal{C}\psi$ is as well. Explicitly for a Dirac spinor this term is:

$$\bar{\psi}_\mathcal{C}\psi = (-i\gamma^2\psi^*)^\dagger\gamma^0\psi = i\psi^T\gamma^0\gamma^2\psi. \tag{5.1.136}$$

Because this term has $\psi$ and $\psi^T$ and no $\psi^\dagger$ or $\psi^*$, it does not have U(1) symmetry. This is a problem for all of our fields with non-zero electric charge or hypercharge but uniquely the right handed neutrino has $Q = Y = 0$ and so we do not need to require the Lagrangian density to have symmetry under U(1) transformations of it. Consequently, the Dirac spinor consisting of just the right handed neutrino with the above term produces a valid term to consider in the Lagrangian density:

$$\psi = \begin{pmatrix} 0 \\ \nu_{eR} \end{pmatrix} \implies \bar{\psi}_\mathcal{C}\psi = i\nu_{eR}^T\sigma_2\nu_{eR}. \tag{5.1.137}$$

As well, to add this term to the Lagrangian density we need to add its complex conjugate and so the valid combination we can consider is the "Majorana mass term":

$$-\frac{M\bar{\psi}_\mathcal{C}\psi + M^*\bar{\psi}\psi_\mathcal{C}}{2} = -\frac{iM}{2}\nu_{eR}^T\sigma_2\nu_{eR} + \frac{iM^*}{2}\nu_{eR}^\dagger\sigma_2\nu_{eR}^*. \tag{5.1.138}$$

While these terms do not have a global U(1) symmetry in transformations of the right handed neutrino, the other terms in the Lagrangian density which involve $\nu_{eR}$ are the derivative terms for a free $\nu_{eR}$ and the Yukawa coupling between $\nu_{eR}$, $\Phi$, and $\psi_L$; the free $\nu_{eR}$ derivative terms do actually have global U(1) symmetry and any global phase transformation of $\nu_{eR}$ in the Yukawa coupling can be absorbed into the definition of $\Phi$ so that that term has global U(1) symmetry as well. Consequently, we can adjust the phase of $\nu_{eR}$ arbitrarily. Doing so, we can make $M$ real and positive by a now familiar technique:

$$\nu'_{eR} = e^{\frac{i}{2}\text{Arg}(M)}\nu_{eR} \implies \frac{iM}{2}\nu_{eR}^T\sigma_2\nu_{eR} = \frac{i|M|}{2}e^{i\text{Arg}(M)}\nu_{eR}^T\sigma_2\nu_{eR} = \frac{i|M|}{2}\nu_{eR}'^T\sigma_2\nu'_{eR}. \tag{5.1.139}$$

Thus, we take by definition $M > 0$. This $M$ is called the "Majorana mass" of $\nu_{eR}$. If we expand the Majorana mass term explicitly in terms of the spin up and spin down components of $\nu_{eR}$, naively it appears to be 0:

$$-i\nu_{eR}^T\sigma_2\nu_{eR} = -\nu_{eR+}\nu_{eR-} + \nu_{eR-}\nu_{eR+}. \tag{5.1.140}$$

If we assume these products commute, the two terms cancel. However, recall that the Dirac spinors will eventually be treated as linear operators (in a way we haven't really explained yet) and those linear operators will anti-commute (meaning swapping their order changes the sign of the expression) rather than commute (the way that complex numbers do). So, this expression is in fact non-zero for such operators (but as a classical field theory of complex valued fields it is in fact 0). So, our Lagrangian density for the leptons is then:

$$\mathfrak{L} = i\bar{\ell}\gamma^\mu\mathscr{D}_\mu\ell - 2\Gamma_e\text{Re}(e_R^\dagger\Phi^\dagger\ell_L) - 2\Gamma_\nu\text{Im}(\nu_{eR}^\dagger\Phi^T\sigma_2\ell_L) + M\text{Im}(\nu_{eR}^T\sigma_2\nu_{eR}). \tag{5.1.141}$$

As we have mentioned before, there are other Dirac spinors in the universe known as quarks. The quarks interact via the strong interaction and comprise the proton and the neutron. The proton, quite mysteriously, seems to have charge exactly equal and opposite that of the electron (and thus has electric charge 1) and the neutron, equally mysteriously, seems to be exactly charge neutral (and thus has electric charge 0). Scattering off of the proton and neutron reveals that they are comprised of 3 quarks each of two different types which we call the up quark $u$ and the down quark $d$. The up and down quarks are Dirac spinors and so we have $u$ comprised of $u_L$ and $u_R$ as well as $d$ comprised of $d_L$ and $d_R$. The proton consists of two up quarks and one down quark while the neutron consists of one up quark and two down quarks. So, if $Q_u$ is the electric charge of the up quark and $Q_d$ the electric charge of the down quark:

$$Q_u + Q_u + Q_d = 1, \qquad Q_u + Q_d + Q_d = 0 \tag{5.1.142}$$

$$\implies Q_u = \frac{2}{3}, \qquad Q_d = -\frac{1}{3}. \tag{5.1.143}$$

Because the quarks are electrically charged, they must interact via some combination of the hypercharge and weak interactions. Just as with the leptons, the left handed quarks form an isospin doublet while the right handed quarks form two isospin singlets. The up quark has higher electric charge than the down quark and so it must go in the upper component of the isospin doublet and the left handed down quark must go in the lower component (hence the names up and down). Thus, $u_L$ has isospin $I = \frac{1}{2}$ and $d_L$ has isospin $I = -\frac{1}{2}$. Given their electric charges we find by (5.1.114) that they both have hypercharge $Y = \frac{1}{3}$. The right handed quarks both are isospin singlets and so have $I = 0$. Then, given their charges $u_R$ has hypercharge $Y = \frac{4}{3}$ and $d_R$ has hypercharge $Y = -\frac{2}{3}$. Due to their SU(2) transformations, the quarks encounter all of the same problems with mass terms that the leptons do. But, we can fix those problems with the Higgs in exactly the same way we did with the leptons, via Yukawa couplings. Because none of the quarks have a hypercharge of 0, they cannot have Majorana masses. So, defining:

$$\mathfrak{q}_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \qquad \mathfrak{q} = \begin{pmatrix} \mathfrak{q}_L \\ u_R \\ d_R \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} u_L \\ d_L \end{pmatrix} \\ u_R \\ d_R \end{pmatrix} \tag{5.1.144}$$

we have that the Lagrangian density for the quarks is given by the same form as the Lagrangian density for the leptons but with the neutrinos replaced by up quarks and the electrons replaced by down quarks:

$$\mathfrak{L} = i\bar{\mathfrak{q}}\gamma^\mu\mathscr{D}_\mu\mathfrak{q} - 2\Gamma_d\text{Re}(d_R^\dagger\Phi^\dagger\mathfrak{q}_L) - 2\Gamma_u\text{Im}(u_R^\dagger\Phi^T\sigma_2\mathfrak{q}_L) \tag{5.1.145}$$

where we can by convention set the overall phases of $u_R$ and $d_R$ so that $\Gamma_u$ and $\Gamma_d$ are real and positive. We should be careful about the Yukawa couplings for the quarks. The Yukawa couplings are by construction SU(2) invariant but we need to also check that they are $\text{U}_\text{Y}(1)$ invariant. The $\Gamma_d$ Yukawa term $d_R^\dagger\Phi^\dagger\mathfrak{q}_L$ has hyper charge given by the hypercharge $\mathfrak{q}_L$ minus the hypercharge of $\Phi$ minus the hypercharge of $d_R$. Thus, it

has hypercharge $\frac{1}{3}-1-(-\frac{2}{3})=0$ and so is $\mathrm{U_Y}$ invariant. The $\Gamma_u$ Yukawa term $u_R^\dagger \Phi^T \sigma_2 \mathfrak{q}_L$ has hypercharge given by the sum of the hypercharges of $\Phi$ and $\mathfrak{q}_L$ minus the hypercharge of $u_R$. Thus, it has hypercharge $1+\frac{1}{3}-\frac{4}{3}=0$ and so is also $\mathrm{U_Y}(1)$ invariant and both of these Yukawa terms are just what we wanted. This combination of hypercharges explains why we do not write down Yukawa terms mixing leptons as quarks; there is no way to do so while preserving hypercharge symmetry.

All together then we can write the combination of our electroweak Lagrangian density together with the lepton and quark Lagrangian densities and the nuclear strong interaction to get a full interaction theory of the hypercharge, weak, and strong interactions together with the Higgs, electron, electron neutrino, up and down quark:

$$\begin{aligned}\mathfrak{L} =& -\mathscr{D}^\mu \Phi^\dagger \mathscr{D}_\mu \Phi - \frac{m_H^2}{2\mathtt{v}^2}\left(\Phi^\dagger\Phi - \frac{\mathtt{v}^2}{2}\right)^2 - \frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}) - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_S^{\mu\nu}\mathcal{F}_{S\mu\nu})\\ &+ i\bar{\ell}\gamma^\mu \mathscr{D}_\mu \ell - 2\Gamma_e \mathrm{Re}(e_R^\dagger \Phi^\dagger \ell_L) - 2\Gamma_\nu \mathrm{Im}(\nu_{eR}^\dagger \Phi^T \sigma_2 \ell_L) + M\mathrm{Im}(\nu_{eR}^T \sigma_2 \nu_{eR})\\ &+ i\bar{\mathfrak{q}}\gamma^\mu \mathscr{D}_\mu \mathfrak{q} - 2\Gamma_d \mathrm{Re}(d_R^\dagger \Phi \mathfrak{q}_L) - 2\Gamma_u \mathrm{Im}(u_R^\dagger \Phi^T \sigma_2 \mathfrak{q}_L).\end{aligned} \tag{5.1.146}$$

This Lagrangian density then fully has $\mathrm{U_Y}(1)$ hypercharge, SU(2) weak, and SU(3) strong gauge symmetry. The only free parameters of this theory are the electromagnetic coupling constant $\mathtt{e}$, Weinberg angle $\theta_W$, strong coupling constant $g_S$, Higgs mass $m_H$, Higgs vev $\mathtt{v}$, electron Yukawa coupling $\Gamma_e$, neutrino Yukawa coupling $\Gamma_\nu$, neutrino Majorana mass $M$, down quark Yukawa coupling $\Gamma_d$, and up quark Yukawa coupling $\Gamma_u$ giving a total of 10 quantities which need to be determined by experiment for this theory.

### Exercises

**Problem 5.1: gauge invariant expressions for W, Z, A and full expansion of the lagrangian density without breaking gauge invariance**

(a)

## 5.2 Generations of Matter and Mass Mixing Matrices

There is another interesting feature of the interactions of elementary fields; not only is there an up quark and a down quark, an electron neutrino and an electron but there are in fact 3 fields for each of these sharing all of the same gauge couplings. That is, there are essentially three copies of the set up quark, down quark, electron neutrino, electron. These copies are called generations of matter and so we say we observe 3 generations of matter. In particular, there is an up quark, a charm quark, and a top quark all of which are Dirac spinors which couple in all of the same ways to all gauge fields. There is a down quark, a strange quark, and a bottom quark all of which are Dirac spinors which couple in all of the same ways to all gauge fields. There is an electron neutrino, a muon neutrino, and a tau neutrino all of which are Dirac spinors which couple in all of the same ways to all gauge fields. And, there is an electron, a muon, and a tauon all of which are Dirac spinors which couple in all of the same ways to all gauge fields. So, we represent these fields as $u_a$, $d_a$, $\nu_a$, and $e_a$ respectively with $a$ ranging from 1 to 3. Then, because these fields have matching transformation rules under Lorentz and gauge transformations across generations, there is a whole matrix of allowed Yukawa terms. Experimentally, these three generations of matter have been observed except for the right handed neutrinos. No right handed neutrinos have ever been observed (because they don't interact through the electromagnetic, weak, or strong interactions if they exist they only interact through the Yukawa interaction and gravity never of which are not strong enough to have so far had an experimentally noticeable effect) but they are the simplest mechanism to explain the existence of neutrino mass, which has been observed. So, we do add them as a possibility though in principle the number of right handed neutrinos may be anything, not necessarily equal to 3 like the other generations. The left handed neutrino masses haven't been measured well either (only differences of masses squared have been measured at this time) but there is enough data to conclude that at least two of the generations have non zero mass and if this mass is to be explained by right handed neutrinos, this requires at least 2 right handed neutrinos. We will for simplicity and essentially without further justification assume there are 3 generations of right handed neutrinos (one can get the minimal case of 2 by setting one of them to be identically 0; we are ignoring recklessly the possibility that there are more than 3). That is, for some $3\times 3$ complex matrices with components $\Gamma_e^{ab}$, $\Gamma_\nu^{ab}$, $M^{ab}$, $\Gamma_u^{ab}$,

$\Gamma_d^{ab}$:

$$\begin{aligned}\mathfrak{L} =& -\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi - \frac{m_H^2}{2\mathrm{v}^2}\left(\Phi^\dagger\Phi - \frac{\mathrm{v}^2}{2}\right)^2 - \frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}) - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_S^{\mu\nu}\mathcal{F}_{S\mu\nu})\\ &+ i\bar{\ell}_a\gamma^\mu\mathscr{D}_\mu\ell_a - 2\mathrm{Re}(\Gamma_e^{ab}e_{Ra}^\dagger\Phi^\dagger\ell_{Lb}) - 2\mathrm{Im}(\Gamma_\nu^{ab}\nu_{Ra}^\dagger\Phi^T\sigma_2\ell_{Lb}) + \mathrm{Im}(M^{ab}\nu_{Ra}^T\sigma_2\nu_{Rb})\\ &+ i\bar{\mathfrak{q}}_a\gamma^\mu\mathscr{D}_\mu\mathfrak{q}_a - 2\mathrm{Re}(\Gamma_d^{ab}d_{Ra}^\dagger\Phi\mathfrak{q}_{Lb}) - 2\mathrm{Im}(\Gamma_u^{ab}u_{Ra}^\dagger\Phi^T\sigma_2\mathfrak{q}_{Lb}). \end{aligned} \tag{5.2.1}$$

(in principle one could also consider linear combinations of the kinetic terms $i\bar{\psi}\gamma^\mu\mathscr{D}_\mu\psi$ but the fact that the Lagrangian density must be real up to a surface term enforces that one make a linear combination of such terms with a self-adjoint matrix of coefficients, then by redefining the lepton and quark fields by appropriate linear combinations one can transform that matrix into the identity matrix.) The Yukawa matrices $\Gamma_e, \Gamma_\nu, \Gamma_d, \Gamma_u$ can be arbitrary complex $3\times 3$ matrices. The Majorana matrix $M^{ab}$ may be complex but must be symmetric. This is because $\nu_{Ra}^T\sigma_2\nu_{Rb} = \nu_{Rb}^T\sigma_2\nu_{Ra}$ which can be proven using by computing the transpose of either side and using that the neutrino fields anti-commute with each other. When we were considering only 1 generation, we were able to fix these matrices to simply each be a real number by appropriate phase redefinitions of the fields. We can do something similar in the case of more than one generation. Consider defining new quark and lepton fields by making some linear combinations:

$$\ell_{La} = U_a{}^b\ell'_{Lb}, \qquad \nu_{Ra} = V_a{}^b\nu'_{Rb}, \qquad e_{Ra} = W_a{}^b e'_{Rb} \tag{5.2.2}$$

$$\mathfrak{q}_{La} = X_a{}^b\mathfrak{q}'_{Lb}, \qquad u_{Ra} = Y_a{}^b u'_{Rb}, \qquad d_{Ra} = Z_a{}^b d'_{Rb} \tag{5.2.3}$$

for $3\times 3$ matrices $U, V, W, X, Y, Z$. These linear combinations are the most general linear combinations which can be made without mixing fields of different transformation properties (such a transformation could not possibly preserve the form of the Lagrangian density). Expanding out the kinetic terms of the Lagrangian density we have:

$$\begin{aligned}&i\bar{\ell}_a\gamma^\mu\mathscr{D}_\mu\ell_a + i\bar{\mathfrak{q}}_a\gamma^\mu\mathscr{D}_\mu\mathfrak{q}_a\\ &\quad= i\ell_{La}^\dagger\mathscr{D}_0\ell_{La} - i\ell_{La}^\dagger\sigma^b\mathscr{D}_b\ell_{La} + i\nu_{Ra}^\dagger\mathscr{D}_0\nu_{Ra} + i\nu_{Ra}^\dagger\sigma^b\mathscr{D}_b\nu_{Ra} + ie_{Ra}^\dagger\mathscr{D}_0 e_{Ra} + ie_{Ra}^\dagger\sigma^b\mathscr{D}_b e_{Ra}\\ &\qquad + i\mathfrak{q}_{La}^\dagger\mathscr{D}_0\mathfrak{q}_{La} - i\mathfrak{q}_{La}^\dagger\sigma^b\mathscr{D}_b\mathfrak{q}_{La} + iu_{Ra}^\dagger\mathscr{D}_0 u_{Ra} + iu_{Ra}^\dagger\sigma^b\mathscr{D}_b u_{Ra} + id_{Ra}^\dagger\mathscr{D}_0 d_{Ra} + id_{Ra}^\dagger\sigma^b\mathscr{D}_b d_{Ra}. \end{aligned} \tag{5.2.4}$$

To condense this expression it is useful to define:

$$\sigma_L^\mu = (\mathbb{1}, -\sigma_1, -\sigma_2, -\sigma_3)^\mu, \qquad \sigma_R^\mu = (\mathbb{1}, \sigma_1, \sigma_2, \sigma_3)^\mu. \tag{5.2.5}$$

Then we have:

$$\begin{aligned}&i\bar{\ell}_a\gamma^\mu\mathscr{D}_\mu\ell_a + i\bar{\mathfrak{q}}_a\gamma^\mu\mathscr{D}_\mu\mathfrak{q}_a\\ &\quad= i\ell_{La}^\dagger\sigma_L^\mu\mathscr{D}_\mu\ell_{La} + i\nu_{Ra}^\dagger\sigma_R^\mu\mathscr{D}_\mu\nu_{Ra} + ie_{Ra}^\dagger\sigma_R^\mu\mathscr{D}_\mu e_{Ra} + i\mathfrak{q}_{La}^\dagger\sigma_L^\mu\mathscr{D}_\mu\mathfrak{q}_{La} + iu_{Ra}^\dagger\sigma_R^\mu\mathscr{D}_\mu u_{Ra} + id_{Ra}^\dagger\sigma_R^\mu\mathscr{D}_\mu d_{Ra} \end{aligned} \tag{5.2.6}$$

$$\begin{aligned}&\quad= i\ell'^\dagger_{La}U_b{}^{a*}\sigma_L^\mu\mathscr{D}_\mu U_b{}^c\ell'_{Lc} + i\nu'^\dagger_{Ra}V_b{}^{a*}\sigma_R^\mu V_b{}^c\mathscr{D}_\mu\nu'_{Rc} + ie'^\dagger_{Ra}W_b{}^{a*}\sigma_R^\mu W_b{}^c\mathscr{D}_\mu e'_{Rc}\\ &\qquad + i\mathfrak{q}'^\dagger_{La}X_b{}^{a*}\sigma_L^\mu X_b{}^c\mathscr{D}_\mu\mathfrak{q}'_{Lc} + iu'^\dagger_{Ra}Y_b{}^{a*}\sigma_R^\mu Y_b{}^c\mathscr{D}_\mu u'_{Rc} + id'^\dagger_{Ra}Z_b{}^{a*}\sigma_R^\mu Z_b{}^c\mathscr{D}_\mu d'_{Rc}. \end{aligned} \tag{5.2.7}$$

In order for this transformation to not change the form of the kinetic terms then we see that each of our matrices must be unitary:

$$i\bar{\ell}_a\gamma^\mu\mathscr{D}_\mu\ell_a + i\bar{\mathfrak{q}}_a\gamma^\mu\mathscr{D}_\mu\mathfrak{q}_a \overset{!}{=} i\bar{\ell}'_a\gamma^\mu\mathscr{D}_\mu\ell'_a + i\bar{\mathfrak{q}}'_a\gamma^\mu\mathscr{D}_\mu\mathfrak{q}'_a \tag{5.2.8}$$

$$\implies U_b{}^{a*}U_b{}^c = \delta^{ac}, \qquad V_b{}^{a*}V_b{}^c = \delta^{ac}, \qquad W_b{}^{a*}W_b{}^c = \delta^{ac},$$
$$X_b{}^{a*}X_b{}^c = \delta^{ac}, \qquad Y_b{}^{a*}Y_b{}^c = \delta^{ac}, \qquad Z_b{}^{a*}Z_b{}^c = \delta^{ac} \tag{5.2.9}$$

$$\implies U^\dagger U = V^\dagger V = W^\dagger W = X^\dagger X = Y^\dagger Y = Z^\dagger Z = \mathbb{1}. \tag{5.2.10}$$

Thus, we can redefine our quarks and leptons in this way without changing the kinetic terms. Doing so turns the Yukawa terms into:

$$\begin{aligned}&-2\mathrm{Re}(\Gamma_e^{ab}e_{Ra}^\dagger\Phi^\dagger\ell_{Lb}) - 2\mathrm{Im}(\Gamma_\nu^{ab}\nu_{Ra}^\dagger\Phi^T\sigma_2\ell_{Lb}) + \mathrm{Im}(M^{ab}\nu_{Ra}^T\sigma_2\nu_{Rb}) - 2\mathrm{Re}(\Gamma_d^{ab}d_{Ra}^\dagger\Phi\mathfrak{q}_{Lb}) - 2\mathrm{Im}(\Gamma_u^{ab}u_{Ra}^\dagger\Phi^T\sigma_2\mathfrak{q}_{Lb})\\ &\quad= -2\mathrm{Re}(\Gamma_e^{cd}W_c{}^{a*}e_{Ra}^\dagger\Phi^\dagger U_d{}^b\ell_{Lb}) - 2\mathrm{Im}(\Gamma_\nu^{cd}V_c{}^{a*}\nu_{Ra}^\dagger\Phi^T\sigma_2 U_d{}^b\ell_{Lb}) + \mathrm{Im}(M^{cd}V_c{}^a\nu_{Ra}^T\sigma_2 V_d{}^b\nu_{Rb})\\ &\qquad - 2\mathrm{Re}(\Gamma_d^{cd}Z_c{}^{a*}d_{Ra}^\dagger\Phi X_d{}^b\mathfrak{q}_{Lb}) - 2\mathrm{Im}(\Gamma_u^{cd}Y_c{}^{a*}u_{Ra}^\dagger\Phi^T\sigma_2 X_d{}^b\mathfrak{q}_{Lb}). \end{aligned} \tag{5.2.11}$$

Then, the components of the Yukawa matrices are equivalently replaced with:

$$\Gamma_e'^{ab} = \Gamma_e^{cd}W_c{}^{a*}U_d{}^b, \qquad \Gamma_\nu'^{ab} = \Gamma_\nu^{cd}V_c{}^{a*}U_d{}^b, \qquad M'^{ab} = M^{cd}V_c{}^aV_d{}^b,$$

$$\Gamma'^{ab}_d = \Gamma^{cd}_d Z_c{}^{a*} X_d{}^b, \qquad \Gamma'^{ab}_u = \Gamma^{cd}_u Y_c{}^{a*} X_d{}^b. \tag{5.2.12}$$

Or, equivalently as matrices:

$$\Gamma'_e = W^\dagger \Gamma_e U, \qquad \Gamma'_\nu = V^\dagger \Gamma_\nu U, \qquad M' = V^T M V, \qquad \Gamma'_d = Z^\dagger \Gamma_d X, \qquad \Gamma'_u = Y^\dagger \Gamma_u X. \tag{5.2.13}$$

Thus we can fix some components of the Yukawa matrices by choosing our unitary matrices $U, V, W, X, Y, Z$ wisely. Consider a generic $N$x$N$ matrix $A$. Then, $AA^\dagger$ is automatically self-adjoint:

$$\left(AA^\dagger\right)^\dagger = (A^\dagger)^\dagger A^\dagger = AA^\dagger. \tag{5.2.14}$$

Thus, by the spectral theorem it has an orthonormal basis of eigenvectors and real eigenvalues. Let $\lambda$ be an eigenvalue of $AA^\dagger$ and $v$ be a vector with that eigenvalue:

$$AA^\dagger v = \lambda v \implies v^\dagger AA^\dagger v = \lambda v^\dagger v \implies \left|A^\dagger v\right|^2 = \lambda |v|^2 \tag{5.2.15}$$

$$\implies \lambda = \frac{\left|A^\dagger v\right|^2}{|v|^2} \geq 0. \tag{5.2.16}$$

Hence, the eigenvalues of $AA^\dagger$ are necessarily nonnegative. Define then the matrix $D$ as the diagonal matrix with the same eigenbasis as $AA^\dagger$ but with the positive square roots of the eigenvalues of $AA^\dagger$ along its diagonal and define the matrix $J$ as the unitary change of basis matrix which diagonalizes $AA^\dagger$. Then, in its diagonal basis the matrix elements of $AA^\dagger$ are just those of $D^2$ and so we have:

$$AA^\dagger = JD^2J^\dagger. \tag{5.2.17}$$

If none of the eigenvalues of $AA^\dagger$ are 0, then define the matrix $K$ by:

$$K = A^\dagger J D^{-1}. \tag{5.2.18}$$

Then:

$$K^\dagger K = D^{-1} J^\dagger AA^\dagger J D^{-1} = D^{-1} J^\dagger J D^2 J^\dagger J D^{-1} = D^{-1} D^2 D^{-1} = \mathbb{1}. \tag{5.2.19}$$

Hence, $K$ is unitary. As well, we can use the definition of $K$ to solve for $A$ as:

$$A^\dagger = KDJ^\dagger \tag{5.2.20}$$

$$\implies A = JDK^\dagger. \tag{5.2.21}$$

Hence, any invertible complex square matrix can be written in this form for unitary matrices $J$ and $K$ and a diagonal matrix $D$ with only positive eigenvalues. This proof can be extended to non-invertible complex square matrices in which case $D$ will have 0 as an eigenvalue (the rank of $D$ is in fact equal to the rank of $A$). This decomposition is called the singular value decomposition of $A$ and the eigenvalues of $D$ are called the singular values of $A$.

Importantly, the singular value decomposition is not unique. We defined $J$ as the change of basis matrix which diagonalizes $AA^\dagger$ however the orthonormal eigenvectors of a self-adjoint operator are only defined up to reordering those vectors, multiplying each vector by an arbitrary complex phase, or taking arbitrary linear combinations of vectors which share the same eigenvalue. So, $J$ is not uniquely defined up to those points. We by convention define $J$ so that the eigenvalues of $D$ are increasing along the diagonal. This settles the reordering issue. We will be working with matrices generically without degenerate eigenvalues here and so we do not have to worry about the ambiguous definition of the corresponding eigenvectors on the degenerate subspace (repeated eigenvalues in the Yukawa matrices would produce different fields of the same mass and such a mass degeneracy is not observed for any of the leptons). This still leaves open the ability to redefine each of the eigenvectors by an arbitrary phase. Let $\iota_n$ be the column vector with 1 in position $n$ and 0 everywhere else and let $\lambda_n$ be the eigenvalue of $D^2$ in position $n$. Then:

$$D^2 \iota_n = \lambda_n \iota_n \tag{5.2.22}$$

$$\implies J^\dagger AA^\dagger J \iota_n = \lambda_n \iota_n \tag{5.2.23}$$

$$\implies AA^\dagger (J\iota_n) = \lambda_n (J\iota_n) \tag{5.2.24}$$

and so we take eigenvector $n$ of $AA^\dagger$, called $v_n$, to be given by:

$$v_n = J\iota_n. \tag{5.2.25}$$

Under a redefinition by arbitrary phase then we must have that for some real $\varphi_n$ and unitary $J'$:

$$v'_n = J'\iota_n = e^{i\varphi_n} v_n. \tag{5.2.26}$$

Define then the matrix $P$ which is diagonal with $e^{i\varphi_n}$ in diagonal entry $n$. Then:

$$P\iota_n = e^{i\varphi_n}\iota_n \tag{5.2.27}$$

$$J'\iota_n = e^{i\varphi_n}v_n = e^{i\varphi_n}J\iota_n = JP\iota_n \tag{5.2.28}$$

$$\implies J' = JP. \tag{5.2.29}$$

So, the ambiguity in the definition of $J$ is captured by the ability to multiply it on the right by an arbitrary diagonal matrix of phases as seen here. Then, this induces a transformation on $K$:

$$K' = A^\dagger J' D^{-1} = A^\dagger JPD^{-1} = A^\dagger JD^{-1}P = KP \tag{5.2.30}$$

where because $D$ and $P$ are simultaneously diagonal, they commute.

Now then, using the singular value decomposition we know that for unitary matrices $J_e$, $J_\nu$, $J_M$, $J_d$, $J_u$, $K_e$, $K_\nu$, $K_M$, $K_d$, $K_u$ and diagonal non-negative definite matrices $D_e, D_\nu, D_M, D_d, D_u$, our Yukawa matrices can be written as:

$$\Gamma_e = J_e D_e K_e^\dagger, \qquad \Gamma_\nu = J_\nu D_\nu K_\nu^\dagger, \qquad M = J_M D_M K_M^\dagger, \qquad \Gamma_d = J_d D_d K_d^\dagger, \qquad \Gamma_u = J_u D_u K_u^\dagger. \tag{5.2.31}$$

Because $M$ is a complex symmetric matrix, it can actually be diagonalized by a single unitary matrix using a variant of the spectral theorem [26]. This allows us to take $K_M = J_M^*$. With this, our ability to redefine the Yukawa matrices gives:

$$\Gamma_e' = W^\dagger J_e D_e K_e^\dagger U, \qquad \Gamma_\nu' = V^\dagger J_\nu D_\nu K_\nu^\dagger U, \qquad M' = V^T J_M D_M J_M^T V,$$
$$\Gamma_d' = Z^\dagger J_d D_d K_d^\dagger X, \qquad \Gamma_u' = Y^\dagger J_u D_u K_u^\dagger X. \tag{5.2.32}$$

Let us choose then:

$$W = J_e, \qquad U = K_e. \tag{5.2.33}$$

Automatically, this diagonalizes $\Gamma_e'$:

$$\Gamma_e' = D_e. \tag{5.2.34}$$

As well, $\Gamma_\nu'$ now becomes:

$$\Gamma_\nu' = V^\dagger J_\nu D_\nu K_\nu^\dagger K_e. \tag{5.2.35}$$

It is not possible in general to diagonalize this expression with just the remaining ability to choose $V$. However, we can write it as a unitary change of basis of a diagonal matrix. Define $U_\nu$ by:

$$U_\nu = K_e^\dagger K_\nu. \tag{5.2.36}$$

Then, let $V$ be given by:

$$V = J_\nu K_\nu^\dagger K_e \tag{5.2.37}$$

$$\implies \Gamma_\nu' = U_\nu D_\nu U_\nu^\dagger. \tag{5.2.38}$$

So, if we choose $\Gamma_e'$ to be diagonal (which we do by convention), $\Gamma_\nu'$ cannot be in general made diagonal but it can be written as a unitary change of basis away from a diagonal matrix. Then, because we have set $U$, $V$, and $W$ we have $M'$ as is. So, define $U_M$ by:

$$U_M = V^T J_M \tag{5.2.39}$$

$$\implies M' = J_M D_M U_M^T. \tag{5.2.40}$$

This tells us nothing as it is simply the generic diagonalization of a complex symmetric matrix $M'$. Hence, $M'$ is just an arbitrary complex symmetric matrix.

While for the leptons we by convention diagonalize the electron Yukawa matrix, for the quarks we by convention diagonalize the up quark Yukawa matrix. So, we choose:

$$Y = J_u, \qquad X = K_u. \tag{5.2.41}$$

Then:

$$\Gamma_u' = D_u. \tag{5.2.42}$$

As well, $\Gamma_d'$ now becomes:

$$\Gamma_d' = Z^\dagger J_d D_d K_d^\dagger K_u. \tag{5.2.43}$$

Just like with the neutrino Yukawa matrix, it is not possible to diagonalize this matrix in general now that we have diagonalized the up quark Yukawa matrix. Though, we can likewise write it as a unitary change of

basis of a diagonal matrix. Define $U_d$ by:

$$U_d = K_u^\dagger K_d. \tag{5.2.44}$$

Then, let $Z$ be given by:

$$Z = J_d K_d^\dagger K_u \tag{5.2.45}$$

$$\implies \Gamma_d' = U_d D_d U_d^\dagger. \tag{5.2.46}$$

So, after fixing $U$, $V$, $W$, $X$, $Y$, $Z$, we have our transformed Yukawa matrices as (dropping the primes, now that we have completed the transformation):

$$\Gamma_e = D_e, \qquad \Gamma_\nu = U_\nu D_\nu U_\nu^\dagger, \qquad \Gamma_d = U_d D_d U_d^\dagger, \qquad \Gamma_u = D_u. \tag{5.2.47}$$

However, recall that the singular value decomposition is not unique and so we can introduce diagonal phase matrices so that:

$$K_e' = K_e P_e, \qquad K_\nu' = K_\nu P_\nu, \qquad K_d' = K_d P_d, \qquad K_u' = K_u P_u. \tag{5.2.48}$$

Using the definitions of $U_\nu$ and $U_d$ this gives:

$$U_\nu' = P_e^\dagger U_\nu P_\nu, \qquad U_d' = P_u^\dagger U_d P_d. \tag{5.2.49}$$

A generic $3 \times 3$ matrix $A$ can be written for $r_{ab} > 0$ and $\varphi_{ab}$ real as:

$$A = \begin{pmatrix} r_{11}e^{i\varphi_{11}} & r_{12}e^{i\varphi_{12}} & r_{13}e^{i\varphi_{13}} \\ r_{21}e^{i\varphi_{21}} & r_{22}e^{i\varphi_{22}} & r_{23}e^{i\varphi_{23}} \\ r_{31}e^{i\varphi_{31}} & r_{32}e^{i\varphi_{32}} & r_{33}e^{i\varphi_{33}} \end{pmatrix}. \tag{5.2.50}$$

Requiring $A$ to be unitary then imposes that each row of $A$ taken as a complex vector is a unit vector and is orthogonal to the other rows of $A$. The same condition is automatically true of the columns of $A$. If we multiply $A$ on the right by a diagonal phase matrix we have:

$$B \stackrel{\text{def}}{=} AP = A \begin{pmatrix} e^{i\alpha_1} & 0 & 0 \\ 0 & e^{i\alpha_2} & 0 \\ 0 & 0 & e^{i\alpha_3} \end{pmatrix} = \begin{pmatrix} r_{11}e^{i(\varphi_{11}+\alpha_1)} & r_{12}e^{i(\varphi_{12}+\alpha_2)} & r_{13}e^{i(\varphi_{13}+\alpha_3)} \\ r_{21}e^{i(\varphi_{21}+\alpha_1)} & r_{22}e^{i(\varphi_{22}+\alpha_2)} & r_{23}e^{i(\varphi_{23}+\alpha_3)} \\ r_{31}e^{i(\varphi_{31}+\alpha_1)} & r_{32}e^{i(\varphi_{32}+\alpha_2)} & r_{33}e^{i(\varphi_{33}+\alpha_3)} \end{pmatrix}. \tag{5.2.51}$$

Let use choose:

$$\alpha_1 = -\varphi_{11}, \qquad \alpha_2 = -\varphi_{12}. \tag{5.2.52}$$

We will leave $\alpha_3$ unspecified for now. Then:

$$B = \begin{pmatrix} r_{11} & r_{12} & r_{13}e^{i(\varphi_{13}+\alpha_3)} \\ r_{21}e^{i(\varphi_{21}-\varphi_{11})} & r_{22}e^{i(\varphi_{22}-\varphi_{12})} & r_{23}e^{i(\varphi_{23}+\alpha_3)} \\ r_{31}e^{i(\varphi_{31}-\varphi_{11})} & r_{32}e^{i(\varphi_{32}-\varphi_{12})} & r_{33}e^{i(\varphi_{33}+\alpha_3)} \end{pmatrix}. \tag{5.2.53}$$

Define then the constants:

$$\delta = -\varphi_{13} - \alpha_3, \qquad \phi_{21} = \varphi_{21} - \varphi_{11}, \qquad \phi_{22} = \varphi_{22} - \varphi_{12},$$
$$\phi_{23} = \varphi_{23} + \alpha_3, \qquad \phi_{31} = \varphi_{31} - \varphi_{11}, \qquad \phi_{32} = \varphi_{32} - \varphi_{12}, \qquad \phi_{33} = \varphi_{33} + \alpha_3. \tag{5.2.54}$$

Then:

$$B = \begin{pmatrix} r_{11} & r_{12} & r_{13}e^{-i\delta} \\ r_{21}e^{i\phi_{21}} & r_{22}e^{i\phi_{22}} & r_{23}e^{i\phi_{23}} \\ r_{31}e^{i\phi_{31}} & r_{32}e^{i\phi_{32}} & r_{33}e^{i\phi_{33}} \end{pmatrix}. \tag{5.2.55}$$

Using that $\phi_{23} = \varphi_{23} + \alpha_3$ we can also see that our freedom to select $\alpha_3$ is equivalent to a freedom to select $\phi_{23}$ arbitrarily. Doing so then fixes values for $\phi_{33}$ and $\delta$. We will return to this. Because $A$ is unitary and $P$ is as well, $B$ is automatically unitary. Therefore, its rows are orthonormal vectors and so:

$$r_{11}^2 + r_{12}^2 + r_{13}^2 = 1. \tag{5.2.56}$$

So, $r_{1a}$ forms the components of a real vector on the sphere in the positive octant. Consequently using spherical coordinates (in a slightly different than normal convention) it can be written for some angles between 0 and $\pi/2$ as:

$$r_{11} = \cos\theta_{12}\cos\theta_{13}, \qquad r_{12} = \sin\theta_{12}\cos\theta_{13}, \qquad r_{13} = \sin\theta_{13}. \tag{5.2.57}$$

With this $B$ becomes:

$$B = \begin{pmatrix} \cos\theta_{12}\cos\theta_{13} & \sin\theta_{12}\cos\theta_{13} & e^{-i\delta}\sin\theta_{13} \\ r_{21}e^{i\phi_{21}} & r_{22}e^{i\phi_{22}} & r_{23}e^{i\phi_{23}} \\ r_{31}e^{i\phi_{31}} & r_{32}e^{i\phi_{32}} & r_{33}e^{i\phi_{33}} \end{pmatrix}. \tag{5.2.58}$$

If we multiply $B$ on the right by a z axis rotation:

$$C \stackrel{\text{def}}{=} BR_z = B\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.2.59}$$

$$\implies C = \begin{pmatrix} \cos(\theta_{12}-\theta)\cos\theta_{13} & \sin(\theta_{12}-\theta)\cos\theta_{13} & e^{-i\delta}\sin\theta_{13} \\ r_{21}e^{i\phi_{21}}\cos\theta + r_{22}e^{i\phi_{22}}\sin\theta & r_{22}e^{i\varphi_{22}}\cos\theta - r_{21}e^{i\varphi_{21}}\sin\theta & r_{23}e^{i\phi_{23}} \\ r_{31}e^{i\phi_{31}}\cos\theta + r_{32}e^{i\phi_{32}}\sin\theta & r_{32}e^{i\phi_{32}}\cos\theta - r_{31}e^{i\phi_{31}}\sin\theta & r_{33}e^{i\phi_{33}} \end{pmatrix}. \tag{5.2.60}$$

Then, choose $\theta = \theta_{12}$:

$$C = \begin{pmatrix} \cos\theta_{13} & 0 & e^{-i\delta}\sin\theta_{13} \\ r_{21}e^{i\phi_{21}}\cos\theta_{12} + r_{22}e^{i\phi_{22}}\sin\theta_{12} & r_{22}e^{i\varphi_{22}}\cos\theta_{12} - r_{21}e^{i\varphi_{21}}\sin\theta_{12} & r_{23}e^{i\phi_{23}} \\ r_{31}e^{i\phi_{31}}\cos\theta_{12} + r_{32}e^{i\phi_{32}}\sin\theta_{12} & r_{32}e^{i\phi_{32}}\cos\theta_{12} - r_{31}e^{i\phi_{31}}\sin\theta_{12} & r_{33}e^{i\phi_{33}} \end{pmatrix}. \tag{5.2.61}$$

Because $B$ is unitary and $R_z$ is as well, $C$ is automatically unitary. Define then the constants $q_{ab} > 0$ and $\zeta_{ab}$ real by:

$$\begin{aligned} q_{21}e^{i\zeta_{21}} &= r_{21}e^{i\phi_{21}}\cos\theta_{12} + r_{22}e^{i\phi_{22}}\sin\theta_{12}, \qquad q_{22}e^{i\zeta_{22}} = r_{22}e^{i\varphi_{22}}\cos\theta_{12} - r_{21}e^{i\varphi_{21}}\sin\theta_{12}, \\ q_{23}e^{i\zeta_{23}} &= r_{23}e^{i\phi_{23}}, \qquad q_{31}e^{i\zeta_{31}} = r_{31}e^{i\phi_{31}}\cos\theta_{12} + r_{32}e^{i\phi_{32}}\sin\theta_{12}, \\ q_{32}e^{i\zeta_{32}} &= r_{32}e^{i\phi_{32}}\cos\theta_{12} - r_{31}e^{i\phi_{31}}\sin\theta_{12}, \qquad q_{33}e^{i\zeta_{33}} = r_{33}e^{i\phi_{33}}. \end{aligned} \tag{5.2.62}$$

Then, $C$ becomes:

$$C = \begin{pmatrix} \cos\theta_{13} & 0 & e^{-i\delta}\sin\theta_{13} \\ q_{21}e^{i\zeta_{21}} & q_{22}e^{i\zeta_{22}} & q_{23}e^{i\zeta_{23}} \\ q_{31}e^{i\zeta_{31}} & q_{32}e^{i\zeta_{32}} & q_{33}e^{i\zeta_{33}} \end{pmatrix}. \tag{5.2.63}$$

Our freedom to select $\alpha_3$ became a freedom to select $\phi_{23} = \zeta_{23}$ and so we still have the ability to select $\zeta_{23}$ arbitrarily. If we multiply $C$ on the right now by a unitary matrix which is like a y axis rotation with a phase adjustment we have:

$$D \stackrel{\text{def}}{=} CR_y = C\begin{pmatrix} \cos\theta & 0 & -e^{-i\delta}\sin\theta \\ 0 & 1 & 0 \\ e^{i\delta}\sin\theta & 0 & \cos\theta \end{pmatrix} \tag{5.2.64}$$

$$\implies D = \begin{pmatrix} \cos(\theta_{13}-\theta) & 0 & \sin(\theta_{13}-\theta) \\ q_{21}e^{i\zeta_{21}}\cos\theta + q_{23}e^{i(\zeta_{23}+\delta)}\sin\theta & q_{22}e^{i\zeta_{22}} & q_{23}e^{i\zeta_{23}}\cos\theta - q_{21}e^{i(\zeta_{21}-\delta)}\sin\theta \\ q_{31}e^{i\zeta_{31}}\cos\theta + q_{33}e^{i(\zeta_{33}+\delta)}\sin\theta & q_{32}e^{i\zeta_{32}} & q_{33}e^{i\zeta_{33}}\cos\theta - q_{31}e^{i(\zeta_{31}-\delta)}\sin\theta \end{pmatrix}. \tag{5.2.65}$$

Then, choose $\theta = \theta_{13}$:

$$D = \begin{pmatrix} 1 & 0 & 0 \\ q_{21}e^{i\zeta_{21}}\cos\theta_{13} + q_{23}e^{i(\zeta_{23}+\delta)}\sin\theta_{13} & q_{22}e^{i\zeta_{22}} & q_{23}e^{i\zeta_{23}}\cos\theta_{13} - q_{21}e^{i(\zeta_{21}-\delta)}\sin\theta_{13} \\ q_{31}e^{i\zeta_{31}}\cos\theta_{13} + q_{33}e^{i(\zeta_{33}+\delta)}\sin\theta_{13} & q_{32}e^{i\zeta_{32}} & q_{33}e^{i\zeta_{33}}\cos\theta_{13} - q_{31}e^{i(\zeta_{31}-\delta)}\sin\theta_{13} \end{pmatrix}. \tag{5.2.66}$$

Because $C$ is unitary and $R_y$ is unitary, $D$ is automatically unitary. This means that the rows of $D$ must be orthogonal to each other. In particular this means that between the first and last rows:

$$\begin{pmatrix} q_{31}e^{i\zeta_{31}}\cos\theta_{13} + q_{33}e^{i(\zeta_{33}+\delta)}\sin\theta_{13} & q_{32}e^{i\zeta_{32}} & q_{33}\cos\theta_{13} - q_{31}e^{i(\zeta_{31}-\delta)}\sin\theta_{13} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 0 \tag{5.2.67}$$

$$\implies q_{31}e^{i\zeta_{31}}\cos\theta_{13} + q_{33}e^{i(\zeta_{33}+\delta)}\sin\theta_{13} = 0 \tag{5.2.68}$$

$$\implies q_{31}e^{i(\zeta_{31}-\delta-\zeta_{33})}\cos\theta_{13} + q_{33}\sin\theta_{13} = 0 \tag{5.2.69}$$

$$\implies q_{31}\sin(\zeta_{31}-\delta-\zeta_{33})\cos\theta_{13} = 0. \tag{5.2.70}$$

Because $\theta_{13}$ is not necessarily $\pi/2$ and $q_{31}$ is not necessarily 0 we must have $\sin(\zeta_{31}-\zeta_{33}-\delta) = 0$. Thus, $e^{i(\zeta_{31}-\zeta_{33}-\delta)}$ must be a real number and so is $\pm 1$. In (5.2.69), because $q_{31}$ and $q_{33}$ are by definition non-negative and $\theta_{13}$ lies in the first quadrant, in order for the two terms to cancel out and give 0 one must

be negative and so $e^{i(\zeta_{31}-\zeta_{33}-\delta)} = -1$. Between the first and second rows we have a very similar situation:

$$\begin{pmatrix} q_{21}e^{i\zeta_{21}}\cos\theta_{13} + q_{23}e^{i(\zeta_{23}+\delta)}\sin\theta_{13} & q_{22}e^{i\zeta_{22}} & q_{23}e^{i\zeta_{23}}\cos\theta_{13} - q_{21}e^{i(\zeta_{21}-\delta)}\sin\theta_{13}\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix} = 0 \tag{5.2.71}$$

$$\implies q_{21}e^{i\zeta_{21}}\cos\theta_{13} + q_{23}e^{i(\zeta_{23}+\delta)}\sin\theta_{13} = 0 \tag{5.2.72}$$

$$\implies q_{21}e^{i(\zeta_{21}-\zeta_{23}-\delta)}\cos\theta_{13} + q_{23}\sin\theta_{13} = 0 \tag{5.2.73}$$

$$\implies q_{21}\sin(\zeta_{21}-\zeta_{23}-\delta)\cos\theta_{13} = 0. \tag{5.2.74}$$

Because $\theta_{13}$ is not necessarily $\pi/2$ and $q_{21}$ is not necessarily 0 we must have $\sin(\zeta_{21}-\zeta_{23}-\delta) = 0$. Thus, $e^{i(\zeta_{21}-\zeta_{23}-\delta)}$ must be a real number and so is $\pm 1$. In (5.2.73), because $q_{21}$ and $q_{23}$ are by definition non-negative and $\theta_{13}$ lies in the first quadrant, in order for the two terms to cancel and give 0 one must be negative and so $e^{i(\zeta_{21}-\zeta_{23}-\delta)} = -1$. Thus, $D$ becomes:

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & q_{22}e^{i\zeta_{22}} & (q_{23}\cos\theta_{13} + q_{21}\sin\theta_{13})e^{i\zeta_{23}} \\ 0 & q_{32}e^{i\zeta_{32}} & (q_{33}\cos\theta_{13} + q_{31}\sin\theta_{13})e^{i\zeta_{33}}\end{pmatrix}. \tag{5.2.75}$$

Then, define the constants $p_{ab} > 0$:

$$p_{22} = q_{22}, \qquad p_{23} = q_{23}\cos\theta_{13} + q_{21}\sin\theta_{13}, \qquad p_{32} = q_{32}, \qquad p_{33} = q_{33}\cos\theta_{13} + q_{31}\sin\theta_{13}. \tag{5.2.76}$$

With this we have:

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & p_{22}e^{i\zeta_{22}} & p_{23}e^{i\zeta_{23}} \\ 0 & p_{32}e^{i\zeta_{32}} & p_{33}e^{i\zeta_{33}}\end{pmatrix}. \tag{5.2.77}$$

Unitary of $D$ then requires of the second row:

$$p_{22}^2 + p_{23}^2 = 1. \tag{5.2.78}$$

Hence, there exists a $\theta_{23}$ between 0 and $\pi/2$ so that:

$$p_{22} = \cos\theta_{23}, \qquad p_{23} = \sin\theta_{23}. \tag{5.2.79}$$

Then, orthogonality with rows two and three requires:

$$\begin{pmatrix} p_{32}e^{i\zeta_{32}} & p_{33}e^{i\zeta_{33}}\end{pmatrix}\begin{pmatrix} e^{-i\zeta_{22}}\cos\theta_{23} \\ e^{-i\zeta_{23}}\sin\theta_{23}\end{pmatrix} = 0 \tag{5.2.80}$$

$$\implies p_{32}e^{i(\zeta_{32}-\zeta_{22})}\cos\theta_{23} + p_{33}e^{i(\zeta_{33}-\zeta_{23})}\sin\theta_{23} = 0 \tag{5.2.81}$$

$$\implies p_{32}e^{i(\zeta_{32}+\zeta_{23}-\zeta_{22}-\zeta_{33})}\cos\theta_{23} + p_{33}\sin\theta_{23} = 0 \tag{5.2.82}$$

$$\implies p_{32}\sin(\zeta_{32}+\zeta_{23}-\zeta_{22}-\zeta_{33})\cos\theta_{23} = 0. \tag{5.2.83}$$

Because $p_{32}$ is not necessarily 0 and $\theta_{23}$ is not necessarily $\pi/2$ we must then have $\sin(\zeta_{32}+\zeta_{23}-\zeta_{22}-\zeta_{33}) = 0$ and so $e^{i(\zeta_{23}+\zeta_{32}-\zeta_{22}-\zeta_{33})}$ is real and thus $\pm 1$. In order for these two terms to cancel because all other factors are positive, we must have $e^{i(\zeta_{23}+\zeta_{32}-\zeta_{22}-\zeta 33)} = -1$. Then:

$$p_{32}\cos\theta_{23} = p_{33}\sin\theta_{23}. \tag{5.2.84}$$

Unitary of $D$ then requires of the third row:

$$p_{32}^2 + p_{33}^2 = 1 \tag{5.2.85}$$

$$\implies p_{33}^2\left(1 + \frac{\sin^2\theta_{23}}{\cos^2\theta_{23}}\right) = 1 \tag{5.2.86}$$

$$\implies p_{33} = \cos\theta_{23}, \qquad p_{32} = \sin\theta_{23}. \tag{5.2.87}$$

Then, $D$ becomes:

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\zeta_{22}}\cos\theta_{23} & e^{i\zeta_{23}}\sin\theta_{23} \\ 0 & -e^{i(\zeta_{22}+\zeta_{33}-\zeta_{23})}\sin\theta_{23} & e^{i\zeta_{33}}\cos\theta_{23}\end{pmatrix}. \tag{5.2.88}$$

If we now multiply $D$ on the right by an x axis rotation we have:

$$E \stackrel{\text{def}}{=} DR_x = D \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{5.2.89}$$

$$\implies E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\zeta_{22}}\cos\theta\cos\theta_{23} + e^{i\zeta_{23}}\sin\theta\sin\theta_{23} & e^{i\zeta_{23}}\sin\theta_{23}\cos\theta - e^{i\zeta_{22}}\cos\theta_{23}\sin\theta \\ 0 & e^{i\zeta_{33}}\sin\theta\cos\theta_{23} - e^{i(\zeta_{22}+\zeta_{33}-\zeta_23)}\cos\theta\sin\theta_{23} & e^{\zeta_{33}}\cos\theta_{23}\cos\theta + e^{i(\zeta_{22}+\zeta_{33}-\zeta_{23})}\sin\theta_{23}\sin\theta \end{pmatrix}. \tag{5.2.90}$$

Now let us finally use our freedom to select $\alpha_3$ and thus set $\zeta_{23}$ to impose $\zeta_{23} = \zeta_{22}$ by definition. Then, now that we have selected $\alpha_3$:

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\zeta_{22}}\cos(\theta_{23}-\theta) & e^{i\zeta_{22}}\sin(\theta_{23}-\theta) \\ 0 & e^{i\zeta_{33}}\sin(\theta-\theta_{23}) & e^{\zeta_{33}}\cos(\theta_{23}-\theta) \end{pmatrix}. \tag{5.2.91}$$

If we select then $\theta = \theta_{23}$ we have that $E$ becomes:

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\zeta_{22}} & 0 \\ 0 & 0 & e^{i\zeta_{33}} \end{pmatrix}. \tag{5.2.92}$$

This is just a diagonal phase matrix. Recalling the definition of $E$:

$$E = DR_x = CR_yR_x = BR_zR_yR_x = APR_zR_yR_x \tag{5.2.93}$$

$$\implies A = ER_x^\dagger R_y^\dagger R_z^\dagger P^\dagger. \tag{5.2.94}$$

Therefore, any $3\times 3$ unitary matrix $A$ for some parameters $\alpha_1$, $\alpha_2$, $\alpha_3$, $\theta_{12}$, $\theta_{13}$, $\theta_{23}$, $\zeta_2$, $\zeta_3$ can be written as:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\zeta_2} & 0 \\ 0 & 0 & e^{i\zeta_3} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} \begin{pmatrix} \cos\theta_{13} & 0 & e^{-i\delta}\sin\theta_{13} \\ 0 & 1 & 0 \\ -e^{i\delta}\sin\theta_{13} & 0 & \cos\theta_{13} \end{pmatrix} \cdot \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{-i\alpha_1} & 0 & 0 \\ 0 & e^{-i\alpha_2} & 0 \\ 0 & 0 & e^{-i\alpha_3} \end{pmatrix}. \tag{5.2.95}$$

One may think there is another phase degree of freedom we could use to transform the 11 component of the first matrix above. However, an overall phase of the first phase matrix can be canceled by a corresponding overall phase in the second phase matrix all the way to the right because $e^{i\theta}\mathbb{1}$ commutes with everything. Thus, without loss of generality we can set the 11 component of the first phase matrix above to be 1. Returning to (5.2.49), we can use the phase matrices $P_e, P_\nu, P_u, P_d$ to cancel the left and right phase matrices which would appear here in this decomposition for $U_\nu$ and $U_d$. Having set those phase matrices so that they are canceled, we have that $U_\nu$ and $U_d$ can be written as:

$$U_\nu = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23\nu} & \sin\theta_{23\nu} \\ 0 & -\sin\theta_{23\nu} & \cos\theta_{23\nu} \end{pmatrix} \begin{pmatrix} \cos\theta_{13\nu} & 0 & e^{-i\delta_\nu}\sin\theta_{13\nu} \\ 0 & 1 & 0 \\ -e^{i\delta_\nu}\sin\theta_{13\nu} & 0 & \cos\theta_{13\nu} \end{pmatrix} \begin{pmatrix} \cos\theta_{12\nu} & \sin\theta_{12\nu} & 0 \\ -\sin\theta_{12\nu} & \cos\theta_{12\nu} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.2.96}$$

$$U_d = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23d} & \sin\theta_{23d} \\ 0 & -\sin\theta_{23d} & \cos\theta_{23d} \end{pmatrix} \begin{pmatrix} \cos\theta_{13d} & 0 & e^{-i\delta_d}\sin\theta_{13d} \\ 0 & 1 & 0 \\ -e^{i\delta_d}\sin\theta_{13d} & 0 & \cos\theta_{13d} \end{pmatrix} \begin{pmatrix} \cos\theta_{12d} & \sin\theta_{12d} & 0 \\ -\sin\theta_{12d} & \cos\theta_{12d} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.2.97}$$

This matrix $U_\nu$ is called the Pontecorvo–Maki–Nakagawa–Sakata matrix or PMNS matrix or neutrino mixing matrix. The $U_d$ matrix is called the Cabibbo–Kobayashi–Maskawa matrix or CKM matrix or quark mixing matrix. Thus, after fixing our field definitions we have the Lagrangian density for our theory as given in (1.9.1) with the Yukawa matrices given as in (5.2.47). Each of the $D$ matrices are diagonal with non-negative eigenvalues and so have 3 parameters each. The PMNS and CKM matrices we see have 4 parameters each: 3 mixing angles and one phase. Thus we have 20 parameters among the four Yukawa matrices. As well, we in principle have 6 complex or equivalently 12 real parameters necessary to fix the Majorana matrix $M$. So, allowing for 3 right handed neutrinos and Majorana masses, we have 32 parameters to fix.

The presence of the Majorana mass terms breaks an otherwise present symmetry of the theory. Notice that every term involving a quark field in (1.9.1) also involves a corresponding complex conjugate of a

quark field. Thus, under global U(1) transformations of all the quarks together that Lagrangian density has symmetry. That is, (1.9.1) has symmetry under the transformation:

$$\mathfrak{q}' = e^{\frac{i}{3}\theta}\mathfrak{q}. \tag{5.2.98}$$

The factor of 3 here is conventional. The corresponding conserved quantity to this symmetry from Noether's theorem is called the "Baryon number". As well, notice that every time involving a lepton field in (1.9.1) also involves a corresponding complex conjugate of a lepton field except for the Majorana mass term. Thus, excluding that term under global U(1) transformations of all the leptons together that Lagrangian density has symmetry. That is, (1.9.1) has symmetry under the transformation:

$$\ell' = e^{i\theta}\ell \tag{5.2.99}$$

except for the Majorana mass term. The corresponding conserved quantity when no Majorana mass term is included is called the "Lepton number". So, should we believe in lepton number conservation? Experimentally so far lepton conservation has been observed and no right handed neutrinos have been observed but their effects are quite small on the other fields here due to the very limited ways in which they are coupled. So, at the moment it is an unsolved problem whether lepton number conservation is a true symmetry of our universe. There are many models with it and many models without it. There are also many models which break both baryon number conservation and lepton number conservation but still conserve the difference of baryon number and lepton number. At the moment the correct path forward is not known to anyone. For simplicity, we will assume lepton number conservation is a true symmetry of our universe in which case no Majorana mass term is allowed in the Lagrangian density in (1.9.1) and we must take $M = 0$. There are many good reasons not to do this! We are only making this assumption so that we can write down a model consistent with the experimental data available at this time with as few parameters as possible. It is certainly not the only model consistent with the data and there are arguments why $M \neq 0$ may provide nice results down the line; thus results just at this time are inaccessible to experiment. Without any Majorana mass term we have then:

$$\begin{aligned}\mathfrak{L} =& -\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi - \frac{m_H^2}{2\mathrm{v}^2}\left(\Phi^\dagger\Phi - \frac{\mathrm{v}^2}{2}\right)^2 - \frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}) - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_S^{\mu\nu}\mathcal{F}_{S\mu\nu})\\ &+ i\bar{\ell}_a\gamma^\mu\mathscr{D}_\mu\ell_a - 2\mathrm{Re}(\Gamma_e^{ab}e_{Ra}^\dagger\Phi^\dagger\ell_{Lb}) - 2\mathrm{Im}(\Gamma_\nu^{ab}\nu_{Ra}^\dagger\Phi^T\sigma_2\ell_{Lb})\\ &+ i\bar{\mathfrak{q}}_a\gamma^\mu\mathscr{D}_\mu\mathfrak{q}_a - 2\mathrm{Re}(\Gamma_d^{ab}d_{Ra}^\dagger\Phi\mathfrak{q}_{Lb}) - 2\mathrm{Im}(\Gamma_u^{ab}u_{Ra}^\dagger\Phi^T\sigma_2\mathfrak{q}_{Lb}).\end{aligned} \tag{5.2.100}$$

This theory then has 20 parameters to be fixed among the different Yukawa matrices. Let us just look at the fermion dependent parts of this Lagrangian density for a moment:

$$\begin{aligned}\mathfrak{L}_f =& i\bar{\ell}_a\gamma^\mu\mathscr{D}_\mu\ell_a - 2\mathrm{Re}(\Gamma_e^{ab}e_{Ra}^\dagger\Phi^\dagger\ell_{Lb}) - 2\mathrm{Im}(\Gamma_\nu^{ab}\nu_{Ra}^\dagger\Phi^T\sigma_2\ell_{Lb})\\ &+ i\bar{\mathfrak{q}}_a\gamma^\mu\mathscr{D}_\mu\mathfrak{q}_a - 2\mathrm{Re}(\Gamma_d^{ab}d_{Ra}^\dagger\Phi\mathfrak{q}_{Lb}) - 2\mathrm{Im}(\Gamma_u^{ab}u_{Ra}^\dagger\Phi^T\sigma_2\mathfrak{q}_{Lb}).\end{aligned} \tag{5.2.101}$$

If we apply the spontaneous symmetry breaking and only work to lowest order in the fields so that we can study this Lagrangian density for weak fields we have:

$$\begin{aligned}\mathfrak{L}_f =& i\bar{\ell}_a\gamma^\mu\partial_\mu\ell_a - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_e^{ab}e_{Ra}^\dagger e_{Lb} - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_e^{ab*}e_{Lb}^\dagger e_{Ra} - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_\nu^{ab}\nu_{Ra}^\dagger\nu_{Lb} - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_\nu^{ab*}\nu_{Lb}^\dagger\nu_{Ra}\\ &+ i\bar{\mathfrak{q}}_a\gamma^\mu\partial_\mu\mathfrak{q}_a - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_d^{ab}d_{Ra}^\dagger d_{Lb} - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_d^{ab*}d_{Lb}^\dagger d_{Ra} - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_u^{ab}u_{Ra}^\dagger u_{Lb} - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_u^{ab*}u_{Lb}^\dagger u_{Ra} + \text{H.O.T.}.\end{aligned} \tag{5.2.102}$$

From (5.2.47) we have made all of the Yukawa matrices self-adjoint. Then, defining:

$$e_a = \begin{pmatrix} e_{La} \\ e_{Ra}\end{pmatrix}, \qquad \nu_a = \begin{pmatrix} \nu_{La} \\ \nu_{Ra}\end{pmatrix}, \qquad d_a = \begin{pmatrix} d_{La} \\ d_{Ra}\end{pmatrix}, \qquad u_a = \begin{pmatrix} u_{La} \\ u_{Ra}\end{pmatrix} \tag{5.2.103}$$

this can all be written as:

$$\begin{aligned}\mathfrak{L}_f =& i\bar{e}_a\gamma^\mu\partial_\mu e_a + i\bar{\nu}_a\gamma^\mu\partial_\mu\nu_a - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_e^{ab}\bar{e}_a e_b - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_\nu^{ab}\bar{\nu}_a\nu_b\\ &+ i\bar{d}_a\gamma^\mu\partial_\mu d_a + i\bar{u}_a\gamma^\mu\partial_\mu u_a - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_d^{ab}\bar{d}_a d_b - \frac{\mathrm{v}}{\sqrt{2}}\Gamma_u^{ab}\bar{u}_a u_b + \text{H.O.T.}.\end{aligned} \tag{5.2.104}$$

Using that we have diagonalized $\Gamma_e$, we see that at lowest order $e_1$, $e_2$, $e_3$ each individually satisfy the Dirac equation but with different masses. Recall that we ordered the eigenvalues of $\Gamma_e$ from least to greatest. Then, $e_1$ is called the electron field, $e_2$ is called the muon field, and $e_3$ is called the tauon field and we can

then see that their masses are given by:

$$m_e = \frac{\mathtt{v}}{\sqrt{2}} D_e^{11}, \qquad m_\mu = \frac{\mathtt{v}}{\sqrt{2}} D_e^{22}, \qquad m_\tau = \frac{\mathtt{v}}{\sqrt{2}} D_e^{33}. \tag{5.2.105}$$

So, we define the electron mass matrix by:

$$M_e = \begin{pmatrix} m_e & 0 & 0 \\ 0 & m_\mu & 0 \\ 0 & 0 & \mu_\tau \end{pmatrix} \implies \Gamma_e = \frac{\sqrt{2}}{\mathtt{v}} M_e. \tag{5.2.106}$$

Similarly, using that we have diagonalized $\Gamma_u$, we see that at lowest order $u_1$, $u_2$, $u_3$ each individually satisfy the Dirac equation but with different masses. We have done the same ordering of eigenvalues of $\Gamma_u$. Then, $u_1$ is called the up quark field, $u_2$ is called the charm quark field, and $u_3$ is called the top quark field and we can see then that their masses are given by:

$$m_u = \frac{\mathtt{v}}{\sqrt{2}} D_u^{11}, \qquad m_c = \frac{\mathtt{v}}{\sqrt{2}} D_u^{22}, \qquad m_t = \frac{\mathtt{v}}{\sqrt{2}} D_u^{33}. \tag{5.2.107}$$

So, we define the up quark mass matrix by:

$$M_u = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix} \implies \Gamma_u = \frac{\sqrt{2}}{\mathtt{v}} M_u. \tag{5.2.108}$$

We have not diagonalized $\Gamma_\nu$ or $\Gamma_d$. Instead, define "mass eigenstate fields":

$$\nu_{Ma} = (U_\nu^\dagger)_a{}^b \nu_b, \qquad d_{Ma} = (U_d^\dagger)_a{}^b d_b \tag{5.2.109}$$

$$\implies \Gamma_\nu^{ab} \bar{\nu}_a \nu_b = D_\nu^{ab} \bar{\nu}_{Ma} \nu_{Mb}, \qquad \Gamma_d^{ab} \bar{d}_a d_b = D_d^{ab} \bar{d}_{Ma} d_{Mb}. \tag{5.2.110}$$

Then, because $U_\nu$ and $U_d$ are unitary:

$$i\bar{\nu}_a \gamma^\mu \partial_\mu \nu_a + i\bar{d}_a \gamma^\mu \partial_\mu d_a = i\bar{\nu}_{Ma} \gamma^\mu \partial_\mu \nu_{Ma} + i\bar{d}_{Ma} \gamma^\mu \partial_\mu d_{Ma}. \tag{5.2.111}$$

So, rewriting $\mathfrak{L}_f$ in terms of the mass eigenstate fields we can see that $\nu_{Ma}$ and $d_{Ma}$ for each $a$ satisfy the Dirac equation with masses:

$$m_{\nu a} = \frac{\mathtt{v}}{\sqrt{2}} D_\nu^{aa}, \qquad m_{da} = \frac{\mathtt{v}}{\sqrt{2}} D_d^{aa}. \tag{5.2.112}$$

So, we define the neutrino and down quark mass matrices by:

$$M_\nu = \begin{pmatrix} m_{\nu l} & 0 & 0 \\ 0 & m_{\nu m} & 0 \\ 0 & 0 & m_{\nu h} \end{pmatrix}, \qquad M_d = \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix} \tag{5.2.113}$$

$$\implies D_\nu = \frac{\sqrt{2}}{\mathtt{v}} M_\nu, \qquad D_d = \frac{\sqrt{2}}{\mathtt{v}} M_d. \tag{5.2.114}$$

Then, we can see that the CKM and PMNS matrices have the interpretations as the change of basis matrices which relate the mass eigenstates of the down-like quarks or neutrinos to the eigenstates of the electroweak interaction (called the flavor eigenstates) of the down quarks or neutrinos. The mass eigenstates are the fields which propagate freely in the limit of no interactions but the weak eigenstates (the conventional $\nu_a$ and $d_a$ fields we have been using) are the fields which transform under SU(2) weak transformations with the electrons and up quarks. It is a common abuse of language to call both the flavor eigenstates and the mass eigenstates of the down-like quarks the down quark, strange quark, and bottom quark (in order of lightest to heaviest for the mass eigenstates or in order of lightest to heaviest up-like quarks associated with them in the SU(2) doublet for the flavor eigenstates). This is because the mixing angles for the CKM matrix are small enough for many purposes to approximately identify the bases with each other. However, it should be made clear that the flavor eigenstates do not have well defined masses; only the mass eigenstates do. So, the physical down quark is defined as the lightest mass eigenstate of the down-like quarks, the strange quark is defined as the middle mass eigenstate, and the bottom quark is defined as the heaviest mass eigenstate. The mixing angles for the neutrinos turn out to be large and so no even approximate identification can be made between the mass eigenstates and the flavor eigenstates. The flavor eigenstates $\nu_1 = \nu_e$, $\nu_2 = \nu_\mu$, $\nu_3 = \nu_\tau$ are called the electron neutrino, mu neutrino, and tau neutrino. The mass eigenstates don't have a consistently established name but we will call them the light neutrino, middle neutrino, and heavy neutrino.

With these identifications, we can rewrite our full Lagrangian density as:

$$\begin{aligned}\mathfrak{L} =& -\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi - \frac{m_H^2}{2\mathrm{v}^2}\left(\Phi^\dagger\Phi - \frac{\mathrm{v}^2}{2}\right)^2 - \frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}) - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_S^{\mu\nu}\mathcal{F}_{S\mu\nu})\\ &+ i\bar{\ell}_a\gamma^\mu\mathscr{D}_\mu\ell_a - \frac{2\sqrt{2}}{\mathrm{v}}\mathrm{Re}(M_e^{ab}e_{Ra}^\dagger\Phi^\dagger\ell_{Lb}) - \frac{2\sqrt{2}}{\mathrm{v}}\mathrm{Im}((U_\nu M_\nu U_\nu^\dagger)^{ab}\nu_{Ra}^\dagger\Phi^T\sigma_2\ell_{Lb})\\ &+ i\bar{\mathfrak{q}}_a\gamma^\mu\mathscr{D}_\mu\mathfrak{q}_a - \frac{2\sqrt{2}}{\mathrm{v}}\mathrm{Re}((U_dM_dU_d^\dagger)^{ab}d_{Ra}^\dagger\Phi\mathfrak{q}_{Lb}) - \frac{2\sqrt{2}}{\mathrm{v}}\mathrm{Im}(M_u^{ab}u_{Ra}^\dagger\Phi^T\sigma_2\mathfrak{q}_{Lb}).\end{aligned} \tag{5.2.115}$$

This Lagrangian density is extremely successful in describing experiment but has a significant number of puzzling characteristics we will need to discuss.

## 5.3 Standard Model in full glory

We have now arrived at the Lagrangian density:

$$\begin{aligned}\mathfrak{L} =& -\mathscr{D}^\mu\Phi^\dagger\mathscr{D}_\mu\Phi - \frac{m_H^2}{2\mathrm{v}^2}\left(\Phi^\dagger\Phi - \frac{\mathrm{v}^2}{2}\right)^2 - \frac{1}{4}\mathcal{F}_Y^{\mu\nu}\mathcal{F}_{Y\mu\nu} - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_W^{\mu\nu}\mathcal{F}_{W\mu\nu}) - \frac{1}{2}\mathrm{Tr}(\mathcal{F}_S^{\mu\nu}\mathcal{F}_{S\mu\nu})\\ &+ i\bar{\ell}_a\gamma^\mu\mathscr{D}_\mu\ell_a - 2\mathrm{Re}(\Gamma_e^{ab}e_{Ra}^\dagger\Phi^\dagger\ell_{Lb}) - 2\mathrm{Im}(\Gamma_\nu^{ab}\nu_{Ra}^\dagger\Phi^T\sigma_2\ell_{Lb})\\ &+ i\bar{\mathfrak{q}}_a\gamma^\mu\mathscr{D}_\mu\mathfrak{q}_a - 2\mathrm{Re}(\Gamma_d^{ab}d_{Ra}^\dagger\Phi\mathfrak{q}_{Lb}) - 2\mathrm{Im}(\Gamma_u^{ab}u_{Ra}^\dagger\Phi^T\sigma_2\mathfrak{q}_{Lb}).\end{aligned} \tag{5.3.1}$$

This is known as the (Mostly) Standard Model Lagrangian density. The Standard Model is precisely this Lagrangian density with the right handed neutrinos set to 0. Then, the left handed neutrinos are automatically massless which is experimentally found to be false. This Lagrangian density is the simplest modification of the Standard Model Lagrangian density which allows for massive neutrinos and it is currently compatible with experiment. I will likely call this Lagrangian density the Standard Model even though it is slightly different. In this theory we have a single spin 0 field called the Higgs $\Phi$. We have twelve spin 1/2 fields coming in three copies of sets of four: the electron-like leptons, the neutrinos, the up-like quarks, and the down-like quarks. We then have three spin 1/2 gauge fields: the hypercharge gauge field, the weak gauge field, and the strong gauge field (called the gluon field). The Standard Model has local symmetry under $\mathrm{U_Y}(1)$, weak SU(2), and strong SU(3) transformations. The representations of $\mathrm{U_Y}(1)$ are specified by hypercharge $Y$, the representations of weak SU(2) are specified by the size of the representation $2I_W + 1$ where $I_W$ takes half integer values and is called the isospin rep, and the representations of strong SU(3) which we use are specified by their number of components (1 for the trivial representation, 3 for the fundamental representation, and 8 for the adjoint representation) called the color rep. We also have global U(1) baryon number symmetry among the quarks and global U(1) lepton number symmetry among the leptons. Then, the field content of the Standard Model is given by:

| Field Content before Spontaneous Symmetry Breaking | | | | | | | |
|---|---|---|---|---|---|---|---|
| Field | Name | Spin | Hypercharge | Isospin Rep | Color Rep | Baryon # | Lepton # |
| $\Phi$ | Higgs | 0 | 0 | 1/2 | 1 | 0 | 0 |
| $\ell_{La}$ | LH Lepton | 1/2 | -1 | 1/2 | 1 | 0 | 1 |
| $\nu_{Ra}$ | RH Neutrino | 1/2 | 0 | 0 | 1 | 0 | 1 |
| $e_{Ra}$ | RH Electron-like Lepton | 1/2 | -2 | 0 | 1 | 0 | 1 |
| $\mathfrak{q}_{La}$ | LH Quark | 1/2 | 1/3 | 1/2 | 3 | 1/3 | 0 |
| $u_{Ra}$ | RH Up-like Quark | 1/2 | 4/3 | 0 | 3 | 1/3 | 0 |
| $d_{Ra}$ | RH Down-like Quark | 1/2 | $-2/3$ | 0 | 3 | 1/3 | 0 |
| $B_\mu$ | Hypercharge Gauge Field | 1 | 0 | 0 | 1 | 0 | 0 |
| $W_\mu^a$ | Weak Gauge Field | 1 | 0 | 1 | 1 | 0 | 0 |
| $G_\mu^a$ | Strong Gauge Field | 1 | 0 | 0 | 8 | 0 | 0 |

Here LH and RH are for left handed and right handed. As well, the subscript $a$ on leptons is to indicate that $a$ can be $1, 2$, or 3 due to the 3 generations of matter. The covariant derivative of any field $\psi$ other than the gauge fields is given by:

$$\mathscr{D}_\mu\psi = \partial_\mu\psi - ig_Y Y B_\mu\psi - ig_W\tau_{Wa}W_\mu^a\psi - ig_S\tau_{Sa}G_\mu^a\psi \tag{5.3.2}$$

where the generators $\tau_{Wa}$ and $\tau_{Sa}$ are the weak and strong generators corresponding to the isospin and color representations of $\psi$. The hypercharge coupling $g_Y$ and weak coupling $g_W$ are related to the fundamental charge $\mathsf{e}$ and Weinberg angle $\theta_W$ by:

$$g_Y = \frac{\mathsf{e}}{2\cos\theta_W}, \qquad g_W = \frac{\mathsf{e}}{\sin\theta_W}. \tag{5.3.3}$$

After spontaneous symmetry breaking we switch to the reduced higgs $\phi$, the $W$ boson, the $Z$ boson, and the $A$ electromagnetic vector potential instead of the Higgs and Hypercharge and Weak gauge fields. They are related by:

$$\Phi = \begin{pmatrix} 0 \\ \frac{\mathsf{v}+\phi}{\sqrt{2}} \end{pmatrix}, \tag{5.3.4}$$

$$W_\mu = \frac{W_\mu^1 - iW_\mu^2}{\sqrt{2}}, \tag{5.3.5}$$

$$Z_\mu = W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W \tag{5.3.6}$$

$$A_\mu = B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W. \tag{5.3.7}$$

After spontaneous symmetry breaking then we have field content:

| Field Content after Spontaneous Symmetry Breaking | | | | | | |
|---|---|---|---|---|---|---|
| Field | Name | Spin | Electric Charge | Color Rep | Baryon # | Lepton # |
| $\phi$ | (reduced) higgs | 0 | 0 | 1 | 0 | 0 |
| $\nu_{La}$ | LH Neutrino | 1/2 | 0 | 1 | 0 | 1 |
| $e_{La}$ | LH Electron-like Lepton | 1/2 | -1 | 1 | 0 | 1 |
| $\nu_{Ra}$ | RH Neutrino | 1/2 | 0 | 1 | 0 | 1 |
| $e_{Ra}$ | RH Electron-like Lepton | 1/2 | -1 | 1 | 0 | 1 |
| $u_{La}$ | LH Up-like Quark | 1/2 | 2/3 | 3 | 1/3 | 0 |
| $d_{La}$ | LH Down-like Quark | 1/2 | $-1/3$ | 3 | 1/3 | 0 |
| $u_{Ra}$ | RH Up-like Quark | 1/2 | 2/3 | 3 | 1/3 | 0 |
| $d_{Ra}$ | RH Down-like Quark | 1/2 | $-1/3$ | 3 | 1/3 | 0 |
| $A_\mu$ | Electromagnetic Gauge Field | 1 | 0 | 1 | 0 | 0 |
| $Z_\mu$ | Z Boson | 1 | 0 | 1 | 0 | 0 |
| $W_\mu$ | W Boson | 1 | 1 | 1 | 0 | 0 |
| $G_\mu^a$ | Strong Gauge Field | 1 | 0 | 8 | 0 | 0 |

The Yukawa matrices are given by:

$$\Gamma_e = \frac{\mathsf{v}}{\sqrt{2}} \begin{pmatrix} m_e & 0 & 0 \\ 0 & m_\mu & 0 \\ 0 & 0 & m_\tau \end{pmatrix} \tag{5.3.8}$$

$$\Gamma_\nu = \frac{\mathsf{v}}{\sqrt{2}} U_\nu \begin{pmatrix} m_{\nu 1} & 0 & 0 \\ 0 & m_{\nu 2} & 0 \\ 0 & 0 & m_{\nu 3} \end{pmatrix} U_\nu^\dagger \tag{5.3.9}$$

$$\Gamma_d = \frac{\mathsf{v}}{\sqrt{2}} U_d \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix} U_d^\dagger \tag{5.3.10}$$

$$\Gamma_u = \frac{\mathsf{v}}{\sqrt{2}} \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix}. \tag{5.3.11}$$

where the neutrino mixing matrix (PMNS) and the down quark mixing matrix (CKM) are given by:

$$U_\nu = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23\nu} & \sin\theta_{23\nu} \\ 0 & -\sin\theta_{23\nu} & \cos\theta_{23\nu} \end{pmatrix} \begin{pmatrix} \cos\theta_{13\nu} & 0 & e^{-i\delta_\nu}\sin\theta_{13\nu} \\ 0 & 1 & 0 \\ -e^{i\delta_\nu}\sin\theta_{13\nu} & 0 & \cos\theta_{13\nu} \end{pmatrix} \begin{pmatrix} \cos\theta_{12\nu} & \sin\theta_{12\nu} & 0 \\ -\sin\theta_{12\nu} & \cos\theta_{12\nu} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.3.12}$$

$$U_d = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23d} & \sin\theta_{23d} \\ 0 & -\sin\theta_{23d} & \cos\theta_{23d} \end{pmatrix} \begin{pmatrix} \cos\theta_{13d} & 0 & e^{-i\delta_d}\sin\theta_{13d} \\ 0 & 1 & 0 \\ -e^{i\delta_d}\sin\theta_{13d} & 0 & \cos\theta_{13d} \end{pmatrix} \begin{pmatrix} \cos\theta_{12d} & \sin\theta_{12d} & 0 \\ -\sin\theta_{12d} & \cos\theta_{12d} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.3.13}$$

Thus, the (Mostly) Standard Model has 25 parameters to be determined by experiment. In Natural Units, there is only one independent unit, a unit of time or equivalently a unit of length or a unit of mass. Conventionally for particle physics, the standard unit is the electron-Volt eV which is the energy equal to the work necessary to push a charge with charge magnitude equal to that of the electron through an emf of 1 Volt. Of course, then one must define the Volt. The historical definition of the Volt is a potential difference of 1 Joule per Coulomb where a Joule is the energy unit $\text{kgm}^2/\text{s}^2$ and the Coulomb is the historical SI unit of arbitrary reference charge. However, SI units are no longer defined in terms of historical reference values but instead defined so that $\hbar$, $c$, $\mathtt{e}$ take exact values in SI units (this implicitly defines the reference units, just like we define $\hbar = c = \epsilon_0 = 1$ which implicitly defines our units). With this, the electron-Volt is defined exactly in terms of the second by the formula:

$$1 \text{ eV} = 2\pi \frac{1.602176634}{6.62607015} 10^{15} \text{ s}^{-1} \approx 1.519267447878626201 \times 10^{15} \text{ s}^{-1}. \tag{5.3.14}$$

This formula uses the SI defined values of the electron charge magnitude and $\hbar$. The second is defined exactly as the duration of 9192631770 periods of the electromagnetic radiation emitted by the transition between the hyperfine levels of the unperturbed ground state of the $^{133}$Cs atom. It is common to use keV (kiloelectron-Volts) for 1000 eV, MeV (megaelectron-Volt) for $10^6$ eV, Gev (gigaelectron-Volts) for $10^9$ eV, and TeV (teraelectron-Volts) for $10^{12}$.

| Parameters of the (Mostly) Standard Model [37] | | |
|---|---|---|
| Symbol | Name | Value |
| $\mathtt{e}$ | Electromagnetic Coupling Constant | .302822120872(23) |
| $g_S$ | Strong Coupling Constant | 1.220(3) |
| $m_W$ | W Boson Mass | 80.379(12) GeV |
| $m_Z$ | Z Boson Mass | 91.1876(21) GeV |
| $m_H$ | Higgs Mass | 125.25(17) GeV |
| $m_e$ | Electron Mass | 510.9989461(31) keV |
| $m_\mu$ | Muon Mass | 105.6583745(24) MeV |
| $m_\tau$ | Tauon Mass | 1.77686(12) GeV |
| $m_{\nu l}$ | Light Neutrino Mass | $< 1.1$ eV |
| $m_{\nu m}$ | Middle Neutrino Mass | $< 1.1$ eV |
| $m_{\nu h}$ | Heavy Neutrino Mass | $< 1.1$ eV |
| $m_u$ | Up Quark Mass | $2.16^{+.49}_{-.26}$ MeV |
| $m_c$ | Charm Quark Mass | 1.27(2) GeV |
| $m_t$ | Top Quark Mass | 172.76(30) GeV |
| $m_d$ | Down Quark Mass | $4.67^{+.48}_{-.17}$ MeV |
| $m_s$ | Strange Quark Mass | $93^{+11}_{-5}$ MeV |
| $m_b$ | Bottom Quark Mass | $4.18^{+.03}_{-.02}$ GeV |
| $\theta_{12\nu}$ | PMNS Light-Middle Mixing Angle | 0.587(14) |
| $\theta_{13\nu}$ | PMNS Light-Heavy Mixing Angle | 0.1488(23) |
| $\theta_{23\nu}$ | PMNS Middle-Heavy Mixing Angle | 0.827(22) |
| $\delta_\nu$ | PMNS Phase | 4.27(56) |
| $\theta_{12d}$ | CKM Down-Strange Mixing Angle | 0.22848(49) |
| $\theta_{13d}$ | CKM Down-Bottom Mixing Angle | 0.00361(10) |
| $\theta_{23d}$ | CKM Strange-Bottom Mixing Angle | 0.04054(72) |
| $\delta_d$ | CKM Phase | 1.196(44) |

The neutrino masses have never been measured directly however magnitude of differences of squared neutrino masses have been observed:

$$|m_{\nu 2}^2 - m_{\nu 1}^2| = 7.53(18) \times 10^{-5} \text{ eV}^2 \tag{5.3.15}$$

$$|m_{\nu 3}^2 - m_{\nu 2}^2| = 2.453(33) \times 10^{-3} \text{ eV}^2. \tag{5.3.16}$$

From these parameters it is possible to calculate the Higgs vev, the Weinberg angle, the hypercharge coupling constant $g_Y$, and the weak coupling constant $g_W$ using the formulas we have established. The parameters listed in this table are the ones more directly accessible in experiment. This model, the Standard Model is the most successful model of the universe in the entire history of humanity. It is able to consistently successfully predict the results of particle physics experiments across a wide spectrum of energies and the

effective theories produced from it are able to predict the structure of atoms and molecules, the material properties of everyday materials, and reduce under appropriate limits to the macroscopically observed physics with which we are all familiar. The Standard Model, first formulated in full around 1974, has passed thousands of different experimental tests and it has proven extremely difficult for physicists to find anyway to modify it to make a successful new prediction without breaking something else. As of 2023, the (Mostly) Standard Model Lagrangian Density is the status quo of our understanding of fundamental physics (when treated with quantum field theory). However, it is known to have some issues.

## 5.4 Standard Model: Hard and Soft Problems

The Standard Model does possess several noteworthy problems. They come in two forms: "hard" problems and "soft" problems. A hard problem is a genuine discrepancy between the model and experimental reality. An exactly true model cannot have any hard problems. A soft problem is a feature of a model that "seems" for a potentially good reason to be peculiar or hard to believe. For example, a term which is allowed by symmetries just turning out to be 0 anyway is quite surprising as it could have been any non-zero value and 0 is quite special (as then there is uniquely no effect). This is often called "unnatural". An exactly true model may very well have soft problems, as they are not true contradictions. The use of soft problems is that historically they have often provided good guidance for how to improve a theory. In particular, one of the Standard Model's biggest hard problems are that it does not have any form of gravity in it. In this book we have yet to attempt to add gravity into this theory but we will unfortunately find it is not very straightforward. The other is that experimentally there is a large abundance of matter out in space which does not interact electromagnetically but does gravitationally. That matter is called dark matter. The only matter present in the Lagrangian density we have presented which does not interact electromagnetically is the set of right hand neutrinos, which have not been directly experimentally observed and are precisely the piece not present in the traditional Standard Model. Even then, these hypothetical right handed neutrinos are heavily disfavored as by experimental evidence as an explanation for the observed properties of dark matter.

In addition to those two hard problems which need to be solved, the Standard Model also has a number of peculiar features. For example, why are there 3 generations of matter? Why do the masses of the leptons and quarks vary over so many orders of magnitude? Why did the Standard Model have to have the gauge group $U_Y(1)xSU(2)xSU(3)$? Are baryon and lepton conservation real symmetries? If not, are there Majorana mass terms needed to be added? How many right handed neutrino species are there? Everything except the weak interaction has parity symmetry; why does the weak interaction alone break that symmetry? Everything except the CKM phase and PMNS phase have symmetry under the combined transformation of parity and charge conjugation. Why do those break it? The Standard Model is not on its own capable of answering any of these questions but perhaps these questions or questions like these can provide guidance on how to create a better model beyond the Standard Model.

## 5.5 Standard Model Effective Field Theory

## 5.6 Pati-Salam Model

## 5.7 Dark Matter

## 5.8 Supersymmetry

## 5.9 Summary and Additional Reading

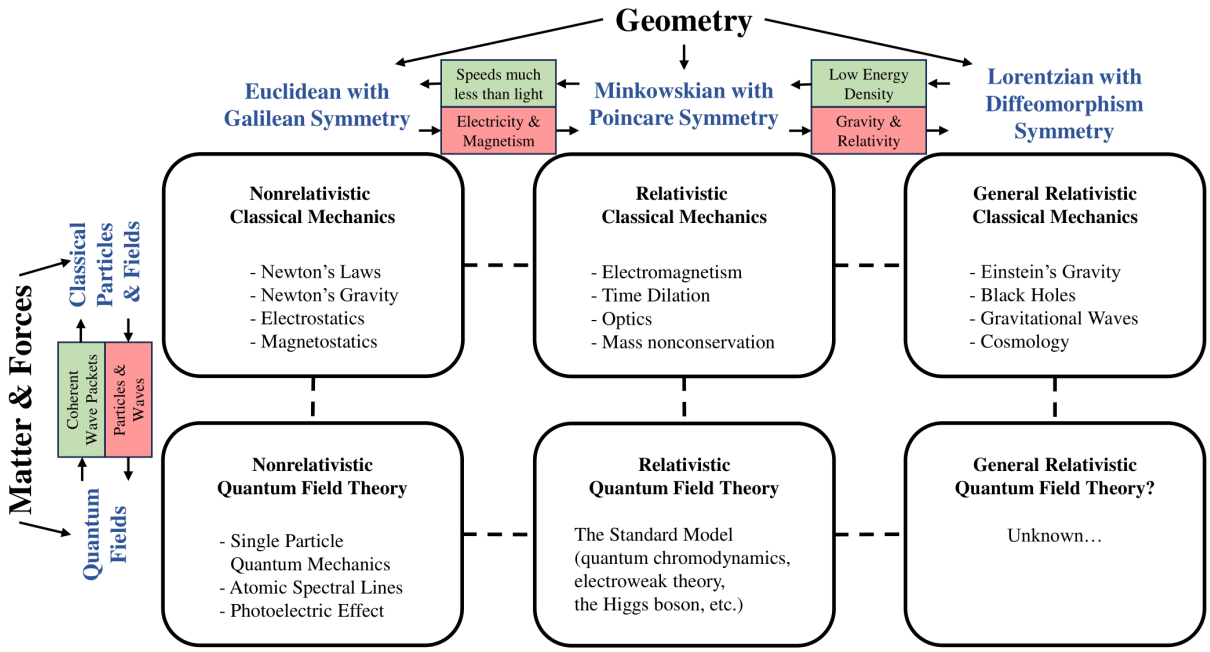



**Figure 5.1:** ???

# Chapter 6

# Quantum Gravity

*Physics is becoming too difficult for the physicists*
– David Hilbert [49]

*In science we must all submit not to what seems to us attractive from one point of view or another, but to what represents an agreement between theory and experiment.*
– Dmitri Mendeleev [32]

lol sike. The rest is up to you. Take inspiration from this poem that I took from Francisco Blanco [4]:

*If some new truth, O Friend! thy toil discover,*
*If thine eyes first by some fair form be blest,*
*Love it for what it is, and as a lover*
*Gaze, or with joy receive thine honoured guest:*
*The new-found Thought, set free, awhile may hover*
*Gratefully near thee, but it cannot rest.*

Sir William Rowan Hamilton [24]

Dirac equation in curved spacetime. Internal Lorentz transformation $\psi' = U\psi$, $U = e^{-i\theta^{AB}\hat{S}_{AB}}$.

$$\partial_\mu U U^{-1} = -i\int_0^1 e^{-i\varepsilon\theta^{AB}\hat{S}_{AB}}\partial_\mu\theta^{CD}\hat{S}_{CD}e^{i\varepsilon\theta^{EF}\hat{S}_{EF}}d\varepsilon \tag{6.0.1}$$

$$e^{-i\varepsilon\theta^{AB}\hat{S}_{AB}}\partial_\mu\theta^{CD}\hat{S}_{CD}e^{i\varepsilon\theta^{EF}\hat{S}_{EF}} = \phi^{AB}(\varepsilon)\hat{S}_{AB} \implies \frac{d\phi^{AB}}{d\varepsilon} = \frac{1}{i}[\hat{S}_{AB},\hat{S}_{CD}]\theta^{AB}\phi^{CD} \tag{6.0.2}$$

$$\frac{1}{i}[\hat{S}_{AB},\hat{S}_{CD}] = \eta_{BD}\hat{S}_{AC} - \eta_{AD}\hat{S}_{BC} - \eta_{BC}\hat{S}_{AD} + \eta_{AC}\hat{S}_{BD} \tag{6.0.3}$$

$$\implies \frac{d\phi^{AB}}{d\varepsilon} = 2\theta^A{}_C\phi^{BC} - 2\theta^B{}_C\phi^{AC} \tag{6.0.4}$$

$$L^A{}_B(\theta) = \left(e^{\theta^{CD}J_{CD}}\right)^A{}_B, \qquad (J_{CD})^A{}_B = \delta^A_D\eta_{BC} - \delta^A_C\delta_{BD} \tag{6.0.5}$$

$$\implies \frac{d\phi^{AB}}{d\varepsilon} = (\theta^{DE}J_{DE})^A{}_C\phi^{CB} + (\theta^{DE}J_{DE})^B{}_C\phi^{AC} \implies \phi^{AB}(\varepsilon) = L^A{}_C(\varepsilon\theta)L^B{}_D(\varepsilon\theta)\phi^{CD}(0) \tag{6.0.6}$$

$$\implies \partial_\mu U U^{-1} = -i\int_0^1 L^A{}_C(\varepsilon\theta)L^B{}_D(\varepsilon\theta)\partial_\mu\theta^{CD}d\varepsilon\hat{S}_{AB} = -iM^{AB}{}_\mu\hat{S}_{AB} \tag{6.0.7}$$

$$\implies \partial_\mu L^A{}_C L_B{}^C = M^{CD}{}_\mu(J_{CD})^A{}_B = -2M^A{}_{B\mu} \implies \partial_\mu U U^{-1} = -\frac{i}{2}L^A{}_C\partial_\mu L^{BC}\hat{S}_{AB} \tag{6.0.8}$$

$$\omega_\mu = K\omega^{AB}{}_\mu\hat{S}_{AB} \quad \text{and} \quad \omega'_\mu = U\omega_\mu U^{-1} + U\partial_\mu U^{-1} \implies K = \frac{i}{2} \tag{6.0.9}$$

$$\implies \nabla_\mu\psi = \partial_\mu\psi + \frac{i}{2}\omega^{AB}{}_\mu\hat{S}_{AB}\psi. \tag{6.0.10}$$

## 6.1 Quantum Field Theory in Curved Spacetime

## 6.2 Naive Quantization

## 6.3 Quadratic Gravity, Renormalizability, and Asymptotic Safety

## 6.4 Loop Quantum Gravity

## 6.5 String Theory

# References


[1] Allahverdyan, A. E., Balian, R., and Nieuwenhuizen, T. M. *A sub-ensemble theory of ideal quantum measurement processes*. *Annals Phys.* **376**. (2017). 324. [1303.7257].

[2] Ballentine, L. E. *Quantum Mechanics: A Modern Developmet*. World Scientific Publishing Company, 2014. 978-981-4578-57-8.

[3] Becker, A. *What Is Real?: The Unfinished Quest for the Meaning of Quantum Physics*. Basic Books, 2018, 978-0465096053.

[4] Blanco, F. M. *Nonperturbative approach to gravitational self-force and Hamiltonian formulation of its conservative dynamics*. *PhD Thesis*. (2025). 1.

[5] Blinder, S. M. *Canonical partition function for the hydrogen atom via the Coulomb propagator*. *Journal of Mathematical Physics*. **36**. (1995). 1208.

[6] Boccara, N. *Functional Analysis: An Introduction for Physicists*. Elsevier, 1990.

[7] Bonilla, L. L. and Guinea, F. *Collapse of the wave packet and chaos in a model with classical and quantum degrees of freedom*. *Phys. Rev. A*. **45**. (1992). 7718.

[8] Boyce, W. E. and DiPrima, R. C. *Elementary Differential Equations and Boundary Value Problems*. Wiley, 2005, 978-0-471-43338-5.

[9] Breuer, H.-P. and Petruccione, F. *Quantum measurement and the transformation from quantum to classical probabilities*. *Phys. Rev. A*. **54**. (1996). 1146.

[10] Breuer, H. and Petruccione, F. *The Theory of Open Quantum Systems*. Oxford University Press, 2002, 9780198520634.

[11] Brown, J. W. and Churchill, R. V. *Complex Variables and Applications*. McGraw Hill, 2013, 978-0073383170.

[12] Coleman, S. *Lecture delivered at Harvard University*, 1975.

[13] Colley, S. *Vector Calculus*. Pearson Education, 2022, 9780136800064.

[14] Corral, M. *Vector Calculus*. Schoolcraft College, 2008, GNU Free Documentation.

[15] Mesures, B. I. *Le Système international d'unités/The International System of Units*. SI, 2025. SI.

[16] Everett, H., III. *The Theory of the Universal Wave Function. PhD Thesis*. (1956). 1.

[17] Fewster, C. J. and Verch, R. *Measurement in Quantum Field Theory*. 2304.13356.

[18] Feynman, R. *Lecture delivered at Cornell University*, 1964.

[19] Feynman, R. *Fun to Imagine*. British Broadcasting Corporation, 1983.

[20] Friedberg, S., Insel, A., and Spence, L. *Linear Algebra*. Featured Titles for Linear Algebra (Advanced) Series. Pearson Education, 2003, 9780130084514.

[21] Giles, R. *The Reconstruction of Gauge Potentials From Wilson Loops*. *Phys. Rev. D*. **24**. (1981). 2160.

[22] Glashow, S. *The Elegant Universe*. NOVA, Public Broadcasting Service, 2003.

[23] Greene, R. E. and Krantz, S. G. *Function Theory of One Complex Variable*. American Mathematical Society, 2006, 978-0821839621.

[24] Hamilton, W. R. *On Unselfishness In The Pursuit Of Truth And Beauty, to Prof. J. C. Adams, Discovery of Neptune*, 1847.

[25] Heslot, A. *Quantum mechanics as a classical theory.* *Phys. Rev. D.* **31**. (1985). 1341.

[26] Horn, R. and Johnson, C. *Matrix Analysis.* Cambridge University Press, 2012, 9781139788885.

[27] Huff, D., Prewett, O., and College, G. A. *The Nature of the Physical Universe: 1976 Nobel Conference.* A Wiley-Interscience publication. Wiley, 1979, 9780471031901.

[28] Koopman, B. *Quantum Theory and the Foundations of Probability.* In *Applied Probability*, volume 7 of *Proceedings of the Symposia in Applied Mathematics.* McGraw-Hill, 1957.

[29] Lamb, W. E., Jr. and Scully, M. O. *The photoelectric effect without photons.* NASA Contractor Report, NGR-10-007-010, ARPA SD-90, NGR-07-004-035, February 1968.

[30] Mandel, L. and Wolf, E. *Optical Coherence and Quantum Optics.* Cambridge University Press, 1995, 978-0521417112.

[31] Manzano, D. *A short introduction to the Lindblad master equation.* *AIP Adv.* **10**. (2020). 025106. [1906.04478].

[32] Mendeleev, D. *Faraday Lecture at the Royal Institution, London*, 1889.

[33] Merzbacher, E. *Quantum Mechanics.* Wiley, 1998, 978-0471887027.

[34] Mostow, M. A. and Shnider, S. *Counterexplains to some results on the existence of field copies.* *Commun. Math. Phys.* **90**. (1983). 521.

[35] Mott, N. F. *The solution of the wave equation for the scattering of particles by a coulombian centre of force.* *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character.* **118**. (1928). 542.

[36] Mott, N. F. *The wave mechanics of -Ray tracks.* *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character.* **126**. (1929). 79.

[37] Navas, S. et al. *Review of particle physics.* *Phys. Rev. D.* **110**. (2024). 030001.

[38] Neumaier, A. *Coherent Quantum Physics: A Reinterpretation of the Tradition.* Texts and Monographs in Theoretical Physics. De Gruyter, 2019, 9783110667363.

[39] Nieuwenhuizen, T. *Is the Contextuality Loophole Fatal for the Derivation of Bell Inequalities?* *Found Phys.* **41**. (2010). 580.

[40] Olver, P. J. *Introduction to Partial Differential Equations.* Undergraduate Texts in Mathematics. Springer, 2013, 978-3319020983.

[41] Olver, P. and Shakiban, C. *Applied Linear Algebra.* Undergraduate Texts in Mathematics. Pearson Prentice Hall, 2018, 9783319910413.

[42] Peskin, M. and Schroeder, D. *An Introduction To Quantum Field Theory.* CRC Press, 2018, 9780429972102.

[43] Planck, M. *Scientific Autobiography: And Other Papers.* Philosophical Library/Open Road, 2014, 9781497675889.

[44] Ralston, J. P. *PT and CPT quantum mechanics embedded in symplectic quantum mechanics.* *Journal of Physics A Mathematical General.* **40**. (2007). 9883.

[45] Ralston, J. P. *Emergent mechanics, quantum and un-quantum.* In Roychoudhuri, C., Kracklauer, A. F., and Raedt, H. D., editors, *The Nature of Light: What are Photons? V*, volume 8832, page 88320W. International Society for Optics and Photonics, SPIE, 2013, https://doi.org/10.1117/12.2025000.

[46] Ralston, J. P. *Quantum Probability.* University of Kansas, 2015, KU.

[47] Ralston, J. P. *Understanding Basic Quantum Mechanics.* University of Kansas, 2016, KU.

[48] Ralston, J. P. *How to Understand Quantum Mechanics.* 2053-2571. Morgan & Claypool Publishers, 2018. 978-1-6817-4226-7.

[49] Reid, C. *Hilbert.* Springer Berlin, Heidelberg, 1970. 978-3-662-27132-2.

[50] Richeson, D. and Wiseman, J. *A fixed point theorem for bounded dynamical systems.* *Illinois Journal of Mathematics.* **46**. (2002). 491. [math/0108064].

[51] Sakurai, J. and Napolitano, J. *Modern Quantum Mechanics.* Cambridge University Press, 2017, 9781108422413.

[52] Salam, A. *Grand Unification of Fundamental Forces*. Nobel Lecture, December 8, 1979.

[53] Salam, A. *Nobel Banquet*. Nobel Prize Outreach, December 10, 1979.

[54] Schonfeld, J. F. *The First Droplet in a Cloud Chamber Track*. *Foundations of Physics*. **51**. (2021). 47.

[55] Schrödinger, E. *Letter to John Lighton Synge*. November 9, 1959.

[56] Schwinger, J. S., DeRaad, L. L., Jr., and Milton, K. A. *Casimir Effect in Dielectrics*. *Annals Phys.* **115**. (1978). 1.

[57] Shankar, R. *Principles of Quantum Mechanics*. Physics and Astronomy. Springer US, 2012, 9781475705768.

[58] Spivak, M. *Calculus*. Calculus. Cambridge University Press, 2006, 9780521867443.

[59] Srednicki, M. *Quantum Field Theory*. Cambridge University Press, 2007, 9781139462761.

[60] Strogatz, S. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. CRC Press, 2018, 9780429961113.

[61] Weinberg, S. *Dreams of a Final Theory: The Scientist's Search for the Ultimate Laws of Nature*. Vintage books. Knopf Doubleday Publishing Group, 1994, 9780679744085.

[62] Weinberg, S. *The Quantum Theory of Fields*. Quantum Theory of Fields, Vol. 2: Modern Applications. Cambridge University Press, 1995, 9780521550017.

[63] Weinberg, S. *Lectures on Quantum Mechanics*. Cambridge University Press, 2015, 9781107111660.

[64] Zemeckis, R. and Gale, B. *Back to the Future*. Universal Pictures, 1985.