THE UNIVERSITY of ADELAIDE

Department of Physics

Isospin Violating Hadronic Mass Splittings using Lattice QCD+QED

BY

Zachary Koumi

Supervisors:

Ross Young & James Zanotti

A thesis submitted towards the degree of
Doctor of Philosophy
at
The Faculty of Sciences
The University of Adelaide

September 6, 2019

Acknowledgements

I would like to thank my supervisors Ross Young and James Zanotti for providing direction and advice throughout my PhD. Waseem Kamleh for his help with Cola in an emergency. Robert Perry, Ryan Bignell, Kim Somfleth, Finn Stokes, Alex Westin, Daniel Murnane and Alex Chambers as well as members of the S4S group for providing advice, and acting as a sounding board when I needed it. Damir, Marco, Anum and all the occupants of my PhD room for contributing significantly to office spirit.

I want to thank QCDSF and CSSM members for their academic support. The physics department teaching and administration staff for making research possible. Finally, I'd like to thank my family and friends for providing support when I was having a bad research day, week or month.

Abstract

The composition and structure of matter has excited scientific thought for millennia. One of the highlights of the previous century was the development of the theories of quantum chromodynamics (QCD) and quantum electrodynamics (QED), which reveal an even deeper layer of structure than the atom and the nucleon.

We use the non-perturbative method of lattice QCD+QED to make precision estimates of the masses and mass splittings of the light ground state hadron spectrum, including pseudoscalar mesons, octet baryons and decuplet baryons. We replicate this same analysis for ground state charmed hadrons. In these studies the QED component is necessary for two reasons. Firstly, it is necessary when attempting to obtain mass results with sub-percent precision. While secondly, it is essential in determining mass splittings between hadrons, QED is a substantial component of the mass splittings within an isospin multiplet, such as the Δ baryons. Our findings provide new insight into these splittings by separating the contributions arising from strong and electromagnetic effects.

We use lattice QCD+QED to determine the flavour-neutral pseudoscalar meson masses, which incorporate disconnected quark line diagrams. We provide estimates of the absolute mass and mass splitting of the lowest two states, near a point of quark mass degeneracy. We show that QED plays an important role in the flavour composition of states around points of approximate quark mass degeneracy, which is important at the physical quark mass and charge.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the Universitys digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Zachary Koumi

Contents

CHAPTER 1

Introduction

The composition and structure of matter has excited scientific thought for millennia, and obtaining an understanding of its composition allows us to understand and predict the world around us. Matter is composed of atoms, including atomic elements such as hydrogen and helium. Atoms are composed of a positively charged nucleus, with electromagnetically bound electrons surrounding it. Most of the mass of an atom is located within the dense, positively charged atomic nucleus, which is known to be a bound state of protons and neutrons, collectively referred to as "nucleons".

Nuclei with more than one nucleon are bound by a very strong force, called the strong force. The strong force binding the atomic nuclei is understood using Quantum Chromodynamics (QCD).

Quantum Chromodyanamics is the quantum field theory which describes (to a great extent) the behaviour of quarks. Quarks are not observed as single particles, rather they appear as constituents of baryons (three quark bound states qqq) and mesons (quark anti-quark bound states $q\bar{q}$). Nucleons (the proton and neutron) are included in the set of baryons, and are the lightest three-quark bound states.

As there is no formal solution to QCD it is difficult to produce predictions of quantities, such as hadron masses. Worse, due to its strong coupling nature at low energies, perturbation theory can not be applied. The only known way to directly probe the observables of QCD at low energies is numerically, which involves direct calculation of the fields at a discrete and finite number of space-time points. This technique is called lattice gauge theory, and can be applied to gauge theories such as QCD.

In this work we determine estimates for the mass splittings between the lowest energy mesons and baryons, as well as mass splittings between hadrons that contain

at least one charm quark, called charmed hadrons. A novel aspect of this study is that we determine mass splittings due to isospin violation, incorporating important Quantum Electrodynamic (QED) effects.

Isospin is the approximate symmetry that makes the nucleons and other isospin multiplets almost degenerate in mass. Though the total mass of baryons and mesons are known to be generated primarily through QCD, Quantum Electrodynamics (QED) becomes important below the 1% level. When calculating the isospin violating mass splittings, the QED components are comparable to the strong QCD components. To determine these splittings we employ lattice QCD while also including Quantum Electrodynamics (QED) on the lattice. This is called lattice QCD+QED.

Isospin violation is relevant to a range of physical phenomena, including the flavour decomposition of nucleon structure [1, 2, 3, 4]; tests of neutrino-nucleus interactions [5, 6]; precision constraints on CKM [7, 8] matrix elements from leptonic [9, 10] and semi-leptonic [11] decay rates; and quark mass parameters [12, 13, 14, 15].

I begin this thesis by describing the quantum field theories of interest, Quantum Chromodynamics and Quantum Electrodynamics in chapter 2[1]. In this chapter I describe how QCD and QED are related and define these theories in terms of the path integral partition function. I also describe SU(3) flavour symmetry and why it is important.

Due to the difficulties in producing solutions from the QCD partition function, and the requirement for regularisation and renormalisation of the field theory we use lattice gauge theory in this thesis to produce estimates of observables. Lattice gauge theory is an numerical approximate method to calculate observables for gauge field theories. In chapter 3 I describe what lattice gauge theory is, how it can be applied to gauge theories with fermions (quarks) and how to determine hadron masses from this discretised field theory. I approach the subject more generally so that it can be applied to either QCD or QED, or both.

It is significantly more time efficient to calculate observables on the lattice with large quark masses than it is to calculate them with small quark masses. This is related to how difficult it is to invert the fermion matrix. In our analysis we use larger than physical quark masses, and using the information obtained at this point in parameter space, we extrapolate our observables to the physical quark mass point. In chapter 4 I describe the methodology for choosing the point in parameter space that we use, called the SU(3) point, how we determine this point

[1] I cover some more fundamental topics in the appendix. I start with the concept of minimisation as applied to physical phenomena and quantum mechanics in appendix A. I then give a very quick introduction to quantum field theory in appendix B and extend the application of field theory to fermions in appendix C. Finally I introduce the concept of a Yang-Mills type gauge field theory in appendix D.

and the analytic expansions that we use to extrapolate from this point to the physical point.

The first chapter of new research in this thesis, chapter 5, presents results for the mass splittings between isospin multiplets of the octet and decuplet baryons including QED. I describe lattice particulars such as quark masses and the action, describe the fitting procedure for the extrapolation and discuss results of the analysis.

I produce a similar analysis for charmed hadrons in chapter 6, producing results for the mass splittings for isospin multiplets as well as hyperfine splittings. In this chapter we also examine the discrepancy of the experimental results of the SELEX collaboration[16] and LHCb[17] and discuss ways to make these observations compatible.

In the last chapter of new research, chapter 7, we investigate the composition of the π^0–η–η' system near a point of exact SU(3) symmetry. Through this analysis we determine the π^0–η mass splitting generated by electromagnetic isospin violation. We study the state composition around this SU(3) point and try to relate this to the behaviour of the mass splitting near the physical quark mass point.

Finally, I conclude this thesis with a summary of our results and outlook to future research in chapter 8.

CHAPTER 2

Quantum Electrodynamics and Chromodynamics

Our goal in this chapter is to introduce the two field theories important to the research in this thesis; Quantum Electrodynamics (QED) and Quantum Chromodynamics (QCD). The two theories are Yang-Mills type gauge theories, where QED has a $U(1)$ gauge symmetry, while QCD has a SU(3) gauge symmetry.

We can show that the two theories produce the same kind of force; in section 2.1 I will show that (in the classical limit) the forces produced by the Yang-Mills type field theories, specifically QCD and QED, are the simplest forces that can exist in a special relativistic theory. To achieve this, I derive the general form for a field strength tensor using special relativity and some assumptions about the force that will be produced, and show this field strength tensor to be of the same form as the Yang-Mills field strength tensor. In section 2.2 I give a summary of electrodynamics and QED, producing the partition function and Feynman rules for QED.

In section 2.3, because there is no classical theory of QCD, I give a historical account of the development of the theory through the progressively more complicated experiments performed to probe the structure of the nucleus of atoms. I then use the hadron spectrum to motivate the introduction of the colour quantum number. The quantum number is recognised as colour charge once the SU(3) local gauge invariance is postulated. In section 2.3.3 I outline the specifics of QCD, and include the partition function and Feynman rules for the theory.

In section 2.4 I explain how SU(3) flavour symmetry presents itself in the hadron spectrum, and explain what causes this symmetry. I then motivate this explanation by using SU(3) flavour symmetry in the quark masses to reproduce the grouping seen in the hadron spectrum. In section 2.4.1 I provide some context to the allocation of quark flavours to hadron states.

2.1 Forces in special relativity

To properly understand the electromagnetic forces we need to understand how the concept of force fits into special relativity. We will begin with a definition of force. We define 4-force as,

$$F = \frac{d}{d\tau}P = \frac{d}{d\tau}(m_0 U) = m_0 A + \frac{dm_0}{d\tau}U \tag{2.1}$$

for 4-momentum P, acceleration A, velocity U , rest mass m_0 and proper time τ. This satisfies our general requirements; P remains constant if there is no applied force. If the rest mass does not change this reduces to something that looks like the Newtonian force law, $F = m_0 A$. Applying the time derivative directly to the momentum vector,

$$F = \gamma \frac{d}{dt}(mc, \vec{p}) = \gamma(u)\left(\frac{1}{c}\frac{dE}{dt}, \vec{f}\right). \tag{2.2}$$

There are two important facts that we need to know before we begin to construct the simplest Lorentz covariant force, which we will show is the type of force that appears in QCD and QED. Firstly, the product,

$$F \cdot U = c^2 \frac{dm_0}{d\tau} = \gamma^2(u)\left(\frac{dE}{dt} - \vec{f}\cdot\vec{u}\right), \tag{2.3}$$

shows that forces that do not increase the particles rest mass, called pure forces, are space-like, hence $F \cdot U = 0$. Conversely, heat-like forces (forces that increase the rest mass) are not space-like. Secondly, by considering the Lorentz transformation equation for forces,

$$f_b' = \frac{f_b - vQ/c^2}{1 - u_b v/c^2}, \tag{2.4}$$

$$f_o' = \frac{f_o}{\gamma(v)(1 - u_b v/c^2)}, \tag{2.5}$$

$$Q' = \frac{Q - v f_b}{1 - u_b v/c^2}, \tag{2.6}$$

where f_b is the force in the boost direction, f_o represents the transformation for the other two space directions orthogonal to u_b and $Q = \frac{dE}{dt}$. Hence any force that is independent of velocity in one frame, has velocity dependence in every other frame.

The simplest conceivable Lorentz covariant force will be $F^\mu = E^{\mu\nu}U_\nu$ (if it exists) for some Lorentz tensor E. We also require that the force is pure, so

there is no frame in which it simply increases the rest mass, $F \cdot U = 0$. This requirement leads to $E_{\mu\nu}U^\mu U^\nu = 0$, for arbitrary velocity vector U, indicating E is antisymmetric, $E_{\mu\nu} = -E_{\nu\mu}$. The force may or may not act on every type of matter, so we associate a property called charge to those particles or waves which the force acts on. The charge on particles could be binary, discrete or continuous; however the force must scale with the amount of charge. We change our force equation to reflect this $F^\mu = E^{\mu\nu}J_\nu$, with new symbol $J_\nu = qU_\nu$ called current.

The tensor E has different values at different space-time points, and so it can be considered a force field. As we have not specified the source of the force field, the distribution of E in space-time is yet to be defined. Conceivably the charges themselves could produce the field, this is the case with known forces such as gravity and electromagnetism. If this was the case we would expect $\partial^\mu E_{\mu\nu} \propto J_\nu$. Automatically, by the antisymmetric property of E, $\partial^\nu J_\nu = 0$, current is conserved. As current is conserved, J has only three independent components, and the same is true for E.

We can write $E^{\mu\nu} = K^{\mu\nu}(\Phi)$ for some function K only dependent on 4-vector $\Phi \propto J$. As $\partial E \propto J$, we should be able to write E as a product of a 4-vector and Φ. We only have two distinct 4-vectors from which we could construct E, $J \propto \Phi$ and ∂. However if we assume the elements of Φ are scalar numbers, $E^{\mu\nu} = J^\mu\Phi^\nu - J^\nu\Phi^\mu = 0$ because J and Φ commute. E can only be $E^{\mu\nu} = \partial^\mu\Phi^\nu - \partial^\nu\Phi^\mu$. Hence E can be written in terms of a four potential Φ, where $\partial^\mu(\partial^\mu\Phi^\nu - \partial^\nu\Phi^\mu) = aJ^\nu$. E satisfies the gauge condition $\Phi^\mu \to \Phi^\mu + \partial^\mu\alpha(x)$ for some arbitrary scalar field α, which shows that Φ only has three degrees of freedom.

Conversely, if we do not assume the elements of Φ commute, then the more general expression for E is given by all relevant antisymmetric two index tensors $E^{\mu\nu} = a(\partial^\mu\Phi^\nu - \partial^\nu\Phi^\mu) + b[\Phi^\mu, \Phi^\nu]$ for constants a and b, which is the same form as the Yang-Mills field strength tensor. With appropriate choice of constants this field strength tensor satisfies the Yang-Mills gauge condition (for details see appendix D).

2.2 Electrodynamics

Maxwell's equations which describe the electromagnetic force fields $\vec{E}$ and $\vec{B}$ can be written in the form [18]

$$\partial_\mu F^{\mu\nu} = J^\nu, \tag{2.7}$$

for current vector $J = (c\rho, \vec{j}) = \rho U$, velocity vector $U = \gamma(c, \vec{u})$, electromagnetic potential $A = (cV, \vec{A})$ and field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. By definition of the Lorentz transformation, Maxwell's equations are Lorentz invariant. As the

equation is Lorentz invariant, and ∂ and U are covariant, we then expect all indices to transform covariantly – hence A is a Lorentz vector. Notice, seemingly we can only get two of Maxwell's equations from the above equation, however the last two Maxwell equations (which can be written as $\frac{1}{2}\partial^\alpha \epsilon_{\alpha\beta\mu\nu} F^{\mu\nu} = 0$) simply enforce the condition that $F_{\mu\nu}$ can be written in terms of a potential A_μ. Hence all four equations are included in the above definition when $F_{\mu\nu}$ is written in terms of a potential A. As $F^{\mu\nu}$ is antisymmetric by definition, Maxwell's equations also imply $\partial_\mu \partial_\nu F^{\mu\nu} = 0$ as the bottom indices are symmetric. Hence $\partial J = 0$, which states charge is conserved locally. One Lagrangian that produces these equations of motion is $\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$. Note that because the indices of $F_{\mu\nu}$ are anti-symmetric under interchange, it is invariant under the transformation $A_\mu \to A_\mu + \partial_\mu \alpha$ for some scalar field α. Hence, electromagnetism is a Yang-Mills type gauge theory.

2.2.1 QED

Quantum electrodynamics or QED is a Yang-Mills type gauge field theory where charged fermions $\psi(x)$ interact via gauge fields $A_\mu(x)$. The gauge transformation is

$$A_\mu \to A_\mu + \partial_\mu \alpha \tag{2.8}$$

for scalar field α, which is the linearisation of the unitary group $U(1)$, $U = e^{i\alpha(x)}$ for $\alpha(x) \in \mathbb{R}$. Here the generator can be represented by $T = 1$. The Lagrangian for QED can be written as

$$\mathcal{L}_{QED} = \bar{\psi}(x)(\not{D} + m)\psi(x) + \frac{1}{4}\mathrm{tr}\left(F^{\mu\nu} F_{\mu\nu}\right), \tag{2.9}$$

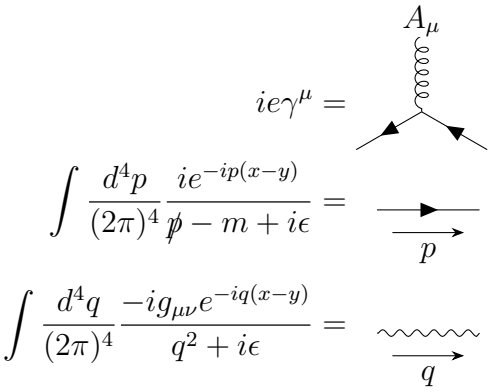
where $F^{\mu\nu} = \partial^\mu A^\nu(x) - \partial^\nu A^\mu(x)$, slash notation and the covariant derivative are defined in appendix C and D. As the gauge group U is single dimensional, it is abelian and the resulting gauge field is as well $[A^\mu(x), A^\nu(x)] = 0$, for $\mu, \nu \in \{0, 1, 2, 3\}$.

The partition function is given by

$$Z = \int D\bar{\psi} D\psi D A e^{i\int d^4x \mathcal{L}_{QED} + \frac{i}{2\xi}\int d^4x (\partial A)^2}, \tag{2.10}$$

where we have included the gauge fixing term[1]. The Feynman rules for this theory in the Feynman gauge $\xi = 1$ are,

[1] More detail on the gauge fixing term can be found in appendix D.



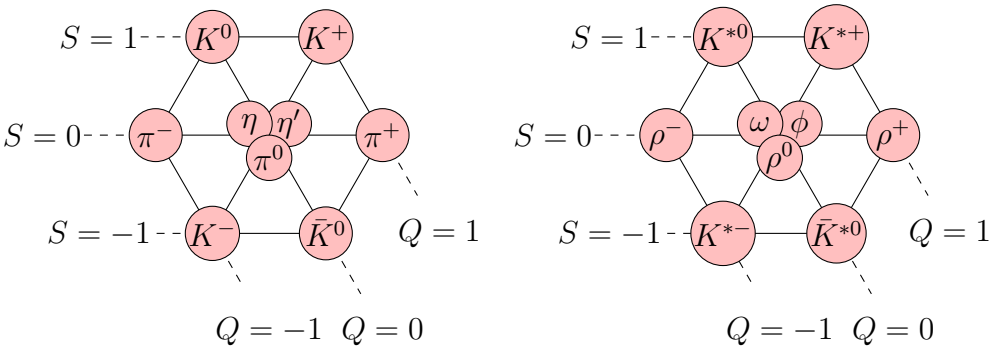
2.3 Hadrons and the Strong Force

Thompson's discovery of the electron in 1897 raised the question of where and how positive charge and mass were distributed inside an atom, as atoms were known to be both heavy and charge neutral. Rutherford's research team found that α-particles would scatter elastically from gold foil, which suggested the existence of a heavy positive charge centre and this heavy charge centre was named the nucleus. High energy elastic scattering experiments using electrons later showed that the nucleus did not act like a point object and diffraction effects due to its finite extent were detected. Inelastic scattering also showed the existence of excited states [19]. The finite extent and excited states of the nucleus are both evidence for a nucleus made of constituents.

The lightest atom discovered, hydrogen has a single electron and its nucleus was named the proton. Over a long period it was thought these protons might be the building blocks for atoms. The force that bound them was thought to be very strong because when building the atomic spectrum, the nucleus was made of increasing numbers of protons with positive electric charge. It was also thought to be of short range because such a force had not previously been detected and seemed to only act in the compact radius of the nucleus. For atoms with increasing proton number, the atomic mass seemed to increase, but not directly with proton mass. In 1932 Chadwick discovered the neutron, which was neutrally charged and had almost the same mass as the proton. This discovery solved the mass discrepancy; to good approximation atomic masses could be predicted using protons and neutrons as constituents [19].

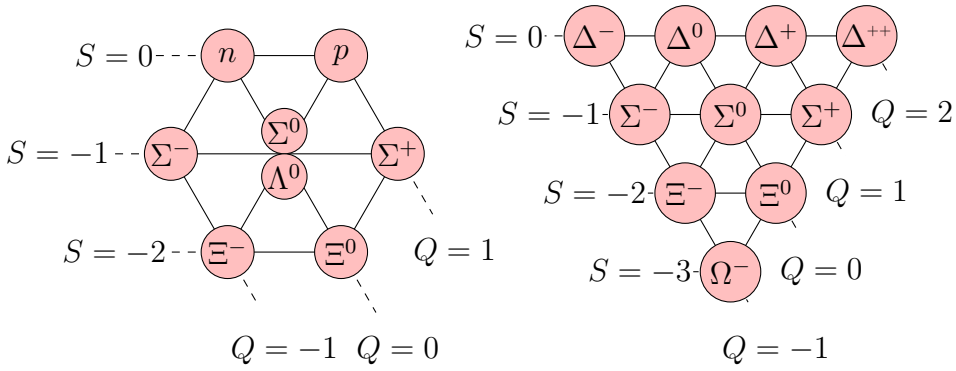
In a cosmic ray experiment in 1947, the kaon was discovered. By the end of the 1950's, many strange particles had been found both mesons and baryons. By the early 1960's the light hadron spectrum was on its way to complete discovery. Interestingly the hadrons seemed to fit into distinct mass and spin groups. Gell-Mann

arranged the particles in these mass and spin groups in terms of the charge and strangeness. This arrangement was called the Eightfold Way. The arrangement of particles is shown in Fig. 2.1 and Fig. 2.2.



(a) (b)

Figure 2.1: S indicates the strangeness of the particles. (a) Spin zero, negative parity, positive charge conjugation, $J^P = 0^-$ hadrons, called pseudoscalar mesons. The central mesons have positive charge conjugation. (b) Spin one, negative parity, negative charge conjugation, $J^P = 1^-$ hadrons, called vector mesons. The central mesons have negative charge conjugation.



(a) (b)

Figure 2.2: (a) Spin $\frac{1}{2}$ positive parity, $J^P = \frac{1}{2}^+$ hadrons, called octet baryons. (b) Spin $\frac{3}{2}$ positive parity, $J^P = \frac{3}{2}^+$ hadrons, called decuplet baryons. S indicates the strangeness of the particles.

Electron scattering experiments by Hofstadter in 1963 revealed that the proton had an exponential electromagnetic charge distribution with a root mean square

radius of approximately 0.8 fm. Both the proton and neutron had a magnetic moment distribution, and these decay with the same curve as the charge distribution when appropriately scaled. The shape of the distributions suggested that both of these particles had spatial extent, in the same way as the atom and the nucleus had shown to [19]. Further inelastic scattering studies showed that states inside the nucleons could be excited [19].

2.3.1 Asymptotic Freedom

The first proton-proton scattering experiments at high energy, above 10 GeV in the centre of mass frame, showed unexpected results. The protons were broken apart into other hadronic matter, mostly pions, but these hadron jets were produced collinear with the collision axis. It was puzzling because the strong force, which people believed to hold the nucleus together, was predicted to be strong and to bind positively charged protons; however, because two hadrons had only a weak interaction, at high relative momentum the proton seemed to behave like a cloud of loosely bound constituents [20].

Further analysis of the proton occurred in the late 1960's at SLAC. In this new experiment 20 GeV electrons were scattered off a hydrogen target and the scattering rate was measured for large deflection angles, called hard scattering [20]. Based on the previous experiment, it was expected that the quarks would again seem to be weakly interacting, resulting in low hard scattering rates. Instead a substantial rate for hard scattering was detected, comparable to if the proton was modeled as a fundamental particle, with the exception that the proton was rarely observed in the final state. The scattering was still mostly deep inelastic, producing a large number of hadrons [20]. This observation suggested the existence of discrete scattering centres within the proton, and the distribution of the scattering electrons exhibited a phenomenon called scale invariance, suggesting the hard scattering centres had no internal structure, and hence were 'pointlike' [21, 22, 23, 24].

The theory proposed by Feynman and Bjorken to explain this discrepancy is called the parton model, in which the proton is assumed to be a loosely bound collection of constituents called partons. These include fermion constituents carrying electric charge, what we now call quarks, and electrically neutral components, which we now call gluons. It was assumed that these constituents were incapable of large momentum exchange through the strong interaction [20]. They can however have large momentum transfer through the electromagnetic interaction. For instance, an electron can knock a quark from a proton. This quark then softly interacts with the remaining parts of the proton pulling the shattered pieces of the proton in the same direction, to produce a hadron jet in the direction of the displaced quark.

The weakness in the strong force at short distances and high energies is called

asymptotic freedom. All field theory couplings change at different length scales. In the case of QED, the coupling gets stronger at close distances. The theory for the strong force needed to be of a type that gets weaker at close distances. When Bjorken scaling was discovered, there were no known asymptotically free field theories in four dimensions. It was later shown in the early 1970s that non-abelian gauge theories exhibit asymptotic freedom, and were shown to be the only type of theory that had the property in four dimensions [20].

2.3.2 The Quark Model

In 1963, Gell-Mann and Zweig proposed the quark model, which explained the spectrum of strongly interacting particles called hadrons in terms of constituents called quarks [20]. Due to the approximate SU(3) flavour symmetry (which I will talk more about in section 2.4), the lightest bound states in the hadron mass spectrum can be nicely described by quantum numbers of isospin and strangeness[2], and form recognisable groupings in terms of their masses. Gell-Mann recognised the hadron spectrum matched the irreducible representations of the SU(3) (flavour symmetry) if bound states were assumed to be of the form $\bar{q}q$ and qqq, for quarks q and anti-quarks $\bar{q}$. Some diagrams of these groupings of particles are shown in Fig. 2.1 and Fig. 2.2.

Mesons were classified as quark–antiquark bound states. The quantum numbers of the real meson states agree with this description, as the lightest mesons come in spin-0 and spin-1 states of odd parity [20]. Baryons were classified as three quark bound states. To obtain the correct number of known baryons and mesons and the correct electric charges, Gell-Mann assumed three types of quarks up (u), down (d) and strange (s), with charges $+2/3, -1/3, -1/3$ respectively. These types are called quark flavours; quark flavours charm (c), bottom (b) and top (t) were later discovered with charges $+2/3, -1/3, +2/3$, respectively.

Despite phenomenological success in predicting the existence of unknown particles and its ability to satisfy the approximate flavour symmetries found in the known hadron spectrum, the spectrum of baryons required that the quarks be symmetric under interchange of quark spin and flavour quantum numbers [20]. However, quarks must also have spin 1/2, and so obey Fermi-Dirac statistics to ensure matter does not collapse; they cannot exist in the same state as an identical quark. The issue is most easily observed in the Δ^{++} bound state, supposedly uuu with spin $+3/2$, and symmetric under interchange.

This led to the introduction of an unobserved quantum number called colour, in which baryons were antisymmetric. Even if a baryon is symmetric in spin and

[2] It was later found that isospin denotes the number of up quarks and strangeness denotes the number of strange quarks. The remaining quarks are assumed to be down quarks unless otherwise indicated.

flavour, it still has an antisymmetric wave function; which requires the existence of at least three colours[3]. The modification to include colour in the wave function is simply $\epsilon^{ijk}q_iq_jq_k$, for colours $i, j, k \in \{1, 2, 3\}$.

It follows directly that the quarks transform under group S_3, a discrete permutation of colours, without affecting the spin-flavour wave function. However, this can be generalised to a continuous SU(3) invariance of the quark colour without relaxing our requirements; we still require the quark colours at the same point in space-time to transform together and to be in orthogonal directions. Hence for a colour transformation $U \in$ SU(3), a baryon transforms as $\epsilon^{ijk}q_i(x)q_j(x)q_k(x) \rightarrow \epsilon^{ijk}U_{ii'}U_{jj'}U_{kk'}q_{i'}(x)q_{j'}(x)q_{k'}(x) = det(U)\epsilon^{ijk}q_i(x)q_j(x)q_k(x) = \epsilon^{ijk}q_i(x)q_j(x)q_k(x)$, as $\det(U) = 1$ [20].

Furthermore, there is no reason to believe that the colours of quarks at all points in the Universe transform together, though it is clear that the colour of quark fields at the same space-time point seem to. For now, consider the colour degree of freedom to be a local SU(3) gauge freedom; this assumption is only validated by the experimental success of the QCD theory. Using this local gauge symmetry, we can create a field theory using the Yang-Mills formalism. In this field theory (QCD), colour turns out to be a type of charge which only quarks posses. Note, the gauge freedom does not generate any forces but the presence of a gauge freedom eliminates many possible terms in the potential and requires a gauge field to ensure the kinetic terms are invariant. The existence of a gauge freedom suggests the existence of a force.

The existence of baryons and mesons as colour neutral objects is not a fundamental property of the theory, but it emerges in QCD due to the increasing strength of the coupling constant at low energies. The property is called confinement. With three colours the simplest colour neutrals are of the form $\bar{q}_iq_j$, $\epsilon^{ijk}q_iq_jq_k$ and $\epsilon^{ijk}\bar{q}_i\bar{q}_j\bar{q}_k$. In the physical world these are the only types we have found and are called meson, baryons and antibaryons respectively.

2.3.3 QCD

Quantum chromodynamics or QCD is a Yang-Mills type gauge field theory where colour charged fermions $\psi_a(x)$, with colour index $a \in \{1, 2, 3\}$, interact via gauge fields $A_\mu(x)$. The gauge transformation is

$$A_\mu \rightarrow UA_\mu U^\dagger - i(\partial_\mu U)U^\dagger = UA_\mu U^\dagger + iU\partial_\mu U^\dagger \tag{2.11}$$

where $U \in SU(3)$ and is generated by $U = e^{iT^a\theta^a}$ where $\theta^a(x) \in \mathbb{R}$. Here the generators T^a, $a \in \{1, ..., 8\}$ can be represented by the Gell-Mann matrices. The

[3]Prior to the development of colour the quark model could be used to calculate $\pi^0 \rightarrow 2\gamma$. The amplitude was a factor of three too small, which suggests there are only three colours [25].

transformation can be linearised to, $A_\mu \to A_\mu + i\theta^a[T^a, A_\mu] + \partial_\mu \theta^a T^a$. The Lagrangian for QCD can be written as

$$\mathcal{L}_{QCD} = \sum_{a,f} \bar{\psi}^f_a(x)(\not{D} + m_f)\psi^f_a(x) + \frac{1}{4}\mathrm{tr}\left(F^{\mu\nu}F_{\mu\nu}\right), \tag{2.12}$$

for colours $a \in \{1,2,3\}$ and flavours $f \in \{u,d,s,c,b,t\}$ where $F^{\mu\nu} = \partial^\mu A^\nu(x) - \partial^\nu A^\mu(x) - i[A^\mu, A^\nu]$.

The partition function is given by

$$Z = \int D\bar{\psi}D\psi DA e^{i\int d^4x \mathcal{L}_{QCD} + iS_{ghost}(c^\dagger,c) - \frac{i}{2\xi}\int d^4x\ tr\left[(\partial A)^2\right]}, \tag{2.13}$$

where we have included the gauge fixing term as defined in appendix D. The Feynman rules for this theory in the Feynman gauge $\xi = 1$ are,

$$\int \frac{d^4p}{(2\pi)^4}\frac{ie^{-ip(x-y)}}{\not{p} - m + i\epsilon} = \quad p$$

$$\int \frac{d^4q}{(2\pi)^4}\frac{-ig_{\mu\nu}e^{-iq(x-y)}}{q^2 + i\epsilon} = \quad q$$

$$ig\gamma^\mu T^a = \quad a^a_\mu$$

$$(q-k)_\mu g_{\alpha\nu} + (p-q)_\nu g_{\alpha\mu} + (k-p)_\alpha g_{\nu\mu} = \quad q,\ p,\ k$$

$$\begin{aligned} -ig^2[&f^{abe}f^{cde}(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho}) \\ &+ f^{ace}f^{bde}(g_{\mu\nu}g_{\rho\sigma} - g_{\mu\sigma}g_{\nu\rho}) \\ &+ f^{ade}f^{bce}(g_{\mu\nu}g_{\rho\sigma} - g_{\mu\rho}g_{\nu\sigma})] = \end{aligned} \quad a^b_\nu,\ a^a_\mu,\ a^c_\rho,\ a^d_\sigma$$

$$\int \frac{d^4q}{(2\pi)^4}\frac{i\delta^{ab}}{q^2} = c_a \xrightarrow[q]{} c_b^\dagger$$

$$-gf^{abc}p^\mu = \quad a_\mu^b,\ p,\ c_a^\dagger,\ c_c$$

2.4 SU(3) Flavour Symmetry

Approximate flavour symmetries have been observed in hadronic bound states since the discovery of the fundamental building blocks of atoms, the proton and the neutron. The symmetry is observed in the closeness of the masses of the proton (938.27 MeV) and neutron (939.57 MeV). Conservation of baryon number and the low mass of the proton make it a stable bound state. The neutron decays to a proton via beta decay, however due to the closeness of their masses, a free neutron has a very long mean lifetime of approximately 15 min. This long decay lifetime and the approximate degeneracy of mass with the proton allows the neutron to form stable bound states with the proton, to form a nucleus. The existence of the neutron is essential in the nuclear stability of nuclei of higher atomic number, and even low atomic number elements have an equal number of protons and neutrons as ground states. Such a configuration is only possible because the binding energy of the nucleus is higher than the proton-neutron mass difference. The mass difference is so small, the electromagnetic effects inside the nucleus result in neutron stability; consider the stability of deuterium nucleus[4].

The symmetry between the proton and neutron, called isospin symmetry, is the result of the quark masses of the up and down quarks being almost degenerate, where the proton is a bound state of quarks uud and the neutron a bound state of quarks ddu. The observation of the mass symmetry between the proton and neutron and other isospin symmetries, as between the differently charged π, Δ, Σ and Ξ, were key motivators in the construction of the quark model.

The small mass of the u, d and s quarks create an approximate symmetry in the hadron mass spectrum, called SU(3) flavour symmetry or $SU(3)_F$, which is similar to isopsin symmetry. SU(3) flavour symmetry is an approximate symmetry found in the QCD Lagrangian for the lightest quark flavours u, d, s. These quarks

[4] Assume the nucleus is a collection of two types of fermions, protons and neutrons, in a potential well. The Fermi exclusion principle and QCD can account for a great deal of the neutron stability. However in the cases where there is an odd number of protons and an odd number of neutrons, neutron stability (especially in the deuteron case) can only be described by electromagnetic effects.

have much lower masses than the other quarks with u and d having masses less than 10 MeV and s having a mass of less than 100 MeV at probing energies typical to hadrons. The close grouping of the lightest hadron masses is an expression of these flavour symmetries in the QCD Lagrangian. Though isospin is the closest to being realised, there are other discrete symmetries that are possible. When d and s quarks are similar in mass, the symmetry is called U-spin, while when u and s quarks are similar it is called V-spin.

The concept behind flavour symmetry is simple; we assume that the quark masses and charges for the lightest three flavours are the same. Then the flavour symmetry is exact and the flavour quantum number for the lightest flavours u, d, s transform under an SU(3) global (flavour) symmetry. I will explain how this works and show that the hadrons form neatly into separate groups, which correspond to the physical spectrum.

The flavours of quarks q transform under the fundamental (or defining) representation, $\psi^i \to \psi'^i \to U^i{}_j \psi^j$, where $U \in SU(3)_F$ and ψ is a vector of quark flavours. Anti-quarks, which are members of the dual space, transform under the conjugate to the fundamental representation $\psi^{*i} \to (U^i{}_j)^* \psi^{*j} = U^\dagger{}_j{}^i \psi^{*j}$. The generators for the conjugate representation are the complex conjugate of the original Lie algebra. We can define ψ_j with a lower index to represent these dual space vectors, hence $*$ lowers indices so that $\psi_i \to \psi'_i = U^{\dagger j}{}_i \psi_j$. By taking the tensor product of vector spaces, we can obtain tensors of the form $\psi^{ik\ldots}{}_{jm\ldots} \in V_i \otimes V_k \ldots \bar{V}_j \otimes \bar{V}_m \ldots$ which transform like

$$\psi^{ik\ldots}{}_{jm\ldots} \to U^i{}_l U^k{}_s \ldots U^{\dagger n}{}_j U^{\dagger o}{}_m \ldots \psi^{ls\ldots}{}_{no\ldots}. \tag{2.14}$$

For simplicity consider a tensor of the form $\psi^i{}_j \in \mathbf{3} \otimes \bar{\mathbf{3}}$, where $\mathbf{3}$ is the vector space associated with the fundamental representation of SU(3) and $\bar{\mathbf{3}}$ is its conjugate. The tensor corresponds to a meson in SU(3) flavour symmetric QCD. This tensor can be broken down into a part that does not transform, it transforms under the 1-dimensional representation of the group SU(3), and a remaining 8-dimensional part

$$\psi^i{}_j \to U^i{}_l U^{\dagger k}{}_j \psi^l{}_k = \delta_{ij} \psi^k{}_k + U^i{}_l U^{\dagger k}{}_j \psi^l{}_k|_{k \neq l}. \tag{2.15}$$

The first part is called a fixed point set, and members of the fixed point set partition the space $\mathbf{3} \otimes \bar{\mathbf{3}}$ because the fixed points remain the same under the transformation. This set has dimension 1. The second part forms a class of orbits under the group action; the class of orbits also partition the space, and has dimension 8. For module X_G of group G and $x, y \in X_G$, an orbit forms an equivalence class, $x \sim y$ if $\exists g \in G$ such that $gx = y$, while fixed point sets form an equivalence class $x \sim y$ if $Gx = x$ and $Gy = y$.

If the orbits $\cup\{0\}$ and fixed point set form subspaces then it is possible to block diagonalise the group transformation by changing basis $\psi \to P\psi$ and $U \to PUP^{-1}$. SU(N) has a special property, if we separate the tensor product space into fixed point, symmetric and anti-symmetric parts, only these sets are both subspaces and orbits $\cup\{0\}$ and hence they transform under irreducible representations of the group SU(N). These are called invariant subspaces. Hence the decomposition in Eq. (2.15) forms invariant subspaces.

If we call the fixed point set $\mathbf{1}$ and the class of orbits $\cup\{0\}$ $\mathbf{8}$ then we can represent the tensor product of vector spaces (which corresponds to a meson in the quark model) as a direct sum of invariant subspaces $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$. This suggests that in the SU(3) limit, there are eight mesons that transform together and will follow the same mass trajectory, while the singlet (η') does not necessary follow that same trajectory.

It is relatively easy to show that if a tensor is symmetric or anti-symmetric that it remains symmetric or anti-symmetric under the group transformation. Consider the tensor $T^{ij} \in \mathbf{N} \otimes \mathbf{N}$ with symmetric part $S^{ij} = \frac{1}{2}(T^{ij} + T^{ji})$. Under the group transformation $S^{ij} \to S'^{ij} = U^{ik}U^{jl}S^{kl} = U^{ik}U^{jl}S^{lk} = S'^{ji}$. Similarly for anti-symmetric part $A^{ij} = \frac{1}{2}(T^{ij} - T^{ji})$, $A^{ij} \to A'^{ij} = U^{ik}U^{jl}A^{kl} = -U^{ik}U^{jl}A^{lk} = -A'^{ji}$. For fixed point elements, as they don't transform, they remain part of this subspace. It is more difficult to show that the groups can not be broken down any further and are hence irreducible.

It is not immediately clear how the subspace $\mathbf{8}$ transforms under the group operation, however as $\mathbf{1}, \mathbf{8}$ are invariant subspaces of the tensor, we must be able to block diagonalise the group representation into a 1-dimensional representation of SU(3) G_1 and an 8-dimensional representation of SU(3) G_8. It follows that G_8 is an irreducible representation of SU(3). This representation is coincidently isomorphic to the regular (or adjoint) representation, which is an 8-dimensional representation constructed from the structure constants of the Lie algebra. It is generally true that subspaces of the same dimension transform under isomorphic representations of SU(N).

Apart from representations of the same dimension, other group representations are isomorphic to each other as a result of the condition $\det(U) = 1$. For SU(N) this is equivalent to,

$$\epsilon_{i_1 i_2 \ldots i_N} U^{i_1}{}_{j_1} U^{i_2}{}_{j_2} \ldots U^{i_N}{}_{j_N} = \epsilon_{j_1 j_2 \ldots j_N} \tag{2.16}$$

noting that N in the identity is the dimension of the fundamental representation. By pre-multiplying by $U^\dagger$ we can show that,

$$\epsilon_{i_1 i_2 \ldots i_N} U^{i_2}{}_{j_2} \ldots U^{i_N}{}_{j_N} = U^{\dagger j_1}{}_{i_1} \epsilon_{j_1 j_2 \ldots j_N}. \tag{2.17}$$

By repeatedly multiplying by $U^\dagger$ we achieve $(N-1)/2$ identities for the group operation.

For SU(3) tensors, $\epsilon_{ijk}\psi^{jk}$ transforms like ψ_i, meaning the anti-symmetric part of ψ^{ij} transform like the conjugate ψ_k. One can also use the original definition to say that the anti-symmetric part of T^{ijk} transforms as $\mathbf{1}$. Using this idea we can show how that flavour symmetric baryons form invariant subspaces under the group operation.

In the quark model a baryon is of the form, qqq and corresponds to the tensor product of spaces $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = T^{ijk}$. With the notation $\{ij\}$ indicating anti-commuting indices and $[ij]$ indicating commuting indices, we break the tensor down into anti-symmetric and symmetric parts,

$$T^{ijk} = T^{\{ijk\}} + T^{\{[ij]k\}} + T^{[\{ij\}k]} + T^{[ijk]}. \tag{2.18}$$

These tensors transform under $\mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$ dimensional representations of SU(3). Hence we can write the tensor product of vector spaces as the direct sum of invariant subspaces $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$. These subspaces transform under irreducible representations of SU(3). Note that there are two $\mathbf{8}$ dimensional subspaces, however due to the closeness in energies and quantum numbers, the physical octet baryon states are formed as superpositions of the corresponding pairs in the two baryon octets. This results in one observable baryon octet. For further discussion refer to Ref. [24].

In the SU(3) limit, the baryons that correspond to each invariant subspace will follow the same mass trajectory. Interestingly, the symmetry properties of the flavour wave functions separate the baryons into the different groups of hyperfine splittings. This is as expected as the total wave function must be antisymmetric.

2.4.1 Flavour Basis

The quark model originated from the spectrum displayed by the light mesons and baryons, and so it is not difficult to understand that the basis we use in the quark model also describes many of the bound states we find in the vacuum. This basis is called the flavour basis. For example, the proton and the neutron are spin 1/2 with flavour composition uud and udd respectively, while the pseudosclar meson spectrum can be written as $u\bar{d}$, $u\bar{s}$, $d\bar{s}$ for π^+, K^+ and K^0, with their antiparticles being conjugates of these states. Quark flavour describes these states well because flavour is conserved in both electromagnetic and strong interactions. If we include weak interactions, flavour is not necessarily conserved, but the states do approximate the flavour basis states due to the weakness of the weak force. In situations where the strong and electromagnetic forces mix the flavour diagonal states, the flavour basis will not describe the physical states, this is the case with π^0, η and η'.

We group particles by flavour because they appeared in the mass spectrum in that way. Perhaps not surprisingly then, as the breaking of the quark symmetry

is by flavour, states that transform under irreducible representations of SU(N) describe the mass spectrum well. For instance determining the invariant subspaces under SU(3) (flavour) provides the possible baryon states grouped by mass. Similarly SU(4) flavour invariant subspaces group particles by mass when the charm quark is included. This is not dependent on the strength of the SU(N) symmetry in the spectrum, it is simply a reflection of the fact that hadron states can be grouped and identified by flavour.

CHAPTER 3

Lattice Gauge Theory

In field theory we normally can not produce an exact solution to the interacting theory. We can use the solution to the free theory and an infinite series in the potential to produce an exact solution. However this 'perturbative' approach implicitly suggests that we can stop calculating terms at some point, truncate the expansion, and achieve a reasonable approximation to the true result. This is not always true. A prime example is the strong coupling theory of QCD at low energy scales.

Regularisation is a technique for changing integrals which give infinite answers into integrals which give finite answers. Regularisation becomes essential when doing field theory because the integrals we try to calculate do not always result in finite answers. These infinities occur because we assume things about the field theory which may not be true. Often we assume the theory is correct up to infinite momentum and over infinite space-time; at the very least this assumption has not been proved by experiment. Alternatively, we may assume all quantities must be finite, however this is only true for observable quantities. In the worst case, the appearance of infinities indicate that our theory may not be valid at some energy scale. When infinities arise, we need to apply a cut-off to the momentum and region we integrate over. When cut-offs are introduced the quantities we produce will, in general, be cut-off dependent as a result.

Certain theories, called renormalisable theories, can be made to always produce a cut-off independent result for observables, and the quantum field theories we will be working with, QCD and QED, are both renormalisable. The process of renormalisation is not unlike setting a boundary condition on an integral. In this case certain quantities are fixed by experimental values, and we enforce these conditions on the quantum field theory using Lagrangian parameters called counter

terms. In the limit of infinite cut-off energy the quantities calculated should remain finite [20]. It is necessary to renormalise if we wish to use physical (observable) parameters such as the mass or charge within our Lagrangian, and to remove cut-off dependence.

All of these issues can be effectively mitigated using lattice gauge theory. This formulation of a gauge field theory discretises space-time and only considers a finite (discrete) slice of space-time, hence the integrals are regulated. In lattice gauge theory, when we set the scale, we renormalise the theory. The quantities calculated using lattice gauge theory are dimensionless and the scale for the calculation can be set by fixing one, or a few, output values to a physical quantity. Strong coupling theories become tractable because lattice gauge theory is not a perturbative method. At no point do we assume the free theory closely approximates the true theory, and the full path integral is calculated, though only at a finite number of points. Non-abelian gauge theories are tractable, though add additional complication as the topology of the gauge field in space-time is non-trivial.

In section 3.1 I explain why discretising space-time is a reasonable thing to do and introduce the action that is used for QCD and QED. In section 3.2 I explain how changing the action to be in terms of Euclidean space-time allows us to prioritise paths based on importance. This form is required if we wish to use importance sampling techniques like Monte-Carlo. In section 3.3 I describe how space-time is discretised in lattice gauge theory.

In section 3.4 I determine the quark action in discrete form to produce Wilson fermions and introduce the notion of a link variable. The hopping parameter and clover improvement to the quark action are explained. In section 3.5 I develop the discretised form of the gauge field kinetic term $F^{\mu\nu}F_{\mu\nu}$.

In section 3.6 I explain how to calculate observables on the lattice using Monte-Carlo sampling. In section 3.7 I describe how to calculate hadron 2-point correlation functions. In section 3.8 I explain how masses can be extracted from the hadron 2-point correlation functions. Note in the following if facts are not referenced they have been obtained from standard text books on lattice QCD, such as [26, 27].

3.1 Lattice gauge theory

The path integral formalism and field theory are based on observations of the universe; particles and waves take all possible paths, some particles are identical and that when left undisturbed they satisfy the Klein Gordon (KG) equation. However these observations were made with finite resolution. To deduce these things, we did not look at all the continuously infinite points in a section of space-time, rather we sampled space-time within our own ability to resolve space-time.

The beauty of the continuum is that we can always look smaller, but as observers, ultimately we look at a finite set of points. Essentially we have been doing what lattice gauge theory sets out to do; sample continuous space-time at finite points to determine what is happening.

We need to sample space-time at a frequency above the highest modes we would expect to be relevant to the system, and this essentially regularises our theory, setting a momentum cut-off. Indeed, a momentum cut-off was implicit in our initial observations. We then renormalise by setting the scale of the lattice using known observables, with the knowledge that observables of a renormalisable theory are cut-off independent. However, because we are generating the fields from discretised space-time there is an error in our observation of the field and in the construction of the field.

When we generate the field theory from a finite set of points, there are artefacts of the discretisation. These artefacts can either be partially or completely removed, or they scale with the lattice spacing, allowing for an estimate within some precision. By choosing a lattice spacing, we are limiting the maximum energy, or mode frequency. This maximum energy needs to be higher than the physics of interest. We also only sample a finite number of points, which puts a lower bound on our allowed frequencies as well. This lower bound can again be related to the lattice spacing and introduces its own discretisation errors when generating fields.

We will implement lattice discretisation on the Yang-Mills theory with fermions as we will use this later for lattice QCD and QED. The action is,

$$S = -\frac{1}{4}\int d^4x F^{\mu\nu}F_{\mu\nu} + \int d^4x \bar{\psi}(i\gamma^\mu D_\mu - m_0)\psi, \tag{3.1}$$

with covariant derivative $D_\mu = \partial_\mu + ig_0 A_\mu$ and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + ig_0[A_\mu, A_\nu]$. The partition function is then,

$$Z = \int DAD\bar{\psi}D\psi e^{iS + i\int d^4x J^\mu A_\mu + i\int d^4x(\bar{\eta}\psi + \bar{\psi}\eta)}. \tag{3.2}$$

Refer to chapter C for detail on the source fields J^μ.

3.2 Euclidean Space-Time

In its current form, the partition function weights all paths equally and from summing all the paths we can determine the probability amplitude at a given point. From looking at the partition function it is not clear which paths will be the most important for obtaining an estimate of Z. With a little knowledge of classical systems we can say that the classical path is the path that minimises the action S, hence paths close to this minimum tend to be the most important for

producing an estimate. If instead the paths were weighted by importance, we could select a random ensemble weighted by this importance to achieve an estimate of Z. This technique is called importance sampling, and we consider it in more detail in section 3.6.

It turns out this form can be achieved, with weights e^{-S}, through a Wick rotation to imaginary time to produce a Euclidean metric. We can obtain this shift by factoring out i in the zeroth component of the Lorentz vectors $x^0 = -ix^4$ and $\partial_0 = +i\partial_4$. As a result we no longer require the notation of up and down indices. This also results in an overall factor of i in the action. We use γ matrices that satisfy $\{\gamma^E_\mu, \gamma^E_\nu\} = \delta_{\mu\nu}$ (where E stands for Euclidean), as we now need them to produce a Euclidean metric when pre-multiplying the Dirac equation by $i\hbar\partial\!\!\!/$. One prescription for doing this is $\gamma^E_4 = \gamma^0$ and $\gamma^E_i = -i\gamma^i$ for $i \in \{1, 2, 3\}$. We now have,

$$S = -i\frac{1}{4}\int d^4x F_{\mu\nu}F_{\mu\nu} - i\int d^4x \bar{\psi}(\gamma^E_\mu D_\mu + m_0)\psi = -i\int d^4x \mathcal{L}_{eucl} = -iS_{eucl} \tag{3.3}$$

and it is equivalent to our original action.

3.3 Discretisation

The discretisation of Euclidean space-time is achieved by the creation of a four-dimensional hypercubic lattice. Each node in this lattice has links to 8 other nodes, none of which can be directly connected to each other. As the lattice is hypercubic, each link length is of a fixed lattice spacing a, and hence the locations within this lattice can be described using a vector $n \in \mathbb{Z}^4$, and the coordinates are defined as $x_\mu = an_\mu$.

The lattice is set to be a finite size so that it is some connected subset of $\mathbb{Z}^4$. The node locations are associated with a field value, which is the value of the (fermion or quark) field at that location. The gauge fields however are associated with any two given nodes, and hence reside on the links between nodes.

The sides and edges of this lattice can have various boundary conditions, such as fixed, symmetric and anti-symmetric. Fixed boundary conditions specify that the node value at a point outside the lattice has a fixed value (usually zero), where as anti-symmetric boundary conditions specify that the node outside the boundary has the negative value of the node on the opposite face or edge, likewise with links.

When calculating field amplitudes on the lattice, integrals are replaced with

sums and derivatives are replaced with finite differences,

$$\int f(x)d^4x \to a^4 \sum_n f(x_n),$$
$$\partial^\mu f(x) \to \frac{1}{2a}(f(x+a^\mu) - f(x-a^\mu)).$$

3.4 Quark Action

Our N component fermion fields transform under the gauge group G as $\psi(n) \to G(n)\psi(n)$ and $\bar{\psi}(n) \to \bar{\psi}(n)G^{-1}(n)$ for lattice site n. The discretised version of the fermion part of the action can be written as,

$$\begin{aligned} S_F = &(M+4r)\sum_n \bar{\psi}(n)\psi(n) \\ &- \frac{1}{2}\sum_{n,\mu}[\bar{\psi}(n)(r-\gamma_\mu)\psi(n+\hat{\mu}) + \bar{\psi}(n+\hat{\mu})(r+\gamma_\mu)\psi(n)] \end{aligned} \tag{3.4}$$

with Wilson parameter r, and is called the free Wilson fermion action. It takes this form because of the fermion doubling problem and the need for a symmetric derivative, details can be found in standard text books.

Something new in this action is two fields at different points multiplied by each other. In continuum field theory the points are quite distinct and transform separately, and generally comparing two points does not make sense, except by derivative. However the need for it in the discrete version of field theory is apparent. To ensure the result of the multiplication is properly defined despite local gauge freedom, we must insist that we parallel transport the fermion field from one tangent space to another. To do this we add a term between them which transforms oppositely to the two fields called a link variable,

$$U(x,y) = e^{ig_0 \int_x^y dx_\mu A_\mu(x)} \tag{3.5}$$

where $U(x,y) \to G(x)U(x,y)G^{-1}(y)$. So we correct our fermion action to instead be,

$$\begin{aligned} S_F^{(W)} = &(M+4r)\sum_n \bar{\psi}(n)\psi(n) \\ &- \frac{1}{2}\sum_{n,\mu}[\bar{\psi}(n)(r-\gamma_\mu)U(n,n+\hat{\mu})\psi(n+\hat{\mu}) + \bar{\psi}(n+\hat{\mu})(r+\gamma_\mu)U^\dagger(n,n+\hat{\mu})\psi(n)] \end{aligned} \tag{3.6}$$

which is called the Wilson action. In the continuum limit as the lattice spacing $a \to 0$, these link variables $U(n, n+\hat{\mu}) = 1 + ig_0 a A_\mu(n) + \mathcal{O}(a^2)$, and so we reproduce our covariant derivative D_μ.

For convenience the fermion action is typically rewritten in terms of the hopping parameter κ,

$$S_F^{(W)} = \frac{1}{2\kappa} \sum_{n,m} \bar{\psi}(n) K_{nm}[U] \psi(m), \tag{3.7}$$

where $\kappa = \frac{1}{8r+2M_0}$. The fermion matrix $K_{nm}[U]$ is defined by our previous expression for $S_F^{(W)}$, however it can be written explicitly as,

$$K_{nm}[U] = \delta_{nm}\mathbb{I} - \kappa \sum_{\mu>0} [(r-\gamma_\mu) U_\mu(n) \delta_{n+\hat{\mu},m} + (r+\gamma_\mu) U_\mu^\dagger(n-\hat{\mu}) \delta_{n-\hat{\mu},m}]. \tag{3.8}$$

If we change variables to $\psi \to \sqrt{2\kappa}\psi$, the correlation functions are invariant under the change. The path integral now yields the original correlation functions multiplied by $(1/2\kappa)^N$. The hopping parameter κ can be used to set the mass of the quark fields. When $r = 1$, the quark mass is defined as,

$$m_q = \frac{1}{2}\left(\frac{1}{\kappa_q} - \frac{1}{\kappa_c}\right) \tag{3.9}$$

where κ_q is the value of κ assigned to a quark q and κ_c is the value of κ when the flavour-neutral pseudoscalar meson mass is zero $M_{q\bar{q}} = 0$.

The action given by Eq. (3.6) is only correct to $\mathcal{O}(a)$, however it can be improved systematically. In this work we use clover improvement, which removes discretisation errors in the derivative to $\mathcal{O}(a^2)$. The improved action is,

$$S_F^{(C)} = S_F^{(W)} + c_{sw} a^5 \sum_{n\in\Lambda} \sum_{\mu<\nu} \bar{\psi}(n) \frac{1}{2} \sigma_{\mu\nu} \hat{F}_{\mu\nu}(n) \psi(n), \tag{3.10}$$

where the coefficient c_{sw} is referred to as the Sheikholeslami-Wohlert (or clover) coefficient [28]. It is calculated non-perturbatively [29]. Though not unique, a convenient choice for $\hat{F}_{\mu\nu}$ is,

$$\hat{F}_{\mu\nu}(n) = \frac{-i}{8a^2}[Q_{\mu\nu}(n) - Q_{\nu\mu}(n)], \tag{3.11}$$

where $Q_{\mu\nu}(n)$ is the sum of plaquettes $U_{\mu\nu}(n)$ (defined in Eq. (3.12), often referred to as the clover term. The clover action employed in this work is called SLiNC and is a small modification where U in Eq. (3.10) are replaced with stout-smeared gauge links [30]. Details can be found in Ref. [31].

3.5 Gluon Action

Our goal now is to find a discrete analogy to $F_{\mu\nu}$ which in the continuum limit results in the Yang-Mills field strength tensor. To this end the plaquette is defined as,

$$U_{\mu\nu}(n) = U(n, n+\hat{\mu})U(n+\hat{\mu}, n+\hat{\mu}+\hat{\nu})U^{\dagger}(n+\hat{\mu}, n+\hat{\mu}+\hat{\nu})U^{\dagger}(n, n+\hat{\nu}). \tag{3.12}$$

for $\mu, \nu \in \{1, ..., 4\}$, such that $\mu \neq \nu$. By using the identity $e^A e^B = e^{A+B+\frac{1}{2}[A,B]+...}$ for matrices A,B, we find

$$U_{\mu\nu}(n) = e^{ia^2 F_{\mu\nu}+\mathcal{O}(a^3)} \tag{3.13}$$

with $F_{\mu\nu} = \frac{1}{a}[(A_\nu(n+\hat{\mu}) - A_\nu(n)) - (A_\mu(n+\hat{\nu}) - A_\mu(n))] + ig_0[A_\nu(n), A_\mu(n)]$ and

$$\begin{aligned} Re\ \mathrm{tr}[e^{ia^2 F_{\mu\nu}+\mathcal{O}(a^3)}] &= Re\ \mathrm{tr}[1 + ia^2 F_{\mu\nu} - \frac{1}{2}a^4 F_{\mu\nu}F_{\mu\nu} + \mathcal{O}(a^5)] \\ &= \mathrm{tr}[1] - \frac{1}{2}a^4 \mathrm{tr}\,[F_{\mu\nu}F_{\mu\nu}] + \mathcal{O}(a^6). \end{aligned} \tag{3.14}$$

So we define the discretised Yang-Mills action as,

$$S_G^{(W)} = \frac{1}{g_0^2}\sum_{n,\mu>\nu} \mathrm{tr}[1 - \frac{1}{2}(U_{\mu\nu}(n) + U_{\mu\nu}^{\dagger}(n))], \tag{3.15}$$

where we have summed over all plaquettes on the lattice. This is called the Wilson gauge action.

As seen from Eq. (3.14), this action is only correct to $\mathcal{O}(a^2)$. The continuum correction to the lattice result can be systematically reduced by including additional terms in the action which converge to zero as we take the limit of zero lattice spacing. The improvement method used in this research is called Symanzik improvement. This method includes the use of larger rectangles of Wilson loops, of dimension 1×2 to reduce the error in the action and details can be found in Ref. [32].

3.6 Calculating Path Integrals on the Lattice

The full discretised partition function is now,

$$Z(\eta,\bar{\eta}) = \int DUD\psi D\bar{\psi}\ \exp\left(-S_{eucl} - a^4\sum_n\ [\bar{\eta}_n\psi_n + \bar{\psi}_n\eta_n]\right) \tag{3.16}$$

$$= \int DUD\psi D\bar{\psi}\ \exp\left(-S_F - a^4\sum_n\ [\bar{\eta}_n\psi_n + \bar{\psi}_n\eta_n]\right)\exp(-S_G) \tag{3.17}$$

$$= \int DU\ \det(K)\exp\left(-S_G - S_\eta\right), \tag{3.18}$$

where K is the fermion matrix defined in Eq. (3.8), the determinant was obtained using Grassmann field identities, $S_\eta = a^8\sum_m\sum_n\ \bar{\eta}(n)K^{-1}_{nm}\eta(m)$ and $\int DU$ is the path integral over the gauge links. Note here the short hand, $\eta_n = \eta(n)$ and $\psi_n = \psi(n)$. When doing field theory we use vacuum expectation values to determine certain observable quantities, which we will talk more about in section 3.7 and section 3.8. These vacuum expectation values take the form,

$$\left\langle\Omega\left|\mathcal{O}[U,\psi,\bar{\psi}]\right|\Omega\right\rangle = \frac{1}{Z}\int DUD\psi D\bar{\psi}\ \mathcal{O}[U,\psi,\bar{\psi}]\exp(-S_{eucl}). \tag{3.19}$$

The action can be combined by defining, $S'_G = \exp(-S_G + \log[\det(K[U])])$. Hence we must solve,

$$\left\langle\Omega\left|\mathcal{O}[U,\psi,\bar{\psi}]\right|\Omega\right\rangle = \frac{1}{Z}\int DU\ \mathcal{O}\left[U,\frac{\partial}{\partial\eta},\frac{\partial}{\partial\bar{\eta}}\right]\exp\left(-S'_G[U] - S_\eta\right). \tag{3.20}$$

However this is just the weighted average of the operator $\mathcal{O}$. If we can produce vacuum configurations of the gauge fields which are distributed with probability $P \propto \exp(-S_{eucl})$, then the weighted average would be just the average of the operator evaluated over this set of gauge fields,

$$\left\langle\Omega\left|\mathcal{O}[U,\psi,\bar{\psi}]\right|\Omega\right\rangle \approx \frac{1}{N}\sum_{n=1}^{N}\ \mathcal{O}[U_n], \tag{3.21}$$

for gauge configuration n. This technique is called Monte-Carlo sampling.

The simplest way to explain the Monte-Carlo sampling is to use an example. Suppose we wish to calculate the integral $I = \int_0^1 f(x)dx$, where $f(x) = x^2$. The domain of the function $f(x)$ is defined as $[0,1]$, an so the range of the function is $[0,1]$. Choose a point randomly in the interval $x \times y : [0,1]\times[0,1]$ and call it (x_1, y_1), where the total area is $A_{total} = 1$. Determine if y_1 is greater than or less

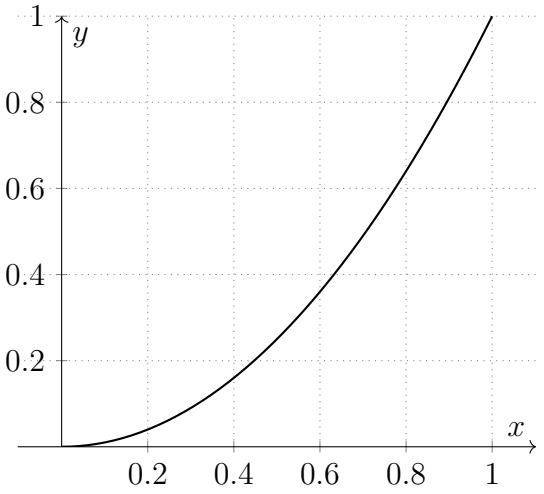


Figure 3.1: Function $f(x) = x^2$ on domain $[0, 1]$.

than $f(x_1)$. If it is less than $f(x_1)$ add one to a count of points N_{below}, otherwise do nothing. If we repeat this N times, the estimate of the area under $f(x) = x^2$ is given by $I \approx \frac{N_{below}}{N} A_{total}$. The estimate gets better the larger the number of samples N. This technique works because the probability for having a point occur below the function value is given by $P = A_{below}/A_{total}$ for areas A. By choosing our samples randomly in $[0, 1] \times [0, 1]$, they are are distributed (below or not below) with that probability, so $\frac{N_{below}}{N} \to P$ as $N \to \infty$.

Similarly, we could have applied the non-weighted tactic to the vacuum expectation value, by choosing a random gauge field and integrating over it by applying Eq. (3.20) directly. In this case the gauge fields are chosen with equal probability. This is still Monte-Carlo, however it is not really importance sampling, all values are equally weighted in importance. The convergence of unweighted Monte-Carlo can be very slow, and it is very slow when doing lattice gauge theory.

In the integral example above ($I = \int_a^b dx\ x^2$) each x sample was weighted with equal probability $1/(b - a)$. Consider instead, if we separated the integral into two parts $\int_a^b dx\ x \cdot x$. Choose the x location of our random sample with the probability distribution $P(x) = \frac{x}{C}$ where $C = \int_a^b dx\ x$ and call it x_1, and choose the y component randomly between 0 and 1 and call in y_1. Notice, $I = C \int_a^b dx\ x \cdot P(x)$. Add a count to N_{below} if the value of y_1 is below $g(x_1) = x_1$, otherwise do nothing. The now weighted average of the samples produces the estimate for the integral $\int_a^b dx\ x^2$ as in the first example.

To apply importance sampling to lattice gauge theory, we have taken advantage of the Euclidean time formulation of the theory. When calculating quantities in lattice gauge theory, the sample is the vacuum expectation value of the operator

calculated on a gauge configuration. The gauge configurations are distributed $P \propto \exp(-S_{eucl})$ hence so are the samples. An estimate of the vacuum expectation is just the average of all samples. Note that lower values of S are preferred in the distribution, and hence any algorithm that implements Monte-Carlo while weighting the samples with $P \propto \exp(-S_{eucl})$ is applying importance sampling.

Quark two point functions are often part of the calculation of operator expectation values on the lattice, so for demonstration purposes, we will show how Monte-Carlo sampling can be applied to calculate the expectation value of these two point functions. The operator takes the form, $\mathcal{O} = \mathrm{T}\{\bar{\psi}(x)\psi(y)\}$ for quark fields ψ, by applying the prescription for the operator in Eq. (3.20), $\mathcal{O} = \mathrm{T}\left\{\frac{\partial}{\partial\bar{\eta}(y)}\frac{\partial}{\partial\eta(x)}\right\}$. The solution to the equation is obtained when the generating fields are set to zero,

$$\left\langle \Omega \left| \mathrm{T}\{\bar{\psi}(x)\psi(y)\} \right| \Omega \right\rangle = \frac{1}{Z}\int DU\ K^{-1}(y-x)\exp(-S'_G) \tag{3.22}$$

$$\approx \frac{1}{N}\sum_{n=1}^{N} K^{-1}(y-x)[U_n] \tag{3.23}$$

where $K^{-1}[U_n]$ is the quark fermion propagator calculated on gauge configuration n. Note that this particular quantity $\left\langle \Omega \left| \mathrm{T}\{\bar{\psi}(x)\psi(y)\} \right| \Omega \right\rangle$ is zero when calculated on gauge configurations that do not have a fixed gauge. The expectation value of quark propagators (by themselves) are not considered in this work, however many quantities that are calculated on the lattice require the calculation of quark propagators, which we explore in the next section, section 3.7.

3.7 Hadron Operators

The masses of hadrons are extracted from the large Euclidean time behaviour of the correlation functions for zero momentum operators. The correlation function $C(t) = \langle P \rangle$ for a given operator P is the amplitude for the state P to exist for time t. If, for instance, the initial state is created at time t_i and the state is destroyed at time t_f then $t = t_f - t_i$. In terms of creation and annihilation of mesons for instance, $P = O_{ab}(t_f){O^\dagger}_{ab}(t_i)$ where $O_{ab}(\tau) = \bar{q}_a(\tau)\Gamma q_b(\tau)$ for quark flavours a, b.

In quantum field theory the hadronic 2-point correlation function is written as,

$$C(\vec{p},t) = \sum_{\vec{x},\vec{y}} e^{-i\vec{p}\cdot(\vec{y}-\vec{x})} \left\langle \Omega \left| \mathrm{T}\{O(y)O^\dagger(x)\} \right| \Omega \right\rangle \tag{3.24}$$

$$= \sum_{\vec{x},\vec{y}} e^{-i\vec{p}\cdot(\vec{y}-\vec{x})} \lim_{T\to\infty} \frac{\int D\bar{\psi}D\psi DA\ O(y)O^\dagger(x)\ \exp\left[-\int_{-T}^{T} d^4x \mathcal{L}_{eucl}\right]}{\int D\bar{\psi}D\psi DA\ \exp\left[-\int_{-T}^{T} d^4x \mathcal{L}_{eucl}\right]}. \tag{3.25}$$

When determining hadron masses, we first write down correlation functions in terms of hadron creation and annihilation operators called interpolators. Interpolators employed in this work are composed of quark fields and Γ matrices. The interpolating field has some overlap with many hadrons which share the same quantum numbers. Something like a hadron propagator is produced by contracting creation and annihilation interpolators. Contracting an interpolator involves Wick contracting the constituent quark fields. This produces a correlation function which is written in terms of Γ matrices and quark propagators. We provide an example of this below considering the π^+.

We wish to construct a correlation function that will produce the hadron mass we wish to study, so we need to know how to construct a relevant interpolator. An initial requirement for a hadron interpolator is gauge invariance, the interpolator must be a colour singlet. To match onto a particular state we desire that the interpolator must have the correct spin, charge conjugation and parity properties J^{PC}.[1]It also needs the correct quark structure $\bar{q}q$ for mesons and qqq for baryons.

The interpolators correspond to hadrons of the corresponding spin, charge and party J^{PC}. Consider, for example SU(3) flavour multiplets of the pseudoscalar mesons, vector mesons, octet baryons and decuplet baryons. For mesons, $O(x) = \bar{q}_1(x)\Gamma q_2(x)$, with quark fields q_1, q_2. A flavour can be associated with a quark field with certain quark mass and charge, which allows us to associate an interpolator with each element in a hadron multiplet. For instance, we can create a independent set of interpolators for the eight pseudoscalar mesons. For π^+ the correlation function is,

$$C(\vec{p},t) = \sum_{\vec{x},\vec{y}} e^{-i\vec{p}(\vec{y}-\vec{x})} \left\langle \Omega \left| \mathrm{T}\{O_{\pi^+}(y)O^{\dagger}_{\pi^+}(x)\} \right| \Omega \right\rangle \tag{3.26}$$

$$= -\sum_{\vec{x},\vec{y}} e^{-i\vec{p}(\vec{y}-\vec{x})} \left\langle \Omega \left| \mathrm{T}\{\bar{d}\gamma_5 u(y)\bar{u}\gamma_5 d(x)\} \right| \Omega \right\rangle . \tag{3.27}$$

By performing Wick contractions between quark fields we produce quark propagators. The contractions only occur between quarks of the same flavour,

$$C(\vec{p},t) \approx \frac{1}{N}\sum_{n=1}^{N}\sum_{\vec{x},\vec{y}} e^{-i\vec{p}(\vec{y}-\vec{x})} - \mathrm{tr}[\gamma_5 K_u^{-1}(x-y)\gamma_5 K_d^{-1}(y-x)][U_n]. \tag{3.28}$$

K_u^{-1}, K_d^{-1} are the fermion propagators of quark flavours u, d respectively, and U_n is the gauge configuration. The Γ matrices which correspond to each type of meson are show in Tab. 3.1.

[1]Note spin is not a well defined quantity on the lattice due to lack of spherical symmetry, however the wave functions of the different spin states have certain symmetries, and emerge as definite spin states in the continuum limit.

To calculate this correlation function on the lattice, we calculate the propagators on each lattice configuration U_n. The correlator from each configuration is summed to produce an estimate for the correlation function, as shown in Eq. (3.28).

Baryon interpolators are of the form

$$O(x) = \epsilon^{abc}(q_1^{aT}(x)C\Gamma_1 q_2^b(x))\Gamma_2 q_3^c(x) \tag{3.29}$$

$$\bar{O}(x) = \epsilon^{abc}(\bar{q}_1^a(x)C\bar{\Gamma}_1 \bar{q}_2^{bT}(x))\bar{q}_3^c(x)\bar{\Gamma}_2, \tag{3.30}$$

here q_i are quark fields, C is the charge conjugation operator and $\bar{\Gamma} = \gamma_0\Gamma^\dagger\gamma_0$. The fields within the brackets are paired and if $(\Gamma_1, \Gamma_2) = (\gamma_5, \mathbb{I})$ then this component of the interpolator has spin-0, while the remainder of the interpolator has spin-$\frac{1}{2}$. This choice of Γ matrices produces interpolators for octet baryons. To produce interpolators for spin-$\frac{3}{2}$ baryons, we instead choose $(\Gamma_1, \Gamma_2) = (\gamma_i, \mathbb{I})$ so that the part in brackets has spin-1, giving a total spin of $\frac{3}{2}$. There is asymmetry in this spin-$\frac{3}{2}$ interpolator as two of the quarks are paired, this is overcome by using three operators added together,

$$\begin{aligned} O(x) = \frac{1}{\sqrt{3}} \big[&\epsilon^{abc}(q_1^{aT}(x)C\gamma_i q_2^b(x))q_3^c(x) \\ &+ \epsilon^{abc}(q_3^{aT}(x)C\gamma_i q_2^b(x))q_1^c(x) \\ &+\epsilon^{abc}(q_3^{aT}(x)C\gamma_i q_1^b(x))q_2^c(x)\big] \end{aligned} \tag{3.31}$$

3.8 Extracting Hadron Mass

The lowest energy hadron which has quantum numbers associated with the interpolator is extracted at the large Euclidean time limit of the correlation function. Consider,

$$C(\vec{p}, t) = \sum_{\vec{x},\vec{y}} e^{-i\vec{p}(\vec{y}-\vec{x})} \left\langle \Omega \left| \mathrm{T}\{O(y)\bar{O}(x)\} \right| \Omega \right\rangle \tag{3.32}$$

we can insert a complete set of energy states [2] $\mathbb{I} = \sum_{\alpha,p,s} |\alpha, p, s\rangle \langle \alpha, p, s|$, with hadron label α, momentum p and spin (helicity on the lattice) s ,

$$C(\vec{p}, t) = \sum_{\vec{x},\vec{y}} \sum_{\alpha,p',s} e^{-i\vec{p}(\vec{y}-\vec{x})} \left\langle \Omega \left| O(y) \right| \alpha, p', s \right\rangle \left\langle \alpha, p', s \left| \bar{O}(x) \right| \Omega \right\rangle, \tag{3.33}$$

α represents all eigenstates with the correct quantum numbers for the interpolator, this includes excited states and multi-hadron states. The operator can be

[2] The set of states is finite instead of uncountably infinite because the scattering states have quantised momentum inside the lattice.

State	J^{PC}	Γ	Particles
Scalar	0^{++}	$\mathbb{I}$, γ_4	$f_0, a_0, K_0^*, ...$
Pseudoscalar	0^{-+}	γ_5, $\gamma_4\gamma_5$	$\pi^{\pm}, \pi^0, \eta, K^{\pm}, K^0, ...$
Vector	1^{--}	γ_i, $\gamma_4\gamma_i$	$\rho^{\pm}, \rho^0, \omega, K^*, \phi, ...$
Axial vector	1^{++}	$\gamma_i\gamma_5$	$a_1, f_1, ...$
Axial vector	1^{+-}	$\gamma_i\gamma_j$	$h_1, b_1, ...$

Table 3.1: A list of Γ matrices used in hadron interpolators.

decomposed using the translation operator,

$$O(x) = e^{Ht_f} e^{-i\vec{P}\cdot\vec{x}} O(0) e^{-Ht_f} e^{+i\vec{P}\cdot\vec{x}}. \tag{3.34}$$

Inserting a complete set of states into the correlation function, and noting that the vacuum has zero energy and momentum,

$$\begin{aligned} C(\vec{p},t) &= \sum_{\vec{x},\vec{y}} \sum_{\alpha,p',s} e^{-i\vec{p}(\vec{y}-\vec{x})} \left\langle \Omega \left| e^{Ht_f} e^{-i\vec{P}\cdot\vec{y}} O(0) e^{-Ht_f} e^{+i\vec{P}\cdot\vec{y}} \right| \alpha, p', s \right\rangle \\ &\quad \left\langle \alpha, \vec{p}', s \left| e^{-i\vec{P}\cdot\vec{x}} e^{Ht_i} \bar{O}(0) e^{-H(-t_i)} e^{+i\vec{P}\cdot\vec{x}} \right| \Omega \right\rangle \\ &= \sum_{\vec{x},\vec{y}} \sum_{\alpha,p',s} e^{-i(\vec{p}-\vec{p'})(\vec{y}-\vec{x})} \left\langle \Omega \left| O(0) \right| \alpha, p', s \right\rangle e^{-E_\alpha(\vec{p'})(t_f-t_i)} \left\langle \alpha, p', s \left| \bar{O}(0) \right| \Omega \right\rangle \\ &= \sum_{\alpha,s} e^{-E_\alpha(\vec{p})(t_f-t_i)} \left\langle \Omega \left| O(0) \right| \alpha, p, s \right\rangle \left\langle \alpha, p, s \left| \bar{O}(0) \right| \Omega \right\rangle \end{aligned} \tag{3.35}$$

To achieve the last line $\sum_{\vec{x},\vec{y}} e^{i(\vec{p}-\vec{p'})(\vec{y}-\vec{x})} = \delta(\vec{p}-\vec{p'})$. In the limit of large time $(t_f - t_i) \to \infty$,

$$C(\vec{p},t) = \lim_{t_f - t_i \to \infty} e^{-E_\alpha(\vec{p})(t_f-t_i)} \left\langle \Omega \left| O(0) \right| \alpha, p, s \right\rangle \left\langle \alpha, p, s \left| \bar{O}(0) \right| \Omega \right\rangle \tag{3.36}$$

The hadron mass is obtained for the case $\vec{p} = 0$ and α now represents the lowest energy state.

When doing calculations on the lattice we create a state at a point, but this state can propagate both forward and backward in time; the particle correlates with points that are forward in time, but also with points that are at earlier times. The reverse time propagating state is the state obtained after applying the time reversal operator τ. The state obtained for mesons is the exact anti-particle, while for baryons it is the anti-particle but with wrong (opposite) parity, which has a different ground state energy than the forward propagating particle with correct parity.

As the spatial and temporal extent of our lattice is finite, the correlation amplitude of the forward propagating states can coincide with the correlation amplitude

of the reverse propagating states. The degree to which this effects the energy state we are trying to measure is dependent on how quickly the correlation function decays below numerical noise.

For meson states we expect to have both a forward and backward propagating state contributing at a single time point. With anti-periodic boundary conditions we expect the form $e^{-E_0(\vec{p})(t-t_i)} + \tau_{\Gamma_1}\tau_{\Gamma_2}e^{-E_0(\vec{p})(T-(t-t_i))}$ for mesons, where T is the lattice extent in the time direction and τ_Γ are ± 1 dependent on the Γ factors inside the interpolators. For baryons the effect is very minor because the state has higher energy, and so the correlation function decays below noise quickly.

Effective mass is defined in terms of correlation function $C(t)$ as,

$$M_{eff}(t) = -\log\left(\frac{C(t+1)}{C(t)}\right), \tag{3.37}$$

which obtains the ground state energy as a plateau when the correlation amplitudes of higher energy states decay below noise. Fits to the effective mass are obtained by fitting the correlation function with the expected form for each type of particle (meson or baryon) in a region which is considered to be free of excited state contamination (at the point of plateau). The parameters of the fit are used to determine the effective mass.

CHAPTER 4

Extrapolating to the Physical Point

Though lattice methods have improved and computers are much faster than 30 years ago, there are often limitations on how large the lattice can be and how close to the physical point the quark masses are; which means after a lattice calculation some extrapolation of observables to the physical point may be required. To do this extrapolation I describe an observable in terms of a Taylor expansion. The observable we are concerned with in this thesis is hadron mass and we create a Taylor expansion in terms of quark mass and charge. To create the Taylor expansions we need to understand which properties of the quarks are important, and also enforce symmetries to make fitting easier and extrapolation more accurate.

Quarks play two different roles in the mass of a hadron. Hadrons are formed from bound states of quarks, interacting (primarily) via QCD. Quarks also play a role in renormalising gluons, the gauge bosons of QCD. We differentiate between these quarks, calling quarks inside the hadron valence quarks and quarks that renormalise the gluons sea quarks. The valence quarks of a hadron correspond to the quark model prescription, for instance π^+ is made of $u\bar{d}$ quarks, and only these flavours of valence quark contribute to the mass of the π^+. However, all sea quark flavours contribute to the mass of a hadron as the charge that gluons see is colour; QCD can not tell the difference between different flavours. Note also, QCD can not act on the flavours of the quarks, which is why flavour is preserved by QCD. A similar argument can be made for QED; the electric charges of the valence quarks contribute to the mass of the particle, but the charges of all the sea quark flavours contribute to the hadron mass. Even though at the physical point the valence and sea quarks are the same particles, their action on hadron mass is different and their symmetries are different and so we will treat them separately.

The point of expansion for the Taylor series is taken to be an SU(3) symmetric

point, which is a point where all quarks have equal mass and equal charge. We choose an SU(3) symmetric point because the light flavours (u, d, s) in the QCD Lagrangian have an approximate SU(3) flavour symmetry, which is evident in the hadron spectrum and a key component in the discovery of QCD as we discussed previously. If we can choose the SU(3) flavour symmetry point closest to the physical configuration then the SU(3) flavour symmetry makes such an expansion highly effective at producing reliable predictions. This is because we know the extrapolation distance is small.

This closest SU(3) point is given by the conditions $\bar{m} = (m_u^* + m_d^* + m_s^*)/3$ and $m_u = m_d = m_s = \bar{m}$, where $\bar{m}$ is the average of the physical quark masses[1]. Using these definitions, the physical point point lies along the line $\delta m_u + \delta m_d + \delta m_s = 0$ from the symmetric point, where here $\delta m_q = m_q - \bar{m}$. A simplified path to this point with $m_u = m_d = m_l$ is shown in Fig. 4.1 and a diagram showing masses of all three flavours is shown in Fig. 4.2.

Another advantage of choosing this point is that the sum of the quark masses remains the same along this path. Hence singlet quantities, which are flavour-neutral objects, remain unchanged to first order along the path. An example of a flavour singlet is the sum of light quark masses $(m_u + m_d + m_s)$, another is the pseudoscalar meson singlet $X_\pi^2 = \frac{1}{6}(M_{K^+}^2 + M_{K^0}^2 + M_{\pi^+}^2 + M_{\pi^-}^2 + M_{\bar{K}^0}^2 + M_{K^-}^2)$. The invariance of singlets is due to the definition of $\delta m_u + \delta m_d + \delta m_s = 0$.

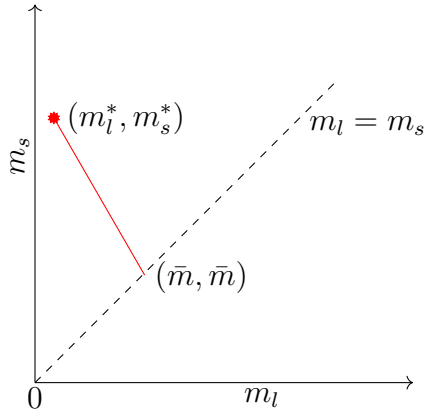


Figure 4.1: The transition from the SU(3) symmetric line to the physical point. The physical point is indicated by the red star. Extrapolation path is shown by the red line. SU(3) symmetric points are along the dotted line, and the SU(3) symmetric point we use is at the intersection of the red and dashed line. In this case we have made the simplification that $m_u = m_d = m_l$, to make the plot two dimensional.

[1]The * superscript is used to denote the physical value.

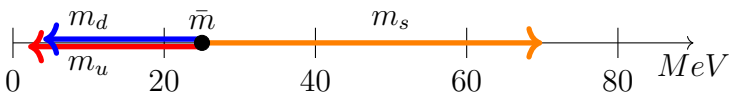


Figure 4.2: The transition from the SU(3) symmetric line to the physical point. The red, blue and orange vectors show the change in the up, down and strange quark masses (respectively) from the SU(3) symmetric point to the physical point. The length of the red plus the blue vector is the same as the length of the orange vector. The vectors sum to zero. Note the quark masses shown in the diagram are approximate and the values change with scale.

At the SU(3) flavour point all hadrons of an irreducible representation have the same mass. If we expand around an SU(3) point, the SU(3) symmetry is broken, however the masses of particles of the same irreducible representation still satisfy S_3 symmetry. We use the permutation symmetries S_3 of quark flavours around an SU(3) symmetric point to put constraints on the coefficients of the hadron mass expansion. We will show that the coefficients of an expansion about an SU(3) flavour symmetry point are the same for all hadrons that transform within an irreducible representation of SU(3) flavour.

In applying symmetries to the Taylor expansion, we reduce the number of coefficients. This allows us to more reliably obtain information from the lattice data by excluding noise in the lattice result (by excluding degrees of freedom). Additionally, extracting information from the lattice becomes easier as less coefficients need to be solved for.

In section 4.1 I define what a singlet quantity is and explain why it is invariant along the path $\delta m_u + \delta m_d + \delta m_s = 0$. I give several examples of hadron mass polynomials that are singlets. As any hadron mass can only be dependent on quark mass polynomials with the same symmetries, in section 4.2 I explore the transformation properties of quark mass polynomials. It is possible to show [33] that these polynomials can be produced to have exact S_3 symmetries, but in general mixed SU(3) symmetries. In section 4.3 I determine hadron mass polynomials that have exact SU(3) symmetries, and (using results from section 4.2) am able to determine an estimate of the magnitude of each order of δm_q in the octet and decuplet baryon quark mass expansions. The decrease in magnitude at each order of δm_q suggests these Taylor series converge.

In section 4.4 I use the hadron mass spectrum to show the valence quark terms in the Taylor expansion can only depend on the quark constituents of the hadron. We then determine the quark mass and charge expansion for pseudoscalar mesons, octet baryons and decuplet baryons respectively using the constraints obtained from the spectrum transforming under S_3. In section 4.5 I use permutation symmetry and QCD flavour blindness to show the sea quark contributions to the masses

occur in the Taylor expansion as flavour singlets. We then determine the sea quark mass and charge expansions using S_3 symmetry arguments, and produce the full expansion with both sea and valence quarks included. In section 4.6 I describe how to determine where the closest symmetric point is, which is the one we expand about in this work, in terms of quark mass and charge $(m_u, m_d, m_s, e_u, e_d, e_s)$. I explain how the analysis is improved by incorporating the Dashen scheme, and describe this scheme in detail.

4.1 Singlet Quantities

A flavour singlet is a quantity that is invariant under the S_3 permutation group acting on the light flavours $\{u, d, s\}$. One of the main advantages of remaining on the $\delta m_u + \delta m_d + \delta m_s = 0$ line to the physical point is that flavour singlet quantities, X_s, do not change much over this path. Consider X_s at the point $(\bar{m}, \bar{m}, \bar{m})$. As the flavours are indistinguishable in every way we expect that,

$$\frac{\partial X_s}{\partial m_u} = \frac{\partial X_s}{\partial m_d} = \frac{\partial X_s}{\partial m_s}. \tag{4.1}$$

Hence $X_s = M_0 + c_1(m_u + m_d + m_s) + \mathcal{O}(m^2)$. Along the line $\delta m_u + \delta m_d + \delta m_s = 0$,

$$\Delta X_s = \frac{\partial X_s}{\partial m}(\delta m_u + \delta m_d + \delta m_s) + \mathcal{O}(m^2) = 0 + \mathcal{O}(m^2). \tag{4.2}$$

It is possible to construct various approximate singlet (flavour S_3 invariant) quantities using hadron masses. Some singlet quantities constructed from hadron masses are shown in Tab. 4.1.

Pseudoscalar[2] meson	$X_\pi^2 = \frac{1}{6}(M_{K^+}^2 + M_{K^0}^2 + M_{\pi^+}^2 + M_{\pi^-}^2 + M_{\bar{K}^0}^2 + M_{K^-}^2)$
	$X_{\eta_8}^2 = \frac{1}{2}(M_{\pi^0}^2 + M_{\eta_8}^2)$
Vector meson	$X_\rho = \frac{1}{6}(M_{K^{*+}} + M_{K^{*+}} + M_{\rho^+} + M_{\rho^-} + M_{\bar{K}^{*0}} + M_{K^{*-}})$
	$X_{\phi_8} = \frac{1}{2}(M_{\rho^0} + M_{\phi_8})$
Octet baryon	$X_N = \frac{1}{6}(M_p + M_n + M_{\Sigma^+} + M_{\Sigma^-} + M_{\Xi^0} + M_{\Xi^-})$
	$X_\Lambda = \frac{1}{2}(M_\Lambda + M_{\Sigma^0})$
Decuplet baryon	$X_\Delta = \frac{1}{3}(M_{\Delta^{++}} + M_{\Delta^-} + M_{\Omega^-})$
	$X_{\Xi^*} = \frac{1}{6}(M_{\Delta^+} + M_{\Delta^0} + M_{\Sigma^{*+}} + M_{\Sigma^{*-}} + M_{\Xi^{*0}} + M_{\Xi^{*-}})$
	$X_{\Sigma^*} = M_{\Sigma^{*0}}$

Table 4.1: List of flavour singlet quantities composed of hadron masses.

[2]We typically analyse pseudoscalar mesons in terms of their mass squared because it is proportional to the quark mass.

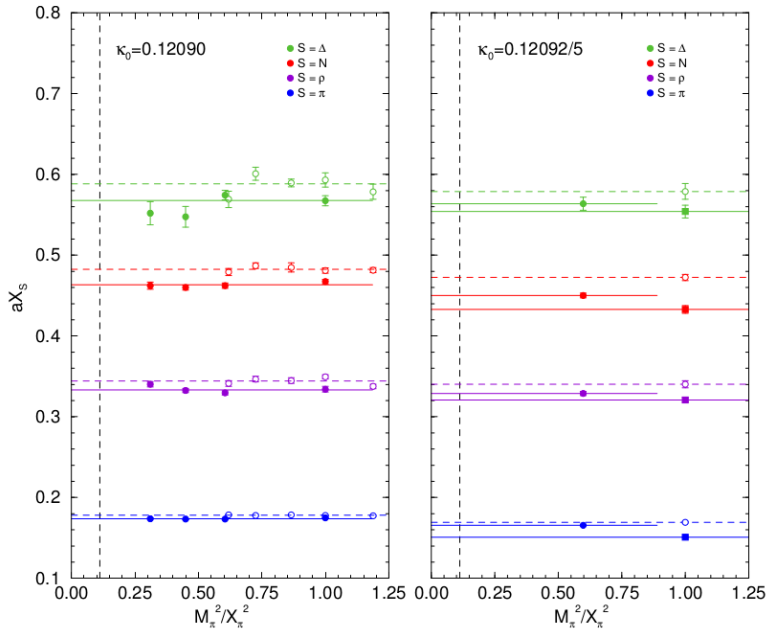


Figure 4.3: SU(3) flavour singlets as a function of pion mass. The dashed vertical line is at the physical quark mass. Filled points are from $32^3 \times 64$ lattice data while open points are from $24^3 \times 48$ lattice data. The horizontal lines are constant fits to the data points. This figure was taken from [33].

Ref. [33] showed using lattice QCD data that singlet quantities along the path $\delta m_u + \delta m_d + \delta m_s = 0$ from the symmetric point remain (within errors) constant, as shown in Fig. 4.3.

4.2 Transformation Properties of Quark Mass Polynomials

The Taylor expansion for a quantity with certain symmetries will only depend on polynomials with these same symmetries. Quark mass polynomials and similarly quark charge polynomials can be designated transformation properties under the flavour permutation group S_3 and SU(3) flavour. The full set of polynomial combinations at any order can always be arranged to have distinct transformation symmetries under S_3, however this is not true for transformation symmetries of SU(3) [33]. In this section we will determine these transformation properties, to later associate these terms with hadron masses.

Consider the quark mass parameters (m_u, m_d, m_s), for the purposes of allocating symmetries to polynomials, we work in the following basis. Our expansion

point is defined as (m_0, m_0, m_0), $\bar{m}_q = (m_u + m_d + m_s)/3$ (not the physical quark masses) and $\delta m_u, \delta m_d, \delta m_s$, where $\delta m_q = m_q - \bar{m}$. Note that $\delta m_u + \delta m_d + \delta m_s = 0$. The polynomials up to $\mathcal{O}(m^3)$ are shown in Tab. 4.2, and full details of how these are obtained are given in [33].

I will demonstrate this division into invariant subspaces for the first order, because it is easy to show without sophisticated group theory techniques. Within the QCD Lagrangian, the mass term is written as $\bar{q}_a M_{ab} q_b$ for light quarks q with flavours $a, b \in \{u, d, s\}$. Under SU(3) flavour symmetry, q is a vector space **3** which transforms under the fundamental representation of SU(3), while $\bar{q}$ transforms like the conjugate space. This means that M is an element of the product space $3 \otimes \bar{3}$. The product space has two invariant subspaces, one composed of fixed points and one composed of orbits that transform like the 8-dimensional representation of SU(3), $3 \otimes \bar{3} = 1 \oplus 8$. We can break M down into basis vectors of these spaces,

$$M = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix} \tag{4.3}$$

$$= \bar{m} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \left. \begin{pmatrix} \delta m_u & 0 & 0 \\ 0 & \delta m_d & 0 \\ 0 & 0 & \delta m_s \end{pmatrix} \right|_{\delta m_u + \delta m_d + \delta m_s = 0} \tag{4.4}$$

$$= \bar{m} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{2}(\delta m_u - \delta m_d) \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$+ \frac{1}{6}(-\delta m_u - \delta m_d + 2\delta m_s) \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{4.5}$$

As the group action on M is $UMU^\dagger$, any matrix proportional to the identity is invariant. The remaining matrices are traceless as a result, and any two matrices that are diagonal, orthogonal to each other and traceless will work as basis vectors for this invariant space. Note that the above can be reduced to the same polynomial as in the table by $2\delta m_s - \delta m_u - \delta m_d = 3\delta m_s$.

4.3 Transformation Properties of Hadron Mass Polynomials

Typically a hadron mass will have distinct permutation symmetries under quark flavours, but not under SU(3). Hence it is only permutation symmetries that are needed when generating these expansions. It is however possible to generate

Polynomial		S_3	SU(3)
1	✓	A_1	1
$(\bar{m}-m_0)$		A_1	1
(δm_s)	✓	E^+	8
$(\delta m_u-\delta m_d)$	✓	E^-	8
$(\bar{m}-m_0)^2$		A_1	1
$(\bar{m}-m_0)\delta m_s$		E^+	1
$(\bar{m}-m_0)(\delta m_u-\delta m_d)$		E^-	1
$\delta m_u^2+\delta m_d^2+\delta m_s^2$	✓	A_1	1, 27
$\delta m_s^2-(\delta m_u-\delta m_d)^2$	✓	E^+	8, 27
$\delta m_s(\delta m_u-\delta m_d)$	✓	E^-	8, 27
$(\bar{m}-m_0)^3$		A_1	1
$(\bar{m}-m_0)^2\delta m_s$		E^+	8
$(\bar{m}-m_0)^2(\delta m_u-\delta m_d)$		E^-	8
$(\bar{m}-m_0)(\delta m_u^2+\delta m_d^2+\delta m_s^2)$		A_1	1, 27
$(\bar{m}-m_0)[3\delta m_s^2-(\delta m_u-\delta m_d)^2]$		E^+	8, 27
$(\bar{m}-m_0)\delta m_s(\delta m_d-\delta m_u)$		E^-	8, 27
$\delta m_u\delta m_d\delta m_s$	✓	A_1	1, 27, 64
$\delta m_s(\delta m_u^2+\delta m_d^2+\delta m_s^2)$	✓	E^+	8,27,64
$(\delta m_u-\delta m_d)(\delta m_u^2+\delta m_d^2+\delta m_s^2)$	✓	E^-	8,27,64
$(\delta m_s-\delta m_u)(\delta m_s-\delta m_d)(\delta m_u-\delta m_d)$	✓	A_2	10, $\overline{10}$

Table 4.2: All quark-mass polynomials up to $\mathcal{O}(m_q^3)$, classified by symmetry properties. The tick (✓) indicates polynomials relevant for $\bar{m}=m_0$, which is along the path $\delta m_u+\delta m_d+\delta m_s=0$ from the physical point. This table was taken directly from [33].

hadron mass combinations which have distinct SU(3) transformation properties. These can be used to give an indication of how well the Taylor expansion is converging, as they can be related to a certain order of δm_q in the expansion. Ref. [33] showed that each additional order of δm_q was smaller than the previous order by about a factor of 10. In this section we will explain briefly the methodology of producing quantities of a pure irreducible representation of SU(3), and tabulate them for reference. We will show that the magnitude of these quantities decreases depending on the SU(3) transformation properties, and discuss the advantages of using our chosen path of $\delta m_u+\delta m_d+\delta m_s=0$.

If we consider the hadrons to be fields, then inside a hadron Lagrangian the mass term looks like $\bar{h}Mh$. The hadron mass matrix transforms like $h\bar{h}$ under SU(3) flavour symmetry. Considering the hadron multiplets, the decuplet mass

Δ^-	Δ^0	Δ^+	Δ^{++}	Σ^{*-}	Σ^{*0}	Σ^{*+}	Ξ^{*-}	Ξ^{*0}	Ω^-	S_3	SU(3)
1	1	1	1	1	1	1	1	1	1	A_1	1
-1	-1	-1	-1	0	0	0	1	1	2	E^+	8
-3	-1	1	3	-2	0	2	-1	1	0	E^-	8
3	-1	-1	3	-1	-3	-1	-1	-1	3	A_1	27
-3	7	7	-3	-5	0	-5	-2	-2	6	E^+	27
-3	-1	1	3	3	0	-3	4	-4	0	E^-	27
2	-3	-3	2	-3	12	-3	-3	-3	2	A_1	64
-1	0	0	-1	3	0	3	-3	-3	2	E^+	64
-1	2	-2	1	1	0	-1	-1	1	0	E^-	64
0	-1	1	0	1	0	-1	-1	1	0	A_1	64

Table 4.3: Decuplet mass matrix, decomposed into basis vectors which transform under irreducible representations. This table is taken from [33].

matrix transforms like $\mathbf{10} \otimes \overline{\mathbf{10}}$, pseudoscalar mesons like $\mathbf{8} \otimes \overline{\mathbf{8}}$ and the octet baryons like $\mathbf{8} \otimes \overline{\mathbf{8}}$. The mass matrix (when we ignore mixing) is diagonal, and we can form an orthogonal basis for the space of diagonal vectors. This basis can be chosen to be vectors that transform under a single irreducible representation of SU(3). These vectors form a basis for the invariant subspaces. The basis vectors for the decuplet baryons are shown in Tab. 4.3. The details of how they were decomposed can be found in [33], together with similar tables for the baryon octet and pseudoscalar meson.

If we input physical masses of the hadrons into the hadron mass polynomials shown in Tab. 4.3 we are able to determine how important each order in the expansion is given that all but the singlet are zero at an SU(3) point. We quote the result given by [33], simplified to exact isospin symmetry,

$$\begin{aligned} 4M_\Delta + 3M_{\Sigma^*} + 2M_{\Xi^*} + M_\Omega &= 13.82 \text{ GeV} && \in \mathbf{1} \propto \delta m^0 \\ -2M_\Delta + M_{\Xi^*} + M_\Omega &= 0.742 \text{ GeV} && \in \mathbf{8} \propto \delta m^1 \\ 4M_\Delta - 5M_{\Sigma^*} - 2M_{\Xi^*} + 3M_\Omega &= -0.044 \text{ GeV} && \in \mathbf{27} \propto \delta m^2 \\ -M_\Delta + 3M_{\Sigma^*} - 3M_{\Xi^*} + M_\Omega &= -0.006 \text{ GeV} && \in \mathbf{64} \propto \delta m^3. \end{aligned}$$

Along a path $\delta m_u + \delta m_d + \delta m_s = 0$ from the physical point, we would expect good convergence of the Taylor expansion for the decuplet baryons from the SU(3) point, as the importance of each order of SU(3) breaking parameter δm significantly decreases. It also indicates we would expect to encompass all effects up to a precision of approximately 1% provided the Taylor expansion is to second order.

Similar hadron mass polynomials can be constructed for the pseudoscalar mesons and octet baryons. We need to be a little more careful of which of these hadron mass polynomials we choose because there is mixing between the Σ^0 and Λ in the

octet baryons and there is mixing between π^0, η, η' in the pseudoscalar mesons. Furthermore, the flavour-neutral pseudoscalar mesons also have disconnected quark line contributions and all meson mass polynomials above the singlet require knowledge of the connected part of η. This means we cannot make a prediction on the convergence of the pseudoscalar Taylor expansion using the physical meson masses, so we do not include these mass polynomials here. The polynomials which help us predict the convergence of the octet baryon are shown below,

$$\begin{aligned}
M_n + M_p + M_{\Sigma^+} + M_{\Sigma^-} & \\
+M_{\Sigma^0} + M_\Lambda + M_{\Xi^-} + M_{\Xi^0} &= 9.209 \text{ GeV} && \in \mathbf{1} \propto \delta m^0 \\
-M_n - M_p + M_{\Xi^-} + M_{\Xi^0} &= 0.758 \text{ GeV} && \in \mathbf{8} \propto \delta m \\
M_n + M_p + M_{\Sigma^+} + M_{\Sigma^-} & \\
-3M_{\Sigma^0} - 3M_\Lambda + M_{\Xi^-} + M_{\Xi^0} &= -0.024 \text{ GeV} && \in \mathbf{27} \propto \delta m^2 \\
M_n - M_p + M_{\Sigma^+} & \\
-M_{\Sigma^-} + M_{\Xi^-} - M_{\Xi^0} &= 0.000071 \text{ GeV} && \in \mathbf{10} \propto \delta m^3.
\end{aligned}$$

Again we see that the magnitude of each additional order in δm is reduced significantly, which suggest we would see good convergence of the Taylor expansion. We would expect the Taylor polynomial to encompass all effects up to a precision of 0.01% for the octet baryons, if we include all polynomials up to second order.

4.4 Valence quark expansion about an SU(3) point

It is possible to generate Taylor expansions in terms of quark mass and charge using the hadron mass basis vectors shown in Tab. 4.3, by solving the linear system. However, it is more intuitive to apply constraints using permutation symmetries S_3 to the expansions directly, and it does not require more than basic group theory knowledge. It also has the benefit of generating the expansion to all orders in quark mass and charge.

We will assume the function that describes hadron masses is analytic and confine ourselves to QCD and mesons for the moment, with sea quark masses held fixed. Under these conditions, the mass of a non-diagonal flavour meson in pure QCD[3]is only a function of its valence quark masses,

$$M_{a\bar{b}} = M(\delta\mu_a, \delta\mu_b) = M_0^{ab} + \sum_{i=1}^{N} c_{ai}\delta\mu_a^i + \sum_{i=1}^{N} c_{bi}\delta\mu_b^i + \sum_{i=1}^{N}\sum_{j=1}^{N} \beta_{ij}\delta\mu_a^i\delta\mu_b^j \quad (4.6)$$

[3]'Pure QCD' here and elsewhere in the thesis means QCD without QED.

where a, b are the quark flavours and $\delta\mu_a$[4]is the change in valence quark mass for flavour a, $\delta\mu_a = m_a - m_0^a$. $M_{a\bar{b}}$ is the mass of a meson composed of quark of flavour a and anti-quark of flavour b. The limit to the sum N, can be as large as required to produce a desired accuracy. The expression is simply the definition of an analytic function of two variables $\delta\mu_a, \delta\mu_b$. If we now consider that we only have three quark flavours, u, d, s and choose the expansion point for valence quark masses $m_u = m_d = m_s = m_0$ so that $\delta\mu_a = m_a - m_0$, then the mass spectrum of meson states satisfies an S_3 symmetry.

When determining meson masses from a lattice simulation, this means that the masses satisfy $M_{a\bar{b}} = M_{c\bar{e}} = M(\delta\mu_a, \delta\mu_b)$ for $a, b, c, e \in \{u, d, s\}$ for all $\delta\mu_a = \delta\mu_c, \delta\mu_b = \delta\mu_e$. Hence we can simplify the expansion about this point. If the above is true, then $c_{ui} = c_{si} = c_{di}$, and $\beta_{ij}^{ab} = \beta_{ij}^{ce}$. To see this consider any two fixed displacements $\delta\mu_1, \delta\mu_2 \in \mathbb{R}$,

$$M_{a\bar{b}} = M_{ab}(\delta\mu_1, \delta\mu_2) = M_0^{ab} + \sum_{i=1}^{N} c_{ai}\delta\mu_1^i + \sum_{i=1}^{N} c_{bi}\delta\mu_2^i + \sum_{i=1}^{N}\sum_{j=1}^{N} \beta_{ij}^{ab}\delta\mu_1^i\delta\mu_2^j = \tag{4.7}$$

$$M_{c\bar{e}} = M_{ce}(\delta\mu_1, \delta\mu_2) = M_0^{ce} + \sum_{i=1}^{N} c_{ci}\delta\mu_1^i + \sum_{i=1}^{N} c_{ei}\delta\mu_2^i + \sum_{i=1}^{N}\sum_{j=1}^{N} \beta_{ij}^{ce}\delta\mu_1^i\delta\mu_2^j. \tag{4.8}$$

First by choosing $\delta\mu_1 = \delta\mu_2 = 0$, we see that the constant is shared by all polynomials. Choosing only $\delta\mu_2 = 0$ we can see $c_{ai} = c_{ci}$ but for $a, c \in \{u, d, s\}$ so $c_{ui} = c_{si} = c_{di}$. Setting $\delta\mu_1, \delta\mu_2$ to be arbitrary, it then follows that $\beta_{ij}^{ab} = \beta_{ij}^{ce}$, hence also $\beta_{ij}^{ab} = \beta_{ij}^{ba} = \beta_{ji}^{ab}$. So the new expansion is,

$$M(\delta\mu_a, \delta\mu_b) = M_0 + \sum_{i=1}^{N} c_i(\delta\mu_a^i + \delta\mu_b^i) + \sum_{i=1}^{N}\sum_{j=1}^{i} \beta_{ij}(\delta\mu_a^i\delta\mu_b^j + \delta\mu_b^i\delta\mu_a^j). \tag{4.9}$$

It is essential that the expansion be about a point $m_u = m_d = m_s$ or the quarks can not be considered indistinguishable and the spectrum will not satisfy S_3.

Using the same arguments we can produce an expression for spin 3/2 baryons.

[4] When the valence and sea quark masses are the same we simply use m while when the valence quark masses are different from the sea quarks, the sea quarks take the symbol m and valence quarks take the symbol μ.

Baryons are dependent on 3 quark masses and satisfy S_3 at a point $m_u = m_d = m_s$,

$$\begin{aligned} M(\delta\mu_a, \delta\mu_b, \delta\mu_c) &= M_0 + \sum_{i=1}^{N} \alpha_i(\delta\mu_a^i + \delta\mu_b^i + \delta\mu_c^i) \\ &+ \sum_{i=1}^{N}\sum_{j=1}^{i} \gamma_{ij}(\delta\mu_a^i\delta\mu_b^j + \delta\mu_a^j\delta\mu_b^i + \delta\mu_a^i\delta\mu_c^j + \delta\mu_a^j\delta\mu_c^i + \delta\mu_b^i\delta\mu_c^j + \delta\mu_b^j\delta\mu_c^i) \\ &+ \sum_{i=1}^{N}\sum_{j=1}^{i}\sum_{k=1}^{j} \gamma_{ijk}(\delta\mu_a^i\delta\mu_b^j\delta\mu_c^k + \delta\mu_a^i\delta\mu_b^k\delta\mu_c^j + \delta\mu_a^j\delta\mu_b^i\delta\mu_c^k \\ &+ \delta\mu_a^j\delta\mu_b^k\delta\mu_c^i + \delta\mu_a^k\delta\mu_b^i\delta\mu_c^j + \delta\mu_a^k\delta\mu_b^j\delta\mu_c^i). \end{aligned} \tag{4.10}$$

The spin 1/2 baryons are a little different. The masses satisfy an S_3 symmetry at a point $m_u = m_d = m_s$ but only in the two quarks that are paired by spin. For instance $M_{a^\uparrow a^\downarrow b^\uparrow} \neq M_{a^\uparrow b^\downarrow a^\uparrow}$, the later state being a higher energy hyperfine state. Hence we can write the expansion as,

$$M(\delta\mu_a^\uparrow, \delta\mu_b^\uparrow, \delta\mu_c^\downarrow) = \sum_{i=0}^{N}\sum_{j=0}^{i}\sum_{k=0}^{N} \gamma_{ijk}(\delta\mu_a^i\delta\mu_b^j + \delta\mu_a^j\delta\mu_b^i)\delta\mu_c^k. \tag{4.11}$$

4.4.1 QCD+QED

We begin as with the QCD case, assume that the function describing the meson mass is analytic, with sea quark masses held fixed, and restrict our selves to non-diagonal flavour mesons. The analytic expansion for the mass is given by,

$$M_{a\bar{b}} = \sum_{i=0}^{N}\sum_{j=0}^{N}\sum_{k=0}^{N}\sum_{m=0}^{N} \beta_{ijkm}^{ab}\delta\mu_a^i\delta\mu_b^j\delta e_a^k\delta e_b^m \tag{4.12}$$

where a, b are the quark flavours and $\delta\mu_a$ is the change in valence quark mass for flavour a, and δe_a is the change in valence quark electric charge of flavour a. $M_{a\bar{b}}$ is the mass of a meson composed of quark of flavour a and anti-quark of flavour b. Now as before let's choose a point $m_u = m_d = m_s$ and $e_u = e_d = e_s$. The spectrum generated about this point from a lattice calculation will satisfy the symmetry $M_{a\bar{b}} = M_{c\bar{e}} = M(\delta\mu_a, \delta\mu_b, \delta e_a, \delta e_b)$ for $a, b, c, f \in \{u, d, s\}$ for all $\delta\mu_a = \delta\mu_c, \delta\mu_b = \delta\mu_f, \delta e_a = \delta e_c, \delta e_b = \delta e_f$. Defining $\delta\mu_1, \delta\mu_2, \delta e_1, \delta e_2 \in \mathbb{R}$,

$$M_{a\bar{b}} = M(\delta\mu_1, \delta\mu_2, \delta e_1, \delta e_2) = \sum_{i=0}^{N}\sum_{j=0}^{N}\sum_{k=0}^{N}\sum_{m=0}^{N} \beta_{ijkm}^{ab}\delta\mu_1^i\delta\mu_2^j\delta e_1^k\delta e_2^m = \tag{4.13}$$

$$M_{c\bar{e}} = M(\delta\mu_1, \delta\mu_2, \delta e_1, \delta e_2) = \sum_{i=0}^{N}\sum_{j=0}^{N}\sum_{k=0}^{N}\sum_{m=0}^{N} \beta_{ijkm}^{ce}\delta\mu_1^i\delta\mu_2^j\delta e_1^k\delta e_2^m. \tag{4.14}$$

As before we must require that $\beta^{ab}_{ijkm} = \beta^{ce}_{ijkm}$ hence also $\beta^{ab}_{ijkm} = \beta^{ba}_{ijkm}$, so the expression is simplified to,

$$\begin{aligned} M_{a\bar{b}} &= M(\delta\mu_a, \delta\mu_b, \delta e_a, \delta e_b) \\ &= \sum_{i=0}^{N}\sum_{j=0}^{i}\sum_{k=0}^{N}\sum_{m=0}^{k} \beta_{ijkm}(\delta\mu_a^i\delta\mu_b^j\delta e_a^k(-\delta e_b)^m + \delta\mu_a^j\delta\mu_b^i\delta e_a^m(-\delta e_b)^k). \end{aligned} \tag{4.15}$$

In this case a further simplification can be made. If the point of expansion is about $e_u = e_d = e_s = 0$, then from the lattice we know that we will get $M(0,0,0,\delta e_1) = M(0,0,\delta e_1,0)$ for $\delta e_1 \in \mathbb{R}$. The two particles have opposite charges but the same mass, we conclude the polynomial of charge must be even. Hence we require that all the odd terms in the charge expansion be dropped. Also we can simplify our notation to $\delta e_a = e_a$,

$$\begin{aligned} M_{a\bar{b}} &= M(\delta\mu_a, \delta\mu_b, e_a, e_b) \\ &= \sum_{i=0}^{N}\sum_{j=0}^{i-1}\underbrace{\sum_{k=0}^{N}}_{even}\sum_{m=0}^{k-1} \beta_{ijkm}(\delta\mu_a^{i-j}\delta\mu_b^j e_a^{k-m}(-e_b)^m + \delta\mu_a^j\delta\mu_b^{i-j} e_a^m(-e_b)^{k-m}). \end{aligned} \tag{4.16}$$

We can generate a similar expression for the spin 3/2 baryons. Consider the case where $M(\delta\mu_1, \delta\mu_2, \delta\mu_3, e_1, e_2, e_3) = M(\delta\mu_1, \delta\mu_2, \delta\mu_3, -e_1, -e_2, -e_3)$ which must be true because positive and negative charge satisfy an S_2 symmetry. Once again we have $m + n + l = even$, and the analytic expansion for the mass is given by,

$$\begin{aligned} M_{abc} &= M(\delta\mu_a, \delta\mu_b, \delta\mu_c, e_a, e_b, e_c) \\ &= \sum_{i=0}^{N}\sum_{j=0}^{i}\sum_{k=0}^{j}\underbrace{\sum_{l=0}^{N}}_{even}\sum_{m=0}^{l}\sum_{n=0}^{min(l-m,m)} \gamma^{lmn}_{ijk} \\ &(\delta\mu_a^i\delta\mu_b^j\delta\mu_c^k e_a^{l-m-n} e_b^m e_c^n + \delta\mu_a^i\delta\mu_b^k\delta\mu_c^j e_a^{l-m-n} e_b^n e_c^m \\ &+ \delta\mu_a^j\delta\mu_b^i\delta\mu_c^k e_a^m e_b^{l-m-n} e_c^n + \delta\mu_a^j\delta\mu_b^k\delta\mu_c^i e_a^m e_b^n e_c^{l-m-n} \\ &+ \delta\mu_a^k\delta\mu_b^i\delta\mu_c^j e_a^n e_b^{l-m-n} e_c^m + \delta\mu_a^k\delta\mu_b^j\delta\mu_c^i e_a^n e_b^m e_c^{l-m-n}). \end{aligned} \tag{4.17}$$

Once again the expansion for spin 1/2 baryons is slightly more complicated,

with an S_3 symmetry only in the quarks of paired spin. The expansion is given by,

$$M(\delta\mu_a^\uparrow, \delta\mu_b^\uparrow, \delta\mu_c^\downarrow, e_a^\uparrow, e_b^\uparrow, e_c^\downarrow) = \sum_{i=0}^{N}\sum_{j=0}^{i}\sum_{k=0}^{N}\underbrace{\sum_{l=0}^{N}}_{even}\sum_{m=0}^{l}\sum_{n=0}^{l-m}\gamma_{ijk}^{lmn}(\delta\mu_a^i\delta\mu_b^j e_a^{l-m-n}e_b^m + \delta\mu_a^j\delta\mu_b^i e_a^m e_b^{l-m-n})\delta\mu_c^k e_c^n. \quad (4.18)$$

4.5 Sea Quark Expansion

Sea quarks can only contribute to the mass of a hadron through the QCD or QED interactions, however the gauge bosons of these two forces can not tell the difference between the flavours. This means that all sea quark flavours contribute to the mass of a hadron, as shown in Fig. 4.4. At a point of exact SU(3) flavour symmetry, all the quarks are identical $m_u = m_d = m_s$ and $e_u = e_d = e_s$, and so around this point we expect hadron masses to have S_3 flavour symmetry in sea quark mass and charge.

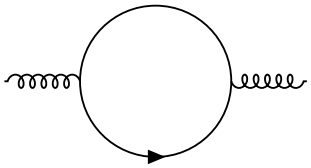


Figure 4.4: Feynman diagram of renormalisation of gluons by sea quarks. The diagram is summed over flavour.

We have already produced an expansion that was similar to this. The sea quark expansion is the same as the spin 3/2 baryon expansion but with the substitutions $\delta\mu \to \delta m$ $a \to u, b \to d, c \to s$ and the sea quark charges e_u, e_d, e_s. Making these substitutions,

$$\begin{aligned} M_{sea} &= M(\delta m_u, \delta m_d, \delta m_s, e_u, e_d, e_s) \\ &= \sum_{i=0}^{N}\sum_{j=0}^{i}\sum_{k=0}^{j}\underbrace{\sum_{l=0}^{N}}_{even}\sum_{m=0}^{l}\sum_{n=0}^{min(l-m,m)}\gamma_{ijk}^{lmn} \\ &(\delta m_u^i\delta m_d^j\delta m_s^k e_u^{l-m-n}e_d^m e_s^n + \delta m_u^i\delta m_d^k\delta m_s^j e_u^{l-m-n}e_d^n e_s^m \\ &+ \delta m_u^j\delta m_d^i\delta m_s^k e_u^m e_d^{l-m-n}e_s^n + \delta m_u^j\delta m_d^k\delta m_s^i e_u^m e_d^n e_s^{l-m-n} \\ &+ \delta m_u^k\delta m_d^i\delta m_s^j e_u^n e_d^{l-m-n}e_s^m + \delta m_u^k\delta m_d^j\delta m_s^i e_u^n e_d^m e_s^{l-m-n}) \end{aligned} \quad (4.19)$$

where M_{sea} is the Taylor expansion of a hadron mass when the valence quark parameters are fixed. For sea quarks we enforce the condition $\delta m_u + \delta m_d + \delta m_s = 0$, which removes the first quark mass term, and reduces every odd term, $\delta m_u^n + \delta m_d^n + \delta m_s^n$ for odd $n \in \mathbb{Z}$. It is important that any ensemble of configurations be generated along the $\delta m_u + \delta m_d + \delta m_s = 0$ to minimise the effect of unphysical sea quark mass on the hadron masses.

We can construct the full expansion for a hadron mass. To help with notation, call the valence quark mass expansion V_e and the sea quark mass expansion S_e. The resulting expansion, which includes both sea and valence quark effects, is simply $M(\delta\mu_a..., e_a..., \delta m_u..., e_u...) = (V_e)(S_e)$.

4.5.1 Particles with Disconnected Quark Line Contributions

An expansion can not be written down (in the same way as above) for particles with disconnected quark line contributions because they do not have a definite composition of valence quarks. The flavour composition of valence quarks changes with the valence quark mass. If the composition is known however, this can be used to determine many of the coefficients. There will still be additional coefficients in the expansion which depend on the disconnected quark line component to the mass. We discuss this in more detail in chapter 7.

4.6 Path to the Symmetric Point

The hadron spectrum satisfies an SU(3) flavour symmetry at any point where $m_u = m_d = m_s$ and $e_u = e_d = e_s$. In the introduction we stated that the closest SU(3) point to the physical point (on the path $\delta m_u + \delta m_d + \delta m_s = 0$) is most ideal because the extrapolation distance is the smallest. We have explored the advantage in determining singlet quantities when choosing an SU(3) symmetric point along the path $\delta m_u + \delta m_d + \delta m_s = 0$ to the physical point. Simply, linear terms in singlet quantities disappear and all odd terms will be reduced in magnitude. The reduction in these odd terms results from inevitable cancellation in terms of form $\delta m_u^n + \delta m_d^n + \delta m_s^n$ for n odd when enforcing the path condition. This becomes useful for more than one reason, in particular it is useful because it helps us define where the symmetric point is, as I will explain presently.

The quark masses are dependent on the scale at which we simulate, which is related to the lattice spacing and hence the QCD coupling coefficient. The quark mass parameter we can control for Wilson-style fermions is the hopping parameter

κ, which is related to the quark mass by,

$$am_q = \frac{1}{2}\left(\frac{1}{\kappa_q} - \frac{1}{\kappa_c}\right), \tag{4.20}$$

where κ_q is the input value and κ_c is the critical value of κ at which a neutral pseudoscalar meson mass $M_{q\bar{q}} = 0$.

As singlet quantities (refer to Tab. 4.1) remain constant along the chosen path, we would expect (for lattice spacing a) $\frac{a^2 X_\pi^2}{(aX_N)^2}$ to be approximately equal to the physical ratio $\left.\frac{X_\pi^2}{(X_N)^2}\right|^*$ at the κ value that corresponds to the SU(3) point. Note that the equality between X_S and X_S^* is only true to first order, however the second order correction in the ratio is reduced. If we define $X_\pi^2 = X_\pi^{2*}(1+\Delta_\pi)$ and $X_N^2 = X_N^{2\,*}(1+\Delta_N)$ then,

$$\frac{X_\pi^2}{X_N^2} = \frac{X_\pi^{2*}(1+\Delta_\pi)}{X_N^{2\,*}(1+\Delta_N)} \tag{4.21}$$

$$\approx \frac{X_\pi^{2*}}{X_N^{2\,*}}(1+\Delta_\pi)(1-\Delta_N), \tag{4.22}$$

$$\approx \frac{X_\pi^{2*}}{X_N^{2\,*}}(1+\Delta_\pi-\Delta_N) \tag{4.23}$$

where Δ_π and Δ_N are the corrections to X_π^2 and X_N^2 respectively. We achieve reduction in error if Δ_π and Δ_N have the same sign, which is a reasonable assumption as they depend on quark mass in a similar way. We can also check all ratios of X_π^2 to other singlets $S = N, \Delta, \rho$ and compare to the physical ratio. Using these ratios we can obtain the location of the SU(3) symmetric point in terms of κ to reasonable accuracy. Some of these ratios for pure QCD are show in Fig. 4.5 as a function of M_π^2/X_π^2. At this point, using ratios like $a = \left.\frac{a^2 X_\pi^2}{aX_N}\,\frac{X_N}{X_\pi^2}\right|^*$, we can obtain the lattice spacing a.

Though the above procedure to determine the SU(3) symmetric point seems clear, it becomes more complicated when we introduce QED. In the pure QCD case, at the SU(3) symmetric point,

$$X_\pi^2 = M_{u\bar{u}}^2 = M_{d\bar{d}}^2 = M_{s\bar{s}}^2 = X_\pi^{2*} \tag{4.24}$$

In the QCD case, all quark flavours share the same κ_c and hence (by Eq. (4.20) and Eq. (4.24)) the same κ at the symmetric point, κ^{sym}. So we can use Eq. (4.24) to define the symmetric point κ^{sym} in pure QCD, and we call this definition the QCD scheme.

When we include QED we do not normally set the charges to be the same, rather we set them to their physical values[5]. This results in the κ_c value being

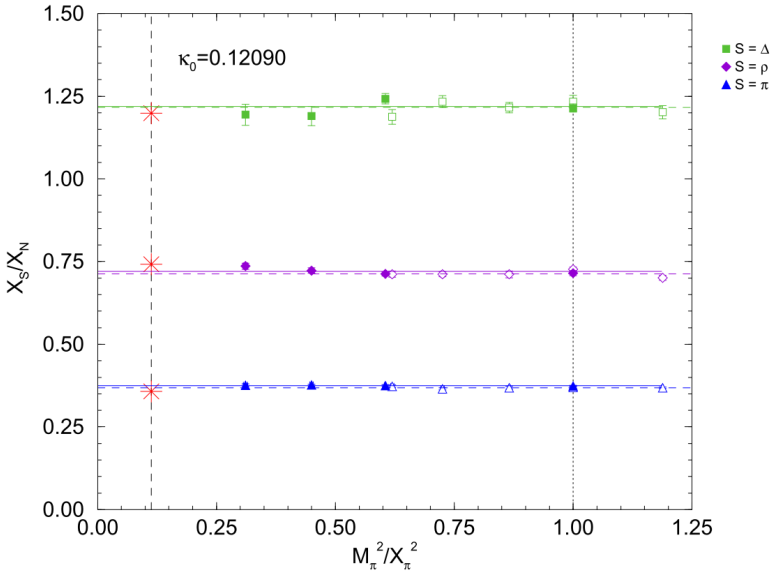


Figure 4.5: SU(3) flavour singlet ratios as a function of pion mass. The dashed vertical line is at the physical quark mass, the dotted line is at the SU(3) symmetric point. Filled points are from $32^3 \times 64$ lattice data while open points are from $24^3 \times 48$ lattice data. The horizontal lines are constant fits to the data points. The red star represents the value of the ratio at the physical quark mass. This figure was taken from [33].

different for quarks of different charge, hence it becomes unclear how we should define the symmetric point for each flavour.

Suppose we define the SU(3) symmetric point as the point in quark coordinates $(m_u, m_d, m_s, e_u, e_d, e_s)$ where the neutral mesons satisfy $M^2_{u\bar{u}} = M^2_{d\bar{d}} = M^2_{s\bar{s}} = M^2_{n\bar{n}} = X^{*2}_\pi$, the same as the QCD definition (the QCD scheme). The n quark is a quark with no charge. The definition is palatable because $X_\pi \approx X^*_\pi$, provided $\delta m_u + \delta m_d + \delta m_s = 0$ and $\delta(e_u^2) + \delta(e_d^2) + \delta(e_s^2) = 0$. When we include charges, and don't choose $e_u = e_d = e_s = e_n$ as the point to expand about (for instance choose the physical charges) if $m_u = m_d = m_s$ [6]then $M_{u\bar{d}} \neq M_{u\bar{u}} \neq M_{d\bar{d}}$ by definition.

If the charges are unequal, then the hadron mass symmetry is broken at any point we could reasonably define as a point of quark mass degeneracy [7]. In this case we would not expect the spectrum to satisfy an S_3 symmetry. By applying Eq. (4.24), we have in fact forced the condition $m_u \neq m_d = m_s \neq m_n$. The S_3

[5]In our research we set $\alpha_{QED} = \frac{e^2}{4\pi} \approx 10\alpha^*_{QED}$ and charges to $2/3, -1/3/, -1/3$ in units of electric charge for flavours u, d, s respectively.

[6]As the charges renormalise the quark masses this becomes difficult to define in terms of hadron mass. However as electromagnetic effects are small, we might consider that these masses are very close to the values for pure QCD.

symmetries fail around such a point, hence the above analytic expansions are not valid for this point.

Doing the extrapolation in this way enforces the assumption that QED effects are negligible, however it is typically these effects we wish to study when doing lattice QCD+QED. Often the results of this type of analysis will not be too far from correct because QED is a small effect, and it is only completely removed at the SU(3) point.

4.6.1 Dashen Scheme

It is however possible to find a better definition (more compatible with the quark definitions we have used in our Taylor expansions) of the SU(3) point in QCD+QED using the information we already have. We wish to make a quark mass definition so that, as we turn on electric charge the value of the quark masses we use in the expansion does not change. We redefine the expansion point to be $m_u = m_d = m_s = m_n = \bar{m}$, where $\bar{m}$ is defined so that $M^2_{n\bar{n}} = X^{2*}_\pi$. We are expanding around $m_u = m_d = m_s = m_n = \bar{m}$, $e_u = e_d = e_s = 0$. Hence the above analytic expansions are valid. The above definition states the QCD parts of the neutral meson masses, $\{M_{q\bar{q}}\}_{QCD}$ for $q \in \{u, d, s\}$ at the symmetric point must be equal, and this QCD part is given by the mass of the neutral meson of the neutral quark $M_{n\bar{n}}$.

Another issue is that the masses of the flavour diagonal mesons do not have the same slope with respect to $\frac{1}{\kappa}$ because of the differently charged quarks. This means a change in δm_u does not produce the same change in the hadron mass $M_{\bar{q}q}$ as a change in δm_d, and is a result of electric charge renormalisation of the quark masses and hadron. The issue is shown diagrammatically in Fig. 4.6.

We can scale the values of δm_q so that these changes are symmetric in flavour, which is essential if we are to use the expansions above. This is achieved by defining a scaling factor Z^q_m so that the flavour diagonal mesons have the same slope in $\frac{1}{\kappa}$.

If we define κ^{sym}_q as the point when $M^2_{q\bar{q}} = X^{2*}_\pi$ for quark flavour $q \in \{u, d, n\}$[8](the QCD scheme) and κ^c_q as the point when $M^2_{q\bar{q}} = 0$. The symmetric point quark mass

[7]This is a slightly false statement. There are two limiting cases and the truth lies somewhere in between. If we assume all of the mass generated from charge in the hadron is from charge renormalisation of quark mass, then any charges can be used and the symmetric point is obtained from the QCD scheme definition. However such an assumption is flawed because it removes the possibility of the quarks interacting electromagnetically with each other, and generating mass in that way. The other limiting case is to assume no renormalisation of quark mass, and attribute all mass gained to electromagnetic effects at a hadron level. The difficulty is that the true case is somewhere in between. Ultimately this is a question of how quark mass is defined when QED is included. The Dashen scheme (discussed below) is a compromise between these two limiting cases.

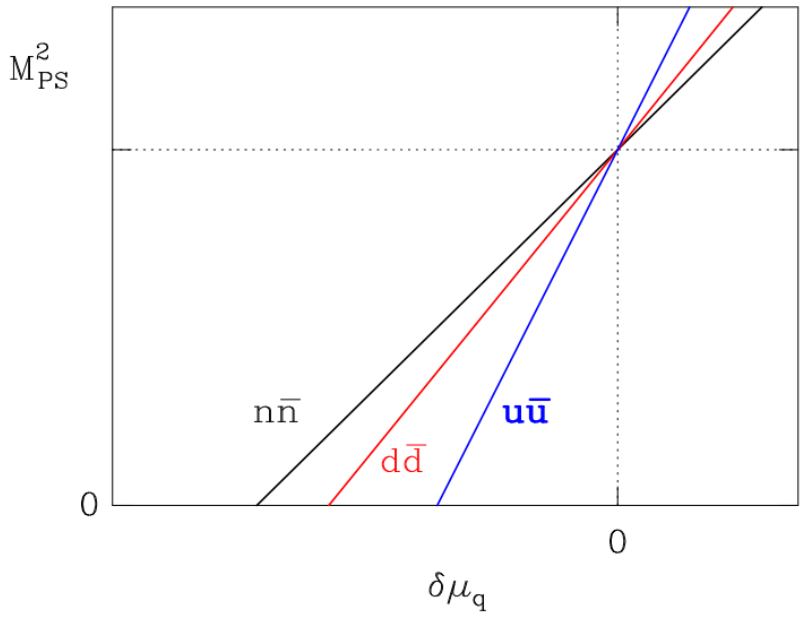


Figure 4.6: Sketch illustrating the dependence of flavour diagonal pseudoscalar meson masses on quark mass. Electric charge renormalisation of the quark masses has effected this dependence making the slopes different [34].

in the QCD scheme is, (using the quark mass definition)

$$am_q^{sym} = \frac{1}{2}\left(\frac{1}{\kappa_q^{sym}} - \frac{1}{\kappa_q^c}\right). \tag{4.25}$$

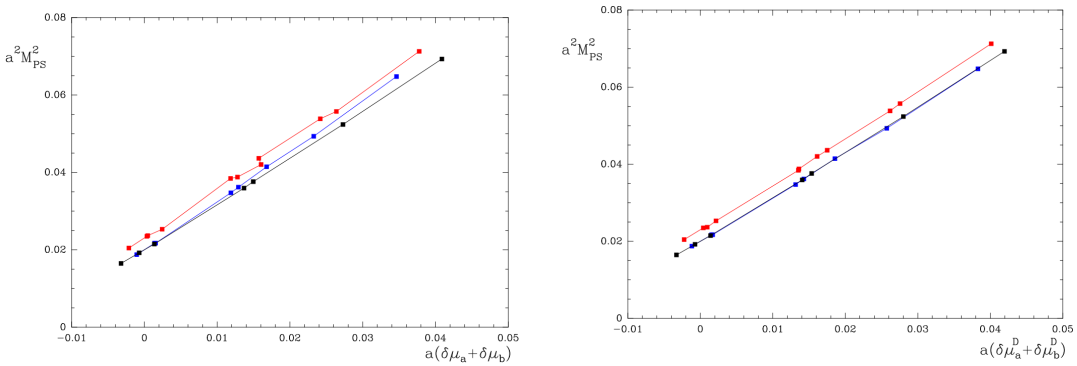
The correction of the slope of each charged δm_q is simply,

$$Z_m^q = \frac{m_n^{sym}}{m_q^{sym}} \tag{4.26}$$

for $q \in \{u, d, n\}$. When applying these scaling factors to the quark masses m_q and mass differences $\delta m_q, \delta\mu_q$ from the QCD scheme, we achieve a new scheme called the Dashen scheme [34]. Note, the Dashen scheme includes as part of its definition a common symmetric point where $m_n^{sym} = Z_m^q m_q^{sym} = \bar{m}$ is defined so that $M^2_{n\bar{n}} = X_\pi^{2*}$.

Ref. [34] showed there is significant improvement in fitting lattice QCD+QED mass spectra with the Dashen scheme against using the QCD scheme. The two spectrum plots from [34] are shown in Fig. 4.7, and demonstrate that it may be impossible to fit the data points by analytic functions if the incorrect scheme is used.

[8]Note $\kappa_d^{sym} = \kappa_s^{sym}$ and $\kappa_d^c = \kappa_s^c$.



(a) (b)

Figure 4.7: Pseudoscalar mesons π^+ (red), $u\bar{u}$ (blue) and $d\bar{d}$ (black). Both plots use the same data points however (a) is in the QCD scheme and (b) is in the Dashen scheme. We can see that the kink in the red points of (a), which can not be fit by regression (or a function), is removed by correctly setting the expansion point to the Dashen scheme. These plots were obtained from [34].

An important caveat about the Dashen scheme and schemes that separate QCD and QED components; the division we have made in the expansion coefficients based on QCD and QED components breaks down as we move away from the symmetric point because the division is scale dependent. This also means the expansion coefficients are scale dependent.

4.6.2 Improving the Dashen Scheme

The other small flaw in the Dashen scheme is that we set the mass of the neutral meson of the neutral quark equal to X_π^*; however X_π^* was calculated using the masses of charged mesons. We should have instead made the equality $M_{q\bar{q}}^{QCD} = X_\pi^{*QCD}$. To explore this idea, consider that at the physical point we can write an expansion for X_π^2,

$$X_\pi^2 = X_0^2 + a_i(\delta\mu_a + \delta\mu_b + \delta\mu_c) + b_i(e_a^2 + e_b^2 + e_c^2) + ... \tag{4.27}$$

to leading order. Notice this is not the general expansion. As $X_\pi^2 = \frac{1}{3}(M_{u\bar{d}}^2 + M_{d\bar{s}}^2 + M_{s\bar{u}}^2)$, X_π^2 is invariant under the S_3 symmetry around any expansion point. Here again we have chosen the $e = 0$ point to expand about, to preserve the mass symmetry about charge. So remaining along $\delta\mu_u + \delta\mu_d + \delta\mu_s = 0$ will remove the first term, and $\delta(e_a^2) + \delta(e_b^2) + \delta(e_c^2) = 0$ will keep the charge term constant, hence it can be ignored. Hence X_π^{2*} remains constant to leading order along the line. If we were to ignore the constraints on charges, then one could choose any charge values and that will not keep singlets approximately constant.

The intersection of this path with the $m_u = m_d = m_s$ and $e_u = e_d = e_s$ path is exactly the point we want. If we knew the physical mass and charges of the quarks it would simply be $(m_u + m_d + m_s)/3 = \bar{m}$ and $(e_u^2 + e_d^2 + e_s^2)/3 = \bar{e^2} = 2/9$. We don't know the masses but we do know the charges. Hence we can use the relation $X_\pi^{2*} \approx X_\pi^2$ and $e_u = e_d = e_s = \bar{e} = \sqrt{2}/3$ to determine the quark masses at the SU(3) point. That is $X_\pi^{2*} = M_{o\bar{o}}^2$ where the o quark has charge $\bar{e}$. Now the e^2 term in both analytic expansions is the same. We call this scheme the improved Dashen scheme. This definition is equivalent to $\{X_\pi^2\}_{QCD} = M_{n\bar{n}}^2$.

If we want to apply this idea for a lattice with $\alpha_{QED} = 10\alpha_{QED}^*$, then as $X_\pi^2 \propto \alpha_{QED} e^2$ we need a charge which is $\bar{e}^2/10$. Calculating what the symmetric point $\bar{e}^2$ should be, $9\bar{e}^2 = \frac{1^2+1^2+2^2}{3\times 10} = 1/5$ so $\bar{e} = \frac{\sqrt{1/5}}{3}$. Hence $M_{o\bar{o}}^2 = \frac{4}{5}M_{n\bar{n}}^2 + \frac{1}{5}M_{d\bar{d}}^2$. Finding the symmetric point using this scheme is shown diagrammatically in Fig. 4.8. Because the squared meson mass is linear in quark mass, our improved Dashen relation is simply $\bar{m} = \frac{4}{5}m_{n\bar{n}}^{sym} + \frac{1}{5}m_{d\bar{d}}^{sym}$, where $m_{q\bar{q}}^{sym}$ is defined at the point $M_{q\bar{q}}^2 = X_\pi^{2*}$. Similarly the scaling factors can be modified to implement this scheme,

$$Z_m^q = \frac{\frac{4}{5}m_{n\bar{n}}^{sym} + \frac{1}{5}m_{d\bar{d}}^{sym}}{m_q^{sym}}. \tag{4.28}$$

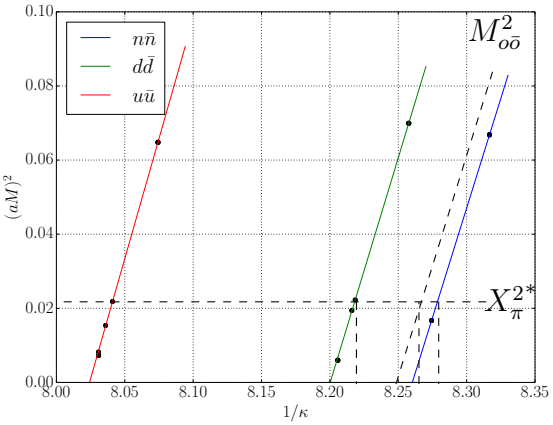


Figure 4.8: Neutral pseudoscalar mesons as a function of $1/\kappa$. Using the simple arguments and no additional data we can account for the fact the X_π^{2*} includes charges by defining the symmetric point to be the quark mass when $M_{o\bar{o}}^2 = X_\pi^{2*}$.

Note here that this only moves the position of the SU(3) symmetric point along the SU(3) line, and any changes in scaling factors are absorbed when we fit coefficients. Now that we understand what the electromagnetic correction to the Dashen scheme is, we can see that it is typically small, and so ignoring this term does not effect the solution greatly.

CHAPTER 5

Light Hadron Spectrum from Lattice QCD+QED

Isospin symmetry is satisfied within the hadron mass spectrum to remarkable accuracy. We observe the strength of the symmetry in the approximate mass degeneracy of isospin multiplets; in most cases isospin symmetry is obeyed to better than $\sim 1\%$[2]. Isospin violation is relevant to a range of physical phenomena, including the flavour decomposition of nucleon structure [1, 2, 3, 4]; tests of neutrino-nucleus interactions [5, 6]; precision constraints on CKM [7, 8] matrix elements from leptonic [9, 10] and semi-leptonic [11] decay rates; and quark mass parameters [12, 13, 14, 15]. In recent years these motivations have prompted extensive effort to introduce electromagnetic effects in numerical lattice QCD studies — building upon the pioneering work of Duncan, Eichten and Thacker [35].

In the present work, we perform simulations in dynamically-coupled QCD+QED [36, 34], where the electric charges of sea-quark loops are included in the fermion determinant. In this work, the hadron spectrum calculations are performed across $32^3 \times 64$ and $48^3 \times 96$ lattices with up to three distinct sea quark mass combinations. Starting from an SU(3) symmetric point inspired by Dashen's relation [37], we use a flavour symmetry breaking expansion [34] to extrapolate to the physical quark masses and interpolate to the physical QED coupling — where our underlying gauge ensembles use an unphysically-large $\alpha_{\text{QED}} \sim 0.1$ to enhance the signal strength in the electromagnetic effects. Partially-quenched correlators are employed to more accurately determine flavour symmetry breaking coefficients in these expansions. In addition to providing isospin splittings among the decuplet multiplets, we also present updated results for the octet baryons [36].

We begin this chapter, in section 5.1 and section 5.2, by describing some lattice specifics including the action and lattice interpolators we used. We then discuss the

finite volume corrections and the use of QED_L for toroidal topologies in section 5.3. In section 5.4 we detail the analytic expansions we used to fit the lattice data points, then describe the fitting procedure in section 5.5.

We conclude the chapter by presenting results and discussion in section 5.6. We present the mass splittings of the isospin multiplets of the octet and decuplet baryons as well as a breakdown of the electromagnetic and strong QCD components of these isospin violating mass splittings. We provide the first direct lattice calculation of the Δ^- mass. The results are summarised in section 5.7.

5.1 Lattice details

The QCD+QED action we are using in this study is given by

$$S = S_G + S_A + S_F^u + S_F^d + S_F^s, \tag{5.1}$$

where S_G is the tree-level Symanzik improved SU(3) gauge action; S_A is the non-compact U(1) gauge action of the photon; and S_F^q is the fermion action for each flavour, q. The photon action is,

$$S_A = \frac{1}{2e^2} \sum_{x,\mu<\nu} (A_\mu(x) + A_\nu(x+\mu) - A_\mu(x+\nu) - A_\nu(x))^2. \tag{5.2}$$

For the fermion action we emply the non-perturbatively $\mathcal{O}(a)$-improved SLiNC action [31],

$$\begin{aligned} S_F^q = \sum_x \Bigg\{ & \frac{1}{2} \sum_\mu \Big[\bar{q}(x)(\gamma_\mu - 1) e^{-ie_q A_\mu(x)} \tilde{U}_\mu(x) q(x+\hat{\mu}) \\ & - \bar{q}(x)(\gamma_\mu + 1) e^{ie_q A_\mu(x)} \tilde{U}_\mu^\dagger(x-\hat{\mu}) q(x-\hat{\mu}) \Big] \\ & + \frac{1}{2\kappa_q} \bar{q}(x) q(x) - \frac{1}{4} c_{SW} \sum_{\mu\nu} \bar{q}(x) \sigma_{\mu\nu} F_{\mu\nu} q(x) \Bigg\} \end{aligned} \tag{5.3}$$

where $\tilde{U}_\mu$ is a single iterated mild stout smeared link. This is a small modification to the standard clover action defined in Eq. (3.10). The clover coefficient c_{SW} has been computed non-perturbatively for pure QCD, we do not include the QED clover term.

Simulations are carried out on lattice volumes of size $32^3 \times 64$ and $48^3 \times 96$. The sea quark κ values are shown in Tab. 5.1 and charges of $e_u = +2/3$, $e_d = e_s = -1/3$, in units of electron charge. The strong coupling was chosen to be $\beta = 5.50$ and the electromagnetic coupling was chosen to be $e^2 = 1.25$, approximately

ten times greater than physical. These choices lead to a lattice spacing of $a = 0.068(1)$fm [34]. Further details can be found in [34] and [36].

In order to better constrain the (*a priori*) unknown coefficients in the flavour-breaking expansions, we employ partially-quenched valence quarks corresponding to neutral pseudoscalar meson masses in the range 225 MeV $\lesssim m_{\bar{q}q} \lesssim$ 765 MeV and valence quark charges $e_{a,b} = 0, -1/3, +2/3$.

#	β	e^2	V	$\kappa_u, \ +2/3$	$\kappa_d, \ -1/3$	$\kappa_s, \ -1/3$
1	5.50	1.25	$32^3 \times 64$	0.124362	0.121713	0.121713
2	5.50	1.25	$32^3 \times 64$	0.124440	0.121676	0.121676
3	5.50	1.25	$32^3 \times 64$	0.124508	0.121821	0.121466
4	5.50	1.25	$48^3 \times 96$	0.124362	0.121713	0.121713
5	5.50	1.25	$48^3 \times 96$	0.124440	0.121676	0.121676

Table 5.1: Summary of lattice details.

Further details can be found in [38] and [36].

5.2 Lattice Interpolators

Hadron masses are computed from two-point correlation functions using conventional techniques, as described in section 3.8. In particular, for baryons we construct zero-momentum two-point functions as

$$C(t) = \sum_{\vec{x}} \mathrm{Tr}\,\Gamma \langle\, T\big(\chi(\vec{x},t)\bar{\chi}(0)\big)\,\rangle \ , \tag{5.4}$$

for some choice of baryon spin projection matrix, Γ, for example for spin-averaged, $\Gamma = (1+\gamma_4)/2$. For octet baryons, we employ the interpolating operator in terms of a doubly-represented quark of flavour, q_1, and a singly-represented quark of flavour, q_2

$$\chi(\vec{x},t) = \epsilon^{abc}\big(q_1^{aT}(\vec{x},t)\, C\gamma_5\, q_2^b(\vec{x},t)\big) q_1^c(\vec{x},t) \ , \tag{5.5}$$

where a, b, c are colour labels. In the following, given the partially quenched nature of our simulations, we distinguish flavour by the electric charge carried by a quark rather than its mass. For example, when the combination *uud* occurs in the following discussion, this refers to an octet baryon where its doubly-represented quark has charge $+2/3$ while the singly-represented quark has charge $-1/3$. For decuplet baryons we choose an explicit spin-projection for the scalar di-quark of

the interpolating operator that contains doubly- and singly-represented quarks

$$\chi(\vec{x},t) = \frac{1}{\sqrt{3}}\epsilon^{abc}\Big[2\big(q_1^{aT}(\vec{x},t)\,C\gamma_-\,q_2^b(\vec{x},t)\big)q_1^c(\vec{x},t) + \big(q_1^{aT}(\vec{x},t)\,C\gamma_-\,q_1^b(\vec{x},t)\big)q_2^c(\vec{x},t)\Big]\,, \tag{5.6}$$

where $\gamma_- = (\gamma_2 + i\gamma_1)/2$.

5.3 Finite Volume Corrections

The observables which we wish to calculate are at physical quark mass, in continuous space-time and in an infinite volume. However when doing a lattice calculation none of these are achieved exactly. By applying effective field theory, estimates of the contribution of these effects on the observable can be calculated and the observable can be corrected.

A consideration when including QED on a lattice is the contribution from constant electromagnetic background fields (or photon zero modes) to the observable. If the average background field in each time slice is non-zero, observables may be altered and hence does not reflect the infinite volume case (where the average background field will be zero over all space). To account for this, we employ the so-called QED$_L$ formulation [39], which is QED defined in space dimensions with toroidal topology $\mathbb{R}\times\mathbb{T}^3$, and remove the zero mode of the photon field on each time slice for the valence quarks.

Additionally, the periodic boundary conditions modify the energy of the state due to its finite extent. In the present work we consider the leading finite volume (FV) corrections associated with the electromagnetic interaction. Strong interaction FV effects are expected to be subdominant as they are exponentially suppressed by $\exp(-m_\pi L)$, whereas electromagnetic FV effects are power law suppressed $\mathcal{O}(1/L)$. At the current volume the electromagnetic FV effects are of the order 1% of the total particle energy. Though it can be a considerable component in the mass splitting of a particle.

The FV correction can be estimated by calculating the difference in electromagnetic contribution to the mass when using QED compared to QED$_L$. The equation used to correct the state energy is [39, 40, 41, 42],

$$M_{corr} = M_{lat}\left[1 + \frac{q^2\alpha_{lat}k}{2M_{lat}L}\left(1 + \frac{2}{M_{lat}L}\right)\right], \tag{5.7}$$

where $k = 2.837297$, α_{lat} is the fine structure constant on the lattice, q is the electric charge of the particle and L is the spatial lattice length. Here we have

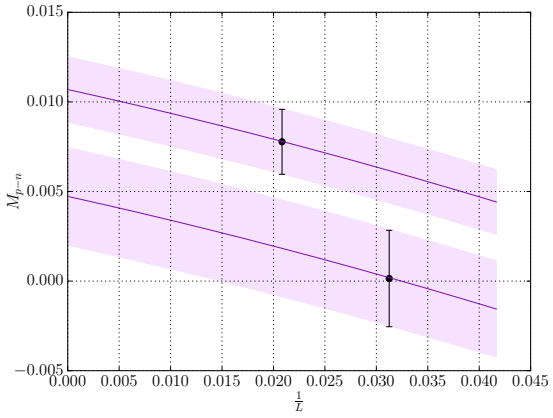


Figure 5.1: Finite volume correction on the lattice quantity M_{p-n} at the symmetric point. Points are from the $32^3 \times 64$ and $48^3 \times 96$ lattice ensembles respectively, and the curves extrapolate to infinite volume using Eq. (5.7).

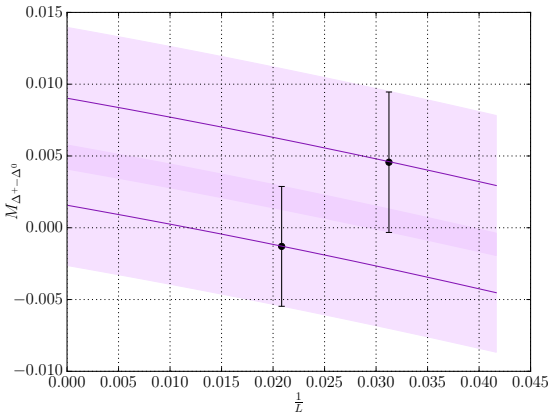


Figure 5.2: Finite volume correction on the lattice quantity $M_{\Delta^+ - \Delta^0}$ at the symmetric point. Points are from the $32^3 \times 64$ and $48^3 \times 96$ lattice ensembles respectively, and the curves extrapolate to infinite volume using Eq. (5.7).

included FV corrections to the mass up to (and including) NLO calculated in QED_L at zero momentum.

As the QCD finite volume effects on similar mass hadrons will be the same, any mass differences will only be sensitive to QED finite volume effects. In Fig. 5.1 and Fig. 5.2 we present results for proton-neutron mass difference and the Δ^+–Δ^0 mass difference at the SU(3) symmetric point on two different volumes ($\sim 2.2, 3.3$ fm). The curves show the extrapolation of the points to infinite volume using Eq. (5.7). We see that within two standard deviations that the two volumes are in agreement. For nucleon and decuplet mass splittings, we are unable to distinguish between statistical and residual finite volume errors, hence we consider the finite volume errors to be under control.

5.4 Mass Expansions

By exploiting the properties of the mass spectrum of pseudoscalar mesons and light baryons around points of exact SU(3) flavour $\{u, d, s\}$ symmetry we are able to reduce significantly the number of coefficients required in an extrapolation to the physical point. To further simplify the expansion, we choose the SU(3) point closest to the physical point, which lies along the line $\delta m_u + \delta m_d + \delta m_s = 0$ from the physical point. Terms in the analytic expansion proportional to $(\delta m_u + \delta m_d + \delta m_s)$ are set to zero. Additionally, singlet quantities along this line remain constant to first order. This allows us to use the singlet quantity $X_\pi^2 = \frac{1}{6}(M_{K^+}^2 + M_{K^0}^2 + M_{\pi^+}^2 + M_{\pi^-}^2 + M_{\bar{K}^0}^2 + M_{K^-}^2)$ to determine the quark mass coordinates of this SU(3) point. The SU(3) point is determined in terms of κ values for each quark of different charge $\{u, d = s\}$ and can be determined for a fictitious quark with zero charge n. For the details refer to chapter 4.

In this section we present these reduced analytic expansions for hadron masses in terms of their variables. The mass expansion for the pseudoscalar mesons is

given by,

$$\begin{aligned}
M^2(a\bar{b}) &= M_0^2 + \alpha(\delta\mu_a + \delta\mu_b) + \beta_1(\delta\mu_a^2 + \delta\mu_b^2) \\
&+ \beta_2\delta\mu_a\delta\mu_b + \beta_1^{EM}(e_a^2 + e_b^2) + \beta_2^{EM}(e_a e_b) \\
&+ \gamma_1^{EM}(e_a^2\delta\mu_a + e_b^2\delta\mu_b) + \gamma_2^{EM}(e_a e_b)(\delta\mu_a + \delta\mu_b) \\
&+ \gamma_3^{EM}(e_b^2\delta\mu_a + e_a^2\delta\mu_b) \\
&+ c_1(\delta m_u + \delta m_d + \delta m_s) \\
&+ c_2(\delta m_u^2 + \delta m_d^2 + \delta m_s^2 - (\delta m_u\delta m_d + \delta m_u\delta m_s + \delta m_d\delta m_s)) \\
&+ c_3(\delta m_u + \delta m_d + \delta m_s)^2 + c_4(e_u^2\delta m_u + e_d^2\delta m_d + e_s^2\delta m_s) \\
&+ c_1^{EM}(e_u^2 + e_d^2 + e_s^2) + c_2^{EM}(e_u e_d + e_u e_s + e_d e_s) \\
&+ c_3^{EM}(e_u^2 + e_d^2 + e_s^2)(\delta\mu_a + \delta\mu_b) + \mathcal{O}(\delta\mu^3, e^4).
\end{aligned} \tag{5.8}$$

In this expansion, the valence quark charges are indicated by $e_{a,b}$ and the sea quark charges by $e_{u,d,s}$. $\delta\mu$ denotes a valence quark mass variation from the SU(3) symmetric point, whereas δm describes the variation of the underlying sea quark masses from the SU(3) symmetric point. These quark mass variations are evaluated in the Dashen scheme [34], where the distance of the symmetric point to the chiral limit m_q^{sym} is defined to be independent of the quark charge, as discussed in section 4.6.

Given that our framework is to approach the physical point along a trajectory that holds the singlet quark mass approximately constant, we can neglect the c_1 and c_3 terms. Furthermore, the span of our sea quark masses are unable to provide any meaningful constraint on terms involving the sea masses. In particular, we neglect c_2 as $\mathcal{O}(\delta m^2)$ and c_4 as $\mathcal{O}(\alpha\delta m)$. The c^{EM} terms could be determined with simulations at different values of the QED gauge coupling, however in our present study these terms are simply absorbed into a redefinition of the relevant expansion parameters.

At the same order of the expansion, we write the expressions for the octet baryons:

$$\begin{aligned}
M(aab) &= M_0 + \alpha_1(2\delta\mu_a + \delta\mu_b) + \alpha_2\delta\mu_a \\
&+ \beta_1(2\delta\mu_a^2 + \delta\mu_b^2) + \beta_2(\delta\mu_a^2 + 2\delta\mu_a\delta\mu_b) + \beta_3(\delta\mu_a^2) \\
&+ \beta_1^{EM}(2e_a^2 + e_b^2) + \beta_2^{EM}(e_a^2 + 2e_a e_b) + \beta_3^{EM}(e_a^2) \\
&+ \gamma_1^{EM}(2e_a^2\delta\mu_a + e_b^2\delta\mu_b) \\
&+ \gamma_2^{EM}(2\delta\mu_a e_a(e_a + e_b) + 2\delta\mu_b e_b e_a)) \\
&+ \gamma_3^{EM}(2\delta\mu_a e_a e_b + \delta\mu_b e_a^2) \\
&+ \gamma_4^{EM}(2\delta\mu_a(e_a^2 + e_b^2) + 2\delta\mu_b e_a^2) \\
&+ \gamma_5^{EM}\delta\mu_a e_a^2 + \gamma_6^{EM}\delta\mu_a e_a e_b
\end{aligned} \tag{5.9}$$

and the decuplet baryons:

$$\begin{aligned}
&M(abc) = M_0 + \alpha_1(\delta\mu_a + \delta\mu_b + \delta\mu_c) \\
&+ \beta_1(\delta\mu_a^2 + \delta\mu_b^2 + \delta\mu_c^2) + \beta_2(\delta\mu_a\delta\mu_b + \delta\mu_a\delta\mu_c + \delta\mu_b\delta\mu_c) \\
&+ \beta_1^{EM}(e_a^2 + e_b^2 + e_c^2) + \beta_2^{EM}(e_a e_b + e_a e_c + e_b e_c) \\
&+ \gamma_1^{EM}(e_a^2\delta\mu_a + e_b^2\delta\mu_b + e_c^2\delta\mu_c) \\
&+ \gamma_2^{EM}(\delta\mu_a e_a(e_b + e_c) + \delta\mu_b e_b(e_a + e_c) + \delta\mu_c e_c(e_a + e_b)) \\
&+ \gamma_3^{EM}(\delta\mu_a e_b e_c + \delta\mu_b e_a e_c + \delta\mu_c e_a e_b) \\
&+ \gamma_4^{EM}(\delta\mu_a(e_b^2 + e_c^2) + \delta\mu_b(e_a^2 + e_c^2) + \delta\mu_c(e_a^2 + e_b^2))
\end{aligned} \tag{5.10}$$

As argued above, we have already dropped the terms involving the sea quark masses and charges.

5.5 Fitting procedure

To use these expressions we must first determine the symmetric point (in the Dashen scheme definition), as described in chapter 4. The symmetric point can be determined from the neutral pseudoscalar meson masses obtained at various values of the hopping parameter. A graph of the fits to the neutral pseudoscalar meson masses is shown in Fig. 5.3. The intercept with the horizontal axis provides a value for κ_a^c for each quark flavour a, while the symmetric point (in the QCD scheme) κ_a^{sym} is defined as the κ_a value when $M_{a\bar{a}}^2 = X_\pi^{2^*}$.

The change in valence quark mass of flavour a (in the QCD scheme) is defined as $\delta\mu_a^Q = \frac{1}{2}\left(\frac{1}{\kappa_a} - \frac{1}{\kappa_a^{sym}}\right)$. The $\delta\mu_a^Q$ are then scaled using the scaling factors Z_m^a as defined in Eq. (4.26) to produce the Dashen scheme definition of $\delta\mu$,

$$\delta\mu_a^D = Z_m^a \delta\mu^Q. \tag{5.11}$$

In most of this thesis the Dashen scheme is assumed and so we do not include the D superscript. More discussion on the Dashen scheme can be found in section 4.6.1.

We can now plot the meson mass spectrum against the Dashen scheme $\delta\mu$, as shown in Fig. 5.5, and fit expansion Eq. (5.8) to these points. The fit is shown using black crosses in the figure.

Once the symmetric and critical points are determined from the meson spectrum, we can also produce mass spectra for the octet and decuplet baryons. These spectra can be fit using Eq. (5.9) and Eq. (5.10) respectively. Plots of the spectra and the fits are shown in Fig. 5.6 and Fig. 5.7. The coefficient values of both mesons and baryons are shown in Tab. 5.2.

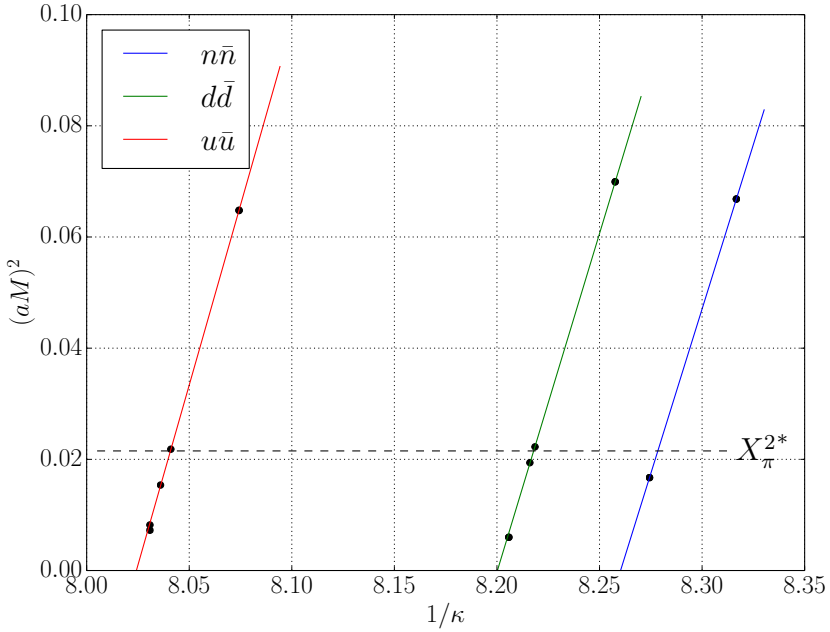


Figure 5.3: The flavour-neutral meson masses for each flavour u, d, n as a function of hopping parameter κ. n is a electrically neutral quark, while u and d have charges $+2/3$ and $-1/3$ with $\alpha_{QED} \approx 10\alpha^*_{QED}$.

As we work with, $\alpha^{lat}_{QED} = \frac{1.25}{4\pi} \approx 10\alpha^*_{QED}$ we correct all coefficients which relate polynomials with charges, by scaling (dividing) them by a factor of

$$Z_{QED} = \frac{Z_3 \alpha^{lat}_{QED}}{\alpha^*_{QED}} \tag{5.12}$$

per factor of e^2, where $Z_3 = 0.9$ is the charge renormalisation factor[43]. The corrected values are the numbers tabulated.

We have also included in Tab. 5.2 the correlated and uncorrelated χ^2/DOF. The χ^2 is the cost function we minimise on to produce the fit parameters. In the uncorrelated case, χ^2 is the Euclidean distance squared between the fit estimate and the data point weighted by the inverse of the variance of the data point. The full χ^2 is the sum of all of these weighted Euclidean distances.

Another metric for determining the minimum considers the correlations between the distinct data points. In our case we use a boot strap method to obtain estimates of errors on all parameters, and so the correlated χ^2_{corr} take into account how the data points change together on each boot strap. To be clear, though we did minimise the χ^2_{corr} to produce a set of parameters, the uncorrelated χ^2 parameters appear in the table. The correlated χ^2_{corr} is only included here to give additional information on whether we have over or under parameterised our model.

The correlated χ^2_{corr} values are greater than one, which suggests that (in some aspect) we have under parameterised our model. This is possibly due to excluding sea quark parameters while including different gauge ensembles 4 and 5. In the correlated χ^2_{corr} case the fit did not place the fit points as close to the data points, however the fit did seem successful and approached the parameters displayed in Tab. 5.2.

The errors on the parameters in Tab. 5.2 are determined from conducting the uncorrelated fit for each bootstrap ensemble and then taking the standard deviation. We also provide normalised correlation matrices for the parameters of the fit on each set of particles (pseudoscalar meson, octet and decuplet) in Fig. 5.4.

We can see from the covariance matrices that certain parameters are typically not determined well by the data set. For instance the data set is very linear, and so it is difficult to determine coefficients that are proportional to $(\delta\mu)^2$. As such, there is high correlation between the α_i coefficients and the β_i coefficients.

The cross correlations tend to form two groups, terms with electromagnetic effects and terms without. There parameter covariance is lowest for the pseudoscalar mesons, then the octet and finally the decuplet. Given that the decuplet masses are the least constrained it could indicate that uncertainty in the data points allows for the changes in the data to be fit using different combinations of parameters. Including heavier quark masses could be a cheap (computational time) way of resolving the quadratic terms better, which may reduce uncertainty overall.

Once the coefficients of the meson spectrum are determined, but before we determine the physical point, we re-adjust the value of κ_c for each flavour u, d, n to align with the horizontal axis intercept of Eq. (5.8). This does not move the symmetric point, but puts it in terms of the intercept of the meson expansion, which is necessary when we determine physical quark masses. The expansions can now be used to determine an estimate for the physical quark mass parameters and the lattice spacing. This is achieved by determining the $\delta\mu_q = \delta m_q$ for $q \in \{u, d, s\}$ and lattice spacing a that achieve the physical meson masses π^+, K^0 and K^+ masses using Eq. (5.8). Note we can use an estimate of the lattice spacing as the starting position of the algorithm using ratios like $a = \frac{a^2 X_\pi^2}{a X_N} \frac{X_n^*}{X_\pi^{2*}}$. This value of lattice spacing does not change much during minimisation.

The results of the minimisation are given in Tab. 5.3 for the $48^3 \times 96$ volume. We note that we are only required to provide three physical inputs to determine the four unknown parameters as we have the additional constraints built into our simulations that $\delta m_u + \delta m_d + \delta m_s = 0$. Using the parameters given in Tab. 5.3 for the $48^3 \times 96$ volume, we are able to provide a prediction for the π^0 mass, which is

[1]Correlated χ^2 is not used to determine parameters, but is included for reference to provide a guide to whether our method is over or under parameterised.

	M_0	α_1	β_1	β_2	β_1^{EM}	β_2^{EM}	γ_1^{EM}	γ_2^{EM}	γ_3^{EM}
M_0	1.00	0.73	-0.72	-0.55	-0.38	-0.53	0.01	0.82	-0.54
α_1	0.73	1.00	-0.97	-0.41	-0.12	-0.20	-0.35	0.34	-0.36
β_1	-0.72	-0.97	1.00	0.31	0.19	0.12	0.43	-0.37	0.25
β_2	-0.55	-0.41	0.31	1.00	-0.16	0.45	-0.01	-0.35	0.10
β_1^{EM}	-0.38	-0.12	0.19	-0.16	1.00	-0.18	-0.08	-0.63	0.42
β_2^{EM}	-0.53	-0.20	0.12	0.45	-0.18	1.00	-0.55	-0.53	0.59
γ_1^{EM}	0.01	-0.35	0.43	-0.01	-0.08	-0.55	1.00	0.24	-0.52
γ_2^{EM}	0.82	0.34	-0.37	-0.35	-0.63	-0.53	0.24	1.00	-0.69
γ_3^{EM}	-0.54	-0.36	0.25	0.10	0.42	0.59	-0.52	-0.69	1.00

(a)

	M_0	α_1	α_2	β_1	β_2	β_3	β_1^{EM}	β_2^{EM}	β_3^{EM}	γ_1^{EM}	γ_2^{EM}	γ_3^{EM}	γ_4^{EM}	γ_5^{EM}	γ_6^{EM}
M_0	1.00	-0.77	-0.37	0.54	0.46	0.30	-0.54	-0.37	0.15	0.15	0.38	-0.19	0.65	0.08	0.29
α_1	-0.77	1.00	-0.19	-0.91	-0.76	0.20	0.15	0.38	0.21	0.12	-0.24	-0.08	-0.29	-0.30	0.00
α_2	-0.37	-0.19	1.00	0.49	0.20	-0.98	0.67	-0.09	-0.56	-0.62	-0.24	0.33	-0.62	0.45	-0.43
β_1	0.54	-0.91	0.49	1.00	0.57	-0.52	0.10	-0.39	-0.36	-0.27	0.26	0.29	0.01	0.36	-0.23
β_2	0.46	-0.76	0.20	0.57	1.00	-0.13	0.05	-0.07	-0.42	-0.19	-0.20	0.12	0.07	0.45	-0.05
β_3	0.30	0.20	-0.98	-0.52	-0.13	1.00	-0.66	0.13	0.60	0.63	0.17	-0.38	0.61	-0.50	0.47
β_1^{EM}	-0.54	0.15	0.67	0.10	0.05	-0.66	1.00	0.27	-0.75	-0.80	-0.45	0.65	-0.91	0.57	-0.76
β_2^{EM}	-0.37	0.38	-0.09	-0.39	-0.07	0.13	0.27	1.00	-0.26	-0.20	-0.75	0.27	-0.12	0.22	-0.28
β_3^{EM}	0.15	0.21	-0.56	-0.36	-0.42	0.60	-0.75	-0.26	1.00	0.60	0.38	-0.81	0.72	-0.78	0.80
γ_1^{EM}	0.15	0.12	-0.62	-0.27	-0.19	0.63	-0.80	-0.20	0.60	1.00	0.59	-0.30	0.52	-0.79	0.42
γ_2^{EM}	0.38	-0.24	-0.24	0.26	-0.20	0.17	-0.45	-0.75	0.38	0.59	1.00	-0.03	0.20	-0.60	0.08
γ_3^{EM}	-0.19	-0.08	0.33	0.29	0.12	-0.38	0.65	0.27	-0.81	-0.30	-0.03	1.00	-0.71	0.37	-0.97
γ_4^{EM}	0.65	-0.29	-0.62	0.01	0.07	0.61	-0.91	-0.12	0.72	0.52	0.20	-0.71	1.00	-0.36	0.81
γ_5^{EM}	0.08	-0.30	0.45	0.36	0.45	-0.50	0.57	0.22	-0.78	-0.79	-0.60	0.37	-0.36	1.00	-0.35
γ_6^{EM}	0.29	0.00	-0.43	-0.23	-0.05	0.47	-0.76	-0.28	0.80	0.42	0.08	-0.97	0.81	-0.35	1.00

(b)

	M_0	α_1	β_1	β_2	β_1^{EM}	β_2^{EM}	γ_1^{EM}	γ_2^{EM}	γ_3^{EM}	γ_4^{EM}
M_0	1.00	-0.85	0.72	0.80	0.01	0.28	-0.33	0.53	-0.28	0.36
α_1	-0.85	1.00	-0.95	-0.84	0.08	-0.15	0.23	-0.45	0.09	-0.43
β_1	0.72	-0.95	1.00	0.65	-0.19	0.02	-0.14	0.29	0.08	0.49
β_2	0.80	-0.84	0.65	1.00	0.26	0.39	-0.41	0.64	-0.44	0.12
β_1^{EM}	0.01	0.08	-0.19	0.26	1.00	0.88	-0.81	0.20	-0.90	-0.75
β_2^{EM}	0.28	-0.15	0.02	0.39	0.88	1.00	-0.91	0.05	-0.79	-0.50
γ_1^{EM}	-0.33	0.23	-0.14	-0.41	-0.81	-0.91	1.00	-0.08	0.67	0.28
γ_2^{EM}	0.53	-0.45	0.29	0.64	0.20	0.05	-0.08	1.00	-0.54	-0.05
γ_3^{EM}	-0.28	0.09	0.08	-0.44	-0.90	-0.79	0.67	-0.54	1.00	0.68
γ_4^{EM}	0.36	-0.43	0.49	0.12	-0.75	-0.50	0.28	-0.05	0.68	1.00

(c)

Figure 5.4: Correlation matrices of the parameters in the fitting model. The label names along the left hand side are also valid along the top. The matrices are colour coded to guide the eye. Red is close to -1, green is close to 1 and the colour scales through to white at 0. (a), (b) and (c) show the correlation matrix for the pseudoscalar mesons, octet baryons and decuplet baryons respectively.

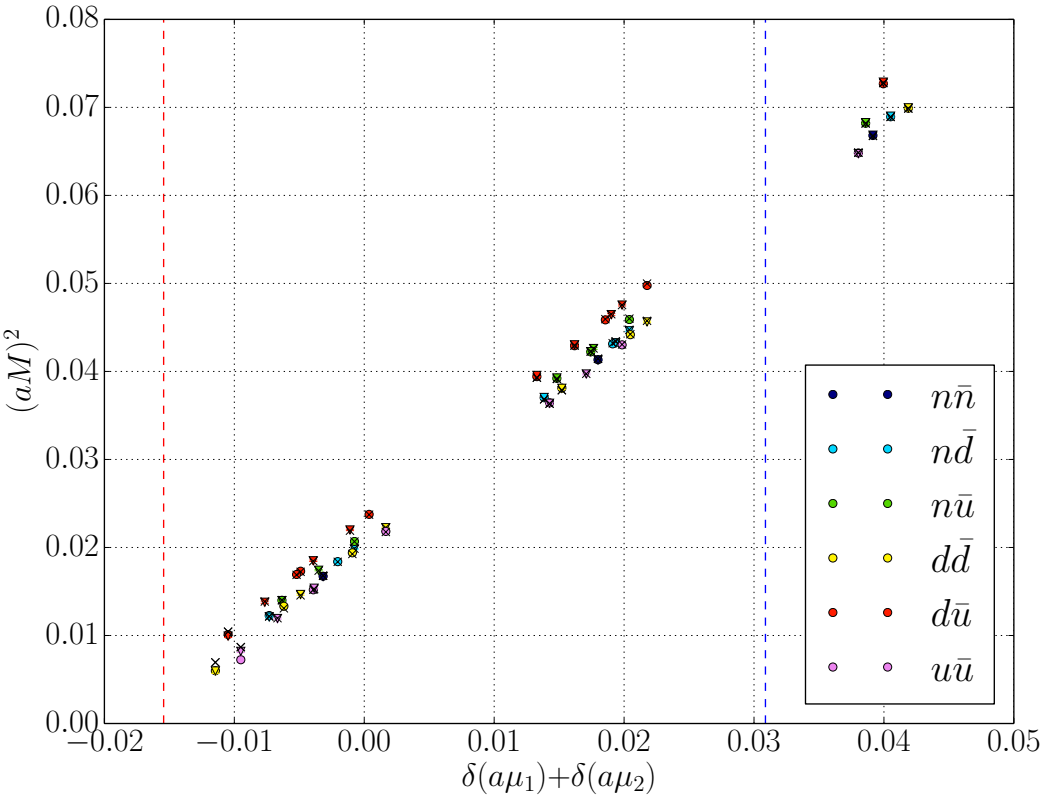


Figure 5.5: The mass spectrum of the pseudoscalar mesons in lattice units as a function of the displacement of the quark masses from the symmetric point value. The black crosses show what the expansion predicts. The circles are ensemble 4 and the triangles are ensemble 5. The red and blue vertical lines are the position of the π^+ and the $s\bar{s}$ mesons at physical quark mass.

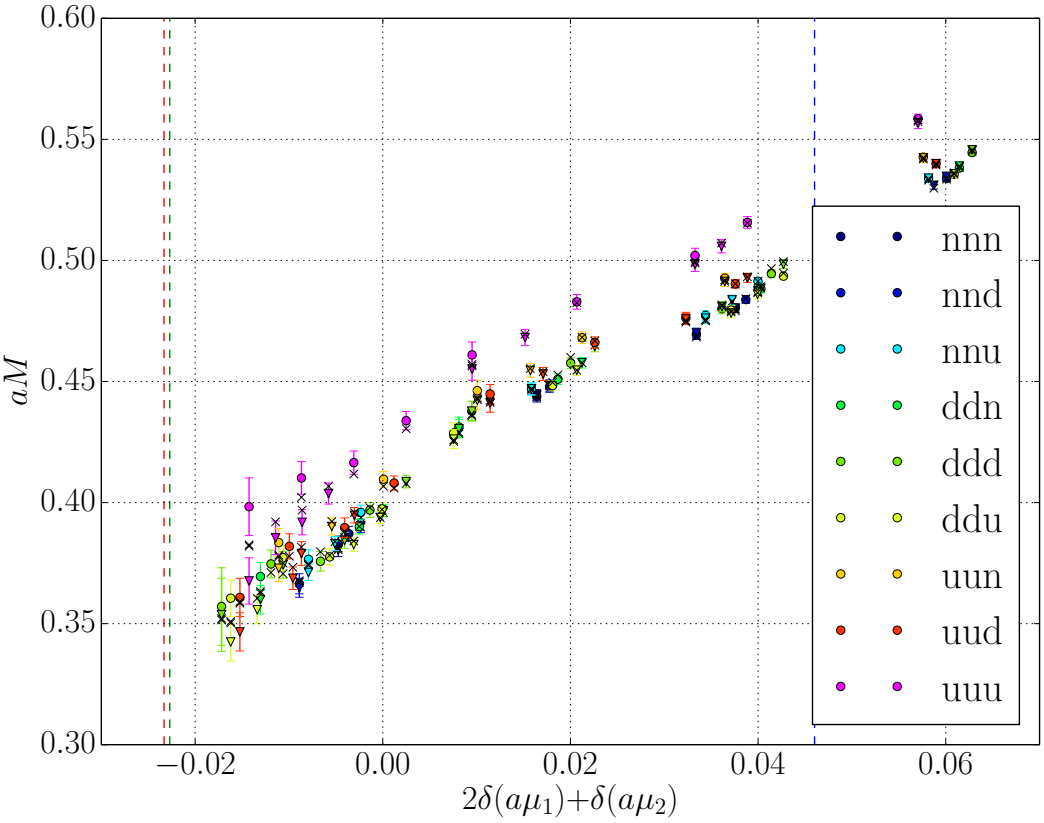


Figure 5.6: The mass spectrum of the octet baryons in lattice units as a function of the displacement of the quark masses from the symmetric point value. The black crosses show what the expansion predicts. The circles are ensemble 4 and the triangles are ensemble 5. The red, green and blue vertical lines are the position of the proton, neutron and sss octet baryons at physical quark mass.

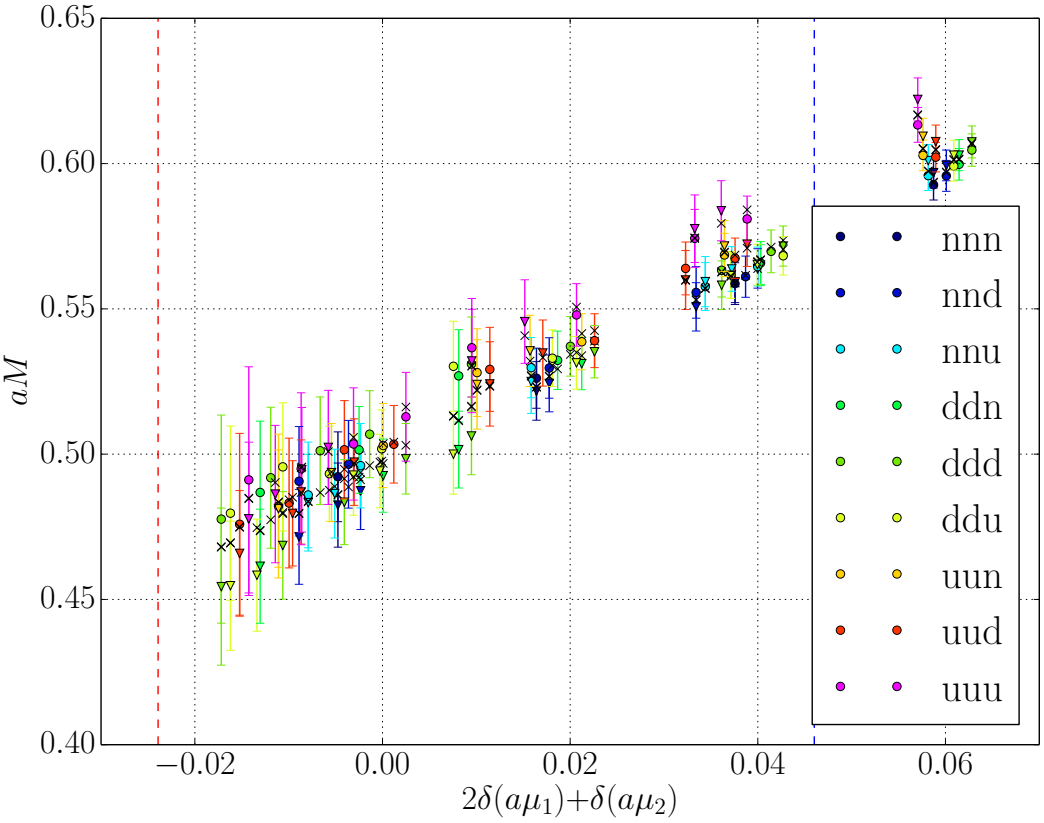


Figure 5.7: The mass spectrum of the decuplet baryons in lattice units as a function of the displacement of the quark masses from the symmetric point value. The black crosses show what the expansion predicts. The circles are ensemble 4 and the triangles are ensemble 5. The red and blue vertical lines are the position of the Δ^{++} and Ω physical quark mass.

	Meson	Octet	Decuplet
M_0	0.020504 (66)	0.3944 (24)	0.494 (11)
α_1	1.1703 (47)	3.32 (11)	1.73 (40)
α_2	–	-1.71 (23)	–
β_1	-0.17 (22)	-20.7 (40)	1.0 (135)
β_2	1.51 (12)	-14.2 (13)	-2.9 (45)
β_3	–	38.0 (11)	–
β_1^{EM}	0.0001975 (47)	0.00083 (17)	0.00064 (53)
β_2^{EM}	-0.0005222 (37)	0.001032 (55)	0.00042 (18)
β_3^{EM}	–	-0.00022 (33)	–
γ_1^{EM}	0.00435 (26)	-0.0041 (44)	0.012 (17)
γ_2^{EM}	-0.00899 (13)	0.0014 (11)	0.003 (5)
γ_3^{EM}	0.00526 (21)	-0.0063 (54)	0.0092 (61)
γ_4^{EM}	–	0.0014 (27)	-0.00011 (70)
γ_5^{EM}	–	0.008 (11)	–
γ_6^{EM}	–	0.016 (12)	–
χ^2	183.74	47.12	20.35
DOF	105	112	118
χ^2/DOF	1.75	0.42	0.172
χ^2_{corr}/DOF[1]	26.48	7.41	7.03

Table 5.2: Coefficients obtained from fitting the mass spectrum of the various particle types.

shown in Tab. 5.4 in the form of a mass splitting from the π^+. Note in this analysis we have ignored disconnected quark line terms and a more detailed analysis of the neutral flavour pseudoscalars is left for chapter 7. The result from the present work is in agreement with that from [34], however we note the improved statistical precision of the current work due to the inclusion of the additional ensembles away from the SU(3) symmetric point.

$a\delta m_u$	$a\delta m_d$	$a\delta m_s$	a^{-1}/GeV
-0.008166 (6)	-0.00754 (1)	0.01571 (1)	2.881 (6)

Table 5.3: Bare quark mass parameters at the physical point and the inverse lattice spacing using the Dashen scheme

Using the physical quark mass parameters and lattice spacing obtained by minimisation, we are able to produce an estimate of octet and decuplet baryon masses at the physical point. We achieve this by inserting the physical quark mass parameters (Tab. 5.3), charges and coefficients (Tab. 5.2) into Eq. (5.9) and Eq. (5.10). The lattice spacing is used to put the masses obtained into physical units of MeV.

5.6 Results

5.6.1 Octet Baryons

Using the above extrapolation procedure, we produce estimates of the splittings among the baryon octet, shown in Tab. 5.4. The table includes a total splitting as well as a breakdown of the QCD and QED components separately. The first uncertainty in the mass splittings is statistical, while the second provides an estimate of the systematic error. We note that since our simulations are performed at only a single value of the lattice spacing, no continuum extrapolation is possible. The electromagnetic splitting between the proton and neutron has seen considerable attention in recent years. Our result for the electromagnetic component to the proton-neutron mass difference agrees well with the dispersive analysis given by [44], $\delta M^\gamma|_{p-n} = 1.30(03)(47)$ and has good agreement with the lattice study by BMW collaboration [45] $\delta M^\gamma|_{p-n} = 1.59(46)$. The octet results of the present study are generally compatible with both phenomenological estimates [46] and the BMW lattice results [40], however these two studies agree with each other far better than with the present work.

Fig. 5.8 shows the composition of the splittings in terms of strong isospin breaking effects and electromagnetic effects graphically and allows us to compare our

	$\pi^+ - \pi^0$	$n - p$	$\Sigma^- - \Sigma^+$	$\Xi^- - \Xi^0$
QED	—	-1.53(25)(50)	-0.29(24)(10)	1.19(15)(20)
QCD	—	2.79(67)(40)	8.58(72)(70)	5.79(28)(80)
Total	5.86(14)(40)	1.27(75)(50)	8.29(77)(25)	6.95(25)(90)
Experiment	4.59	1.30	8.08	6.85

Table 5.4: Predicted mass splittings for π^+ and octet baryons in the Dashen scheme. Results from fitting $48^3 \times 96$ lattice data. π^0 assumed to be the state $(u\bar{u} - d\bar{d})/\sqrt{2}$. All values quoted in MeV. The first error is statistical, while the second error is an estimate of the systematic error, which is obtained by the difference in the $32^3 \times 64$ results to those of the $48^3 \times 96$.

	p	n	Σ^+
This work	939(14)(56)	940(14)(56)	1165(11)(23)
Experiment	938.3	939.6	1189.4
	Σ^-	Ξ^0	Ξ^-
This work	1173(10)(23)	1276(6)(19)	1283(6)(19)
Experiment	1197.5	1314.8	1321.7

Table 5.5: Predicted masses from the $48^3 \times 96$ ensembles for the octet baryons. All values quoted in MeV. The first error is statistical, while the second error is an estimate of the systematic error, which is obtained by the difference in the $32^3 \times 64$ results to those of the $48^3 \times 96$.

results with BMW estimates found in [40]. The lines in Fig. 5.8 represent a constraint placed by the experimentally observed mass splittings. A direct comparison for the QED component of the Σ in [40] is not possible since this is taken as input for the BMW analysis. Our results for both volumes show good agreement with the experimental constraints, but there is perhaps some discrepancy with the BMW collaboration over the correct ratio of forces in the $n - p$ splitting. Additionally, the assumption that the $\Sigma^- - \Sigma^+$ has zero electromagnetic contribution may not be accurate, however we are unable to resolve this within our statistical errors.

Fig. 5.9 shows the predicted mass splittings between the octet baryons for both lattice volumes. We can see on this plot that both volumes agree with each other within statistical errors, and with experiment. In this figure the precision on the splittings is higher due to a large degree of correlation in the mass determinations.

A table of the masses of the octet baryons in given in Tab. 5.5. The $48^3 \times 96$ show remarkable agreement with the nucleon physical masses (with significant

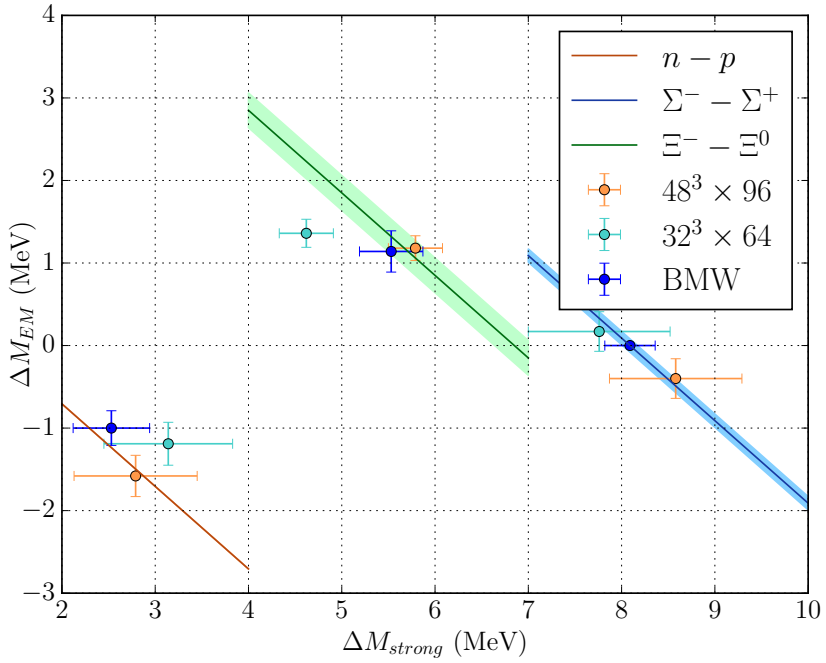


Figure 5.8: A break down of the octet splittings in terms of EM and isospin breaking effects. The BMW lattice points are from [40]. The lines represent a constraint placed by the experimentally observed mass splittings.

statistical error), but with decreasing accuracy for Σ and Ξ. Note that the systematic error is determined by the difference between the $48^3 \times 96$ and those of the $32^3 \times 64$. The $32^3 \times 64$, which are not tabulated, were systematically heavier than the $48^3 \times 96$ results. The uncertainty in the scale is approximately 2%, and most likely contributes to this difference. We do not expect these systematics to effect the isospin mass splittings of the states. These results provide an update of previous work by QCDSF [36], which was based on a single set of sea quark masses, ensembles 1 and 4.

5.6.2 Decuplet Baryons

For the decuplet baryons, we perform an extrapolation to the physical point based on the polynomial expansion about the SU(3) symmetric point. No attempt has been made in the present work to incorporate the effects of the resonant nature of the decuplet baryons at the physical quark masses — which necessarily lead to branch point singularities in the quark mass extrapolation [47, 48].

The absolute masses of the decuplet baryons at the physical quark mass are tabulated in Tab. 5.6. The absolute masses themselves do not compare so favourably

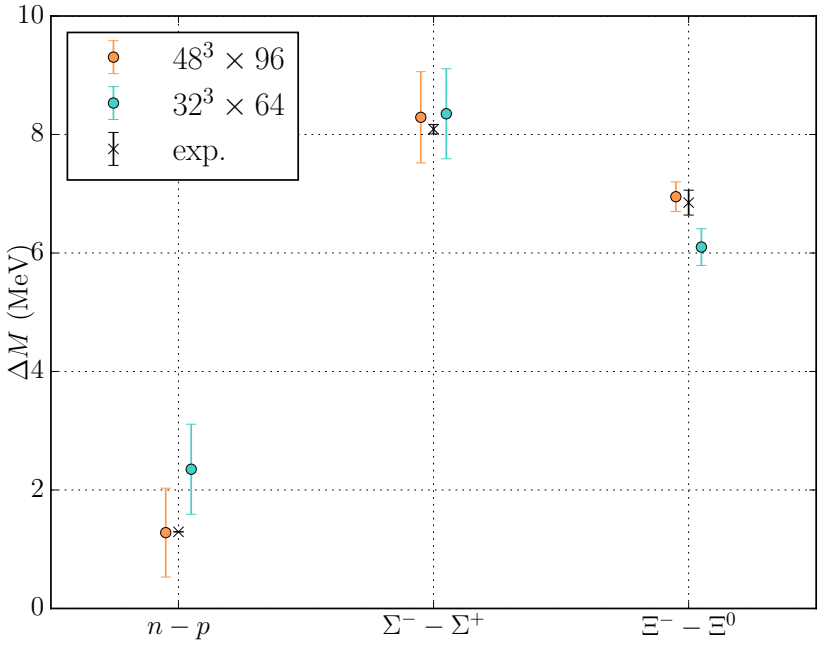


Figure 5.9: Octet mass splittings. This includes EM effects. The black crosses are experimental data. The coloured points are estimates generated from our lattice analysis.

with experimental determinations, however within the quoted uncertainties we observe that the absolute masses at the physical point are compatible with the experimental masses.

Nevertheless, it is possible that there is a systematic uncertainty that is causing an underestimate of the SU(3) breaking between these states. This systematic could be due to the fact that our simulations are performed at and around the SU(3) symmetric point where the Δ and Σ^* states are stable three quark states. However, in the physical system the Δ and Σ^* states are unstable and decay, e.g. to $\Delta \to \pi + N$, where the net mass of the $\pi + N$ system is significantly lower than the three quark state. The opening of these decay channels is certainly anticipated to affect the extrapolation to the physical point [47]. We possibly see the emergence of this feature in Fig. 5.7 by the difference between the two lattice ensembles 4 and 5 at lighter valence quark mass. Ensemble 4 is at the SU(3) point, while ensemble 5 has u, d quarks lighter than the SU(3) point. We do not expect the threshold effects to have a strong influence on the mass splitting generated by isospin symmetry violation.

Some selected splittings of phenomenological interest are tabulated in Tab. 5.7. The combination $\Delta^+ + +\Delta^- - (\Delta^+ + \Delta^0)$ eliminates the leading order strong

	Δ^{++}	Δ^{+}	Δ^{0}	Δ^{-}
This work	1304(59)(6)	1306(58)(6)	1308(57)(6)	1311(56)(6)
Cutkosky [49]	1231.43	1230.95	1233.40	1235.68
	Σ^{*+}	Σ^{*0}	Σ^{*-}	
This work	1425(38)(8)	1427(38)(8)	1431(37)(8)	
Experiment	1382.8	1383.7	1387.2	
	Ξ^{*0}	Ξ^{*-}	Ω	
This work	1542(26)(9)	1546(25)(9)	1656(21)(8)	
Experiment	1531.78	1535.2	1672.45	

Table 5.6: Predicted masses from the $48^3 \times 96$ ensembles for the decuplet baryons. All values quoted in MeV. The first error is statistical, while the second error is an estimate of the systematic error, which is obtained by the difference in the $32^3 \times 64$ results to those of the $48^3 \times 96$.

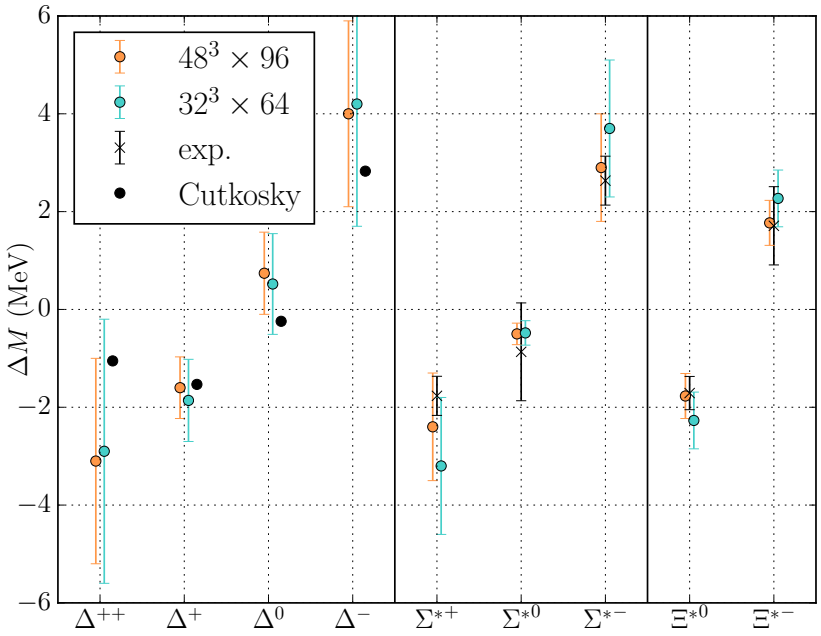


Figure 5.10: Decuplet mass splittings with the average isospin multiplet mass subtracted. The black crosses are experimental data, and for the Δ baryons, a fit to experimental data. The coloured points are estimates generated from our lattice analysis.

	$\Delta^{++}+\Delta^{-}-\Delta^{+}-\Delta^{0}$	$\Delta^{0}-\Delta^{++}$	$\Delta^{-}-\Delta^{++}+\frac{1}{3}\left(\Delta^{0}-\Delta^{+}\right)$
QED	1.7(14)(10)	-2.5(20)(13)	-2.7(26)(20)
QCD	-0.006(11)(6)	6.3(24)(5)	10.5(40)(10)
Total	1.7 (14)(10)	3.8(31)(5)	7.8(46)(5)
Cutkosky [49]	2.84–3.55	0.81–1.53	4.31–4.92
Exp./Pheno.	—	2.86(30) [50]	4.6(2) [51]
	$\Sigma^{*+}+\Sigma^{*-}-2\Sigma^{*0}$	$\Sigma^{*-}-\Sigma^{*+}$	$\Xi^{*-}-\Xi^{*0}$
QED	1.5(7)(1)	-0.8(11)(7)	0.61(51)(60)
QCD	-0.0032(56)(30)	6.1(22)(2)	2.92(98)(1)
Total	1.5(7)(1)	5.3(23)(10)	3.54(98)(8)
Cutkosky [49]	1.42	4.56	3.09
PDG [52]	2.6(21)	4.4(6)	3.2(6)

Table 5.7: Mass splittings for decuplet baryons. All values quoted in MeV. The first error is statistical, while the second error is an estimate of the systematic error, which is obtained by the difference in the $32^3 \times 64$ results to those of the $48^3 \times 96$.

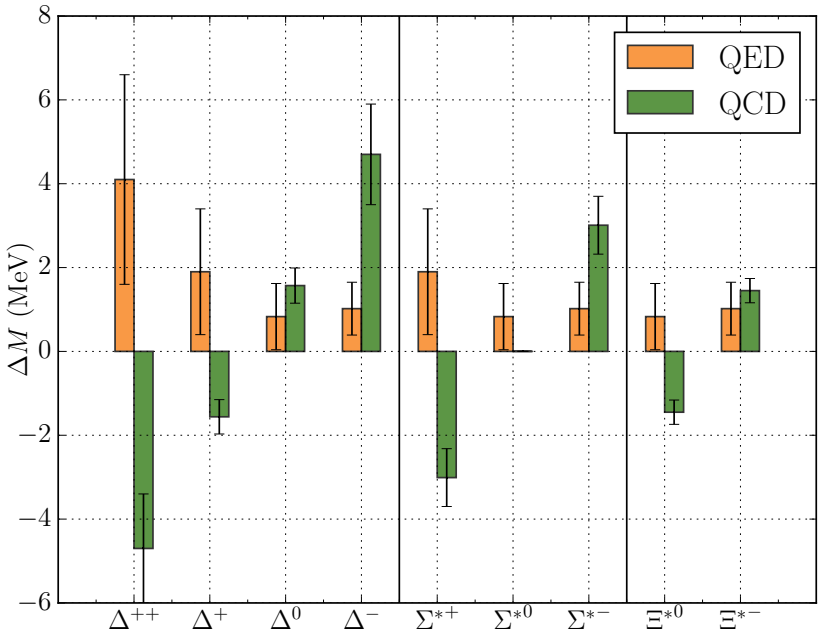


Figure 5.11: QCD and QED components of the decuplet mass splittings with the average isospin multiplet mass subtracted, for the $48^3 \times 96$ ensembles. The error bars represent symmetric statistical errors on the central value.

isospin violation, and hence isolates QED effects. The difference $\Delta^0 - \Delta^{++}$ has been reported by the PDG [53]. $\Delta^- - \Delta^{++} + \frac{1}{3}(\Delta^0 - \Delta^+)$ can be determined by considering the difference between π^- and π^+ cross sections on deuteron targets as found in [51]. The $\Sigma^{*+} + \Sigma^{*-} - 2\Sigma^{*0}$ combination removes leading order strong ispspin breaking components, giving a good estimate of electromagnetic effects. Here we've separated the mixed terms (γ^{EM}) into either QCD isospin breaking ($\delta m_u - \delta m_d$) or QED isospin breaking ($e_u - e_d$) and added them to the pure QCD or QED parts respectively. The Σ^* and Ξ^* splittings are both experimentally observed. Our results agree with the experimental results, typically within one standard deviation.

Fig. 5.10 shows the mass splittings at the physical quark mass for all decuplet baryons on both volumes. The mass splitting within each isospin multiplet is given as the difference of each mass from the average of its respective isospin multiplet. For the Σ^* and Ξ^* experimental values are available for us to compare with directly, these are shown in black crosses. The Δ baryons only have some values obtained experimentally, we are only able to compare to a fit of the data [54]. The Δ^- is not reported experimentally because it is difficult to study π^-–n scattering. Our lattice results agree with the experimental values. While the magnitude of the decuplet baryon masses were overestimated, possibly due to the resonance structure of the baryons, the splittings seem generally unaffected. In Fig. 5.11 we see how the splittings are generated from the QCD and QED parts of the particle mass. It is clear from this that there is significant uncertainty in the QED components, and obtaining greater precision in QED will be the focus of future studies.

5.7 Summary

We were able to produce estimates including both QCD and QED interactions for the octet and decuplet mass spectrum splittings. The estimates we produced agreed with experiment where it was available, and provided further insight where experimental data was not available.

In particular new insight was obtained in the break down of QCD and QED components in the octet spectrum splittings. The value of the updated total splitting (built on previous work by QCDSF [36]) agreed with experiment, and the predictions for the absolute value of the physical octet masses were within two standard deviations of the experimental values.

Similarly for the decuplet baryons we were able to produce estimates for the splittings using QCD+QED, and the first direct estimate for the Δ^-. Where experimental data was available our estimates agreed within two standard deviations. A breakdown of the splitting in terms of strong and electromagnetic components was included for the decuplet. It revealed that the QED components are a major

source of uncertainty in the splittings.

CHAPTER 6

Charmed Hadron Spectrum from Lattice QCD+QED

In the past 15 years interest in the charmed baryon spectrum has increased. This increase was triggered by the first observation of the doubly charmed baryons $\Xi_{cc}^{++}(3460), \Xi_{cc}^{+}(3443), \Xi_{cc}^{+}(3520), \Xi_{cc}^{++}(3541), \Xi_{cc}^{++}(3780)$ by SELEX [16]. The $\Xi_{cc}^{+}(3520)$ result was later confirmed by SELEX [55], however other experimental groups including BaBar [56] , Belle [57] and LHCb [58], were unable to detect $\Xi_{cc}^{+}(3520)$ or the other states. The observations have not been excluded because the production environments are different in each experiment [17].

If the states exist and have been properly identified, presumably the lower energy states are spin $J^P = \frac{1}{2}^+$ and higher states are $J^P = \frac{3}{2}^+$, this would lead to either electromagnetic charge splittings of $\sim 20\,\text{MeV}$ or strong isospin breaking effects of $\sim 60\,\text{MeV}$, both of which (from a theoretical standpoint) are unlikely. These splittings are some of the quantities we produce estimates for in our work using lattice QCD+QED. More recently a $\Xi_{cc}^{++}(3621)$ baryon has been detected and confirmed by LHCb [17], with unknown spin. However it is not clear if this observation is compatible with the SELEX observation of the $\Xi_{cc}^{+}(3520)$.

Theoretical calculations including lattice QCD calculations [59], [60], [61], [62],[63] have predicted the Ξ_{cc}^{+} mass to be $100 - 200\,\text{MeV}$ heavier than the experimental result $\Xi_{cc}^{+}(3520)$, with only ETMC producing the single estimate that agrees [64]. In this study we produce estimates for the mass splittings within the charmed multiplets using lattice QCD+QED. We also produce some estimates of mass splittings between different spin states and provide absolute mass values for the spectrum.

As the naming convention for the charmed baryons is a little confusing we have tabulated them in Tab. 6.1. The particles are arranged in columns of singly

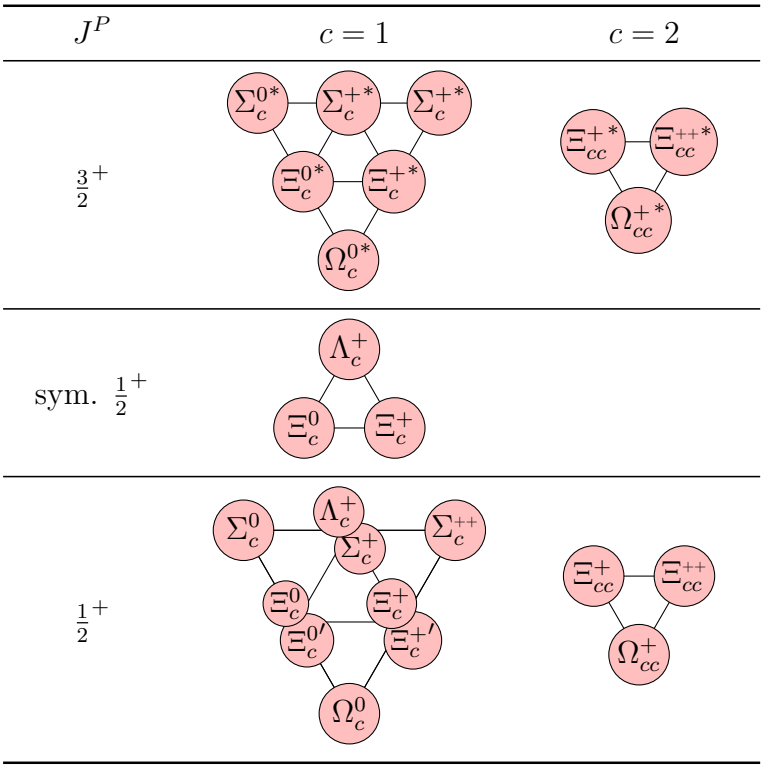


Table 6.1: The charmed baryon spectrum.

and doubly charmed baryons. Baryons on the top line of each diagram have zero strangeness, with the number of strange quarks increasing by one on each level below. Maximal isospin states are on the right of each row, with isospin decreasing by one from right to left. From this table we study all particles except the symmetric wave function spin $\frac{1}{2}$ triplet. We also study D mesons, which are not included in the table. These are singly charmed mesons with the other quark being either u, d or s.

Our approach is much the same as in chapter 5. We construct SU(3) flavour breaking expansions within hadron multiplets that transform under irreducible representations of SU(3). The coefficients in the flavour breaking expansions are dependent on the number of charm quarks within the hadron. This allows us to use just the physical charm quark, and no extrapolation of the charm quark to its physical mass is necessary. The physical charm quark is tuned to produce the physical $\eta_{c\bar{c}}$ mass, and when we tune we use the physical charm charge. The charm quark does not play a significant role as a sea quark [65, 63], and we do not include it in the sea. In this way we can consider the charm to be partially quenched.

The lattice specifics and finite volume corrections are the same as in the previous analysis (section 5.1 and section 5.3) and so I will not repeat it here. The

exception is that only a subset of ensembles are used. The lattices used are shown in Tab. 6.2.

#	β	e^2	V	$\kappa_u,\ +2/3$	$\kappa_d,\ -1/3$	$\kappa_s,\ -1/3$
1	5.50	1.25	$32^3 \times 64$	0.124362	0.121713	0.121713
2	5.50	1.25	$48^3 \times 96$	0.124362	0.121713	0.121713

Table 6.2: Summary of lattice details

This chapter is organised as follows. In section 6.1 I describe the interpolators used for the charmed hadrons. In section 6.2 we discuss the mass expansions we use to fit to the lattice data. In section 6.3 we discuss how the charm quark was implemented on the lattice, including the determination of its mass and charge. In section 6.4 we discuss how the mass spectra from the lattice were fit using the mass expansions, the procedure is similar to that used in the light hadron spectrum. The differences in procedure are highlighted. In section 6.5 we present the results of the analysis including determination of isospin violating splittings using lattice QCD+QED. Finally, I summarise the results of the chapter in section 6.6.

6.1 Lattice Interpolators

Hadron masses are computed from two-point correlation functions using conventional techniques, as described in section 3.8. In particular, for baryons we construct zero-momentum two-point functions as

$$C(t) = \sum_{\vec{x}} \mathrm{Tr}\,\Gamma \langle\, T\big(\chi(\vec{x},t)\bar{\chi}(0)\big)\,\rangle\ , \tag{6.1}$$

for some choice of baryon spin projection matrix, Γ. As the interpolators for spin $\frac{1}{2}$ baryons are more complicated in charm sector we have tabulated them in Tab. 6.3.

For spin $\frac{3}{2}^+$ baryons we choose an interpolating operator symmetric in spin that contains quarks $q_1 \in \{n,d,u,c\}$, $q_2 = c$ and $q_3 \in \{n,d,u\}$

$$\begin{aligned}\chi(\vec{x},t) = \frac{1}{\sqrt{3}}\epsilon^{abc}\Big[&\big(q_1^{aT}(\vec{x},t)\,C\gamma_-\,q_2^b(\vec{x},t)\big)q_3^c(\vec{x},t)\\ &+\big(q_2^{aT}(\vec{x},t)\,C\gamma_-\,q_3^b(\vec{x},t)\big)q_1^c(\vec{x},t)\\ &+\big(q_3^{aT}(\vec{x},t)\,C\gamma_-\,q_1^b(\vec{x},t)\big)q_2^c(\vec{x},t)\Big]\ ,\end{aligned} \tag{6.2}$$

where $\gamma_- = (\gamma_2 + i\gamma_1)/2$.

baryon	wave function
$\Sigma_c(ll'c)$	$\frac{1}{\sqrt{2}}\epsilon[(l'^T C\gamma_5 c)l + l^T C\gamma_5 c)l']$
$\Xi'_c(lsc)$	$\frac{1}{\sqrt{2}}\epsilon[(s^T C\gamma_5 c)l + l^T C\gamma_5 c)s]$
$\Omega_c(ssc)$	$\epsilon[(s^T C\gamma_5 c)l + l^T C\gamma_5 c)s]$
$\Lambda_c(ll'c)$	$\frac{1}{\sqrt{6}}\epsilon[2(l^T C\gamma_5 l')c + (l^T C\gamma_5 c)l' - (l'^T C\gamma_5 c)l]$
$\Xi_c(lsc)$	$\frac{1}{\sqrt{6}}\epsilon[2(l^T C\gamma_5 s)c + (l^T C\gamma_5 c)s - (s^T C\gamma_5 c)l]$
$\Xi_c(ccl)$	$\epsilon(c^T C\gamma_5 l)c$
$\Omega_c(ccs)$	$\epsilon(c^T C\gamma_5 s)c$

Table 6.3: Spin $+\frac{1}{2}$ charmed hadron interpolators, $l, l' \in \{u, d, n\}$. Though the s quark is degenerate with the d quark on the lattice, we include it to show the wave functions of Ξ and Ω.

QCD and QED are flavour blind, and the masses of the quarks change for the different ensembles, and within an ensemble as we partially quench. So, in the following we distinguish flavour by the electric charge. The $+0.632/3$ charged quark is identified as a charm quark, this charge choice is explained in more detail in section 6.3. The other quarks can be determined by the charge $0, -1/3, +2/3$ for n, d, u respectively.

6.2 Mass expansions

We could use the same strategy to extrapolate as in the previous chapter, only in this case expand around an SU(4) point. Points on the SU(4) trajectory are defined by $(m_u + m_d + m_s + m_c)^{lat} = (m_u + m_d + m_s + m_c)^{phys}$, and the SU(4) symmetric point occurs when all lattice quark masses are equal. The SU(4) symmetric point, lying along $\delta m_u + \delta m_d + \delta m_s + \delta m_c = 0$ from the physical point has an extrapolation distance which is larger for all quarks, when compared with its SU(3) counterpart.

We could instead define $\delta\mu_c$ from the light quark (u, d, s) SU(3) point, which maintains the $X_\pi^2 = \frac{1}{3}(M_{K^0}^2 + M_{K^+}^2 + M_{\pi^+}^2)$ constant property. In this case we would extrapolate along the $(m_u + m_d + m_s)^{lat} = (m_u + m_d + m_s)^{phys}$ path. As we typically only consider the light quarks $\{u, d, s\}$ to be important in the QCD vacuum, it also keeps the properties of the vacuum constant to first order during the extrapolation. From this expansion point, the extrapolation to the light quarks is easier at the expense of the charm, which is easier to generate. This idea has been explored in pure QCD [66].

As the extrapolation distance is higher when extrapolating from the SU(4) point and the SU(3) point, both of these methods require higher order expansions

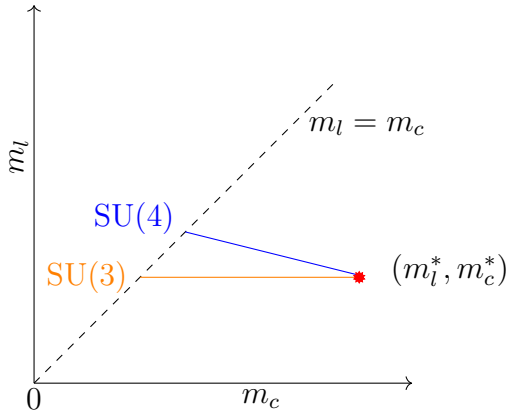


Figure 6.1: The transition from the SU(4) symmetric line to the physical point. The vertical axis represents the light quark masses $m_l = m_u = m_d = m_s$. The physical point is indicated by the red star. Extrapolation path for an SU(4) expansion is shown in blue, while an expansion from the SU(3) point is shown in orange. SU(4) symmetric points are along the dotted line, and the SU(3) and SU(4) symmetric points we use are at the intersection of (respectively) the orange and blue lines with dashed line. The fixed charm expansion point corresponds to the red star (physical point).

and hence more coefficients. Furthermore, because we can simulate at the physical charm mass, it is perhaps excessive to attempt an SU(4) or modified SU(3) extrapolation.

A safer route is to construct separate parameterisations for the different SU(3) irreducible representations shown in Tab. 6.1, in terms of spin and charm. The coefficients will then depend on how many charm quarks we include in the hadrons we fit with the expansion. The charm is fixed to the physical charm mass. This fixed charm approach is what we have chosen for our current work. The paths we've discussed are shown graphically in Fig. 6.1. In this diagram the physical point (red star) corresponds to the fixed charm approach while the blue and orange lines correspond to SU(4) and SU(3) expansions respectively.

There are several advantages to the fixed charm approach. Firstly, we know that the spectrum will satisfy the expansion symmetries in u, d, s flavours in a similar way as in chapter 5 and will be primarily linear. We would not expect the charmed extrapolation to behave as linearly as the light spectrum, because SU(4) flavour symmetry is not a good approximation, hence there is significantly more uncertainty in the order of expansion required for a combined fit.

Another advantage of the fixed charm approach is the charm quark can be simulated easily at its physical point and so we do not have to provide intermediate charm quark values; we have to produce and analyse less data. If we were to

attempt an SU(4) or similar combined fit with charm extrapolation we would need data on how the spectrum changes on its way to the charm physical value.

A disadvantage of the fixed charm approach is that we can no longer fit all the particles together. We require more quark propagators to obtain the same level of accuracy on the coefficients in each charm expansion. However, because we only need to extrapolate a small distance, less coefficients are required. Additionally, parameters can be excluded due to the $\delta\mu_c = 0$ condition and the charms static presence in the expansion, which is addressed presently.

The charmed meson expansion is slightly different from the one given for the S_3 symmetric quarks, Eq. (5.8). We will not assume that the charm quark satisfies the symmetry, but also we will set $\delta\mu_c = 0$ as we are already at the charm physical point. The expansion is given by,

$$\begin{aligned} M^2(a\bar{c}) &= M_0^2 + \alpha(\delta\mu_a) + \beta_1(\delta\mu_a^2) + \beta_1^{EM}(e_a^2 + e_c^2) + \beta_2^{EM} e_a e_c \\ &+ \gamma_1^{EM}(e_a^2\delta\mu_a) + \gamma_2^{EM}\delta\mu_a e_a e_c. \end{aligned} \tag{6.3}$$

Where here the c represents the charm quark, while $a \in \{u, d, n\}$. The γ_3^{EM} coefficient is excluded because $e_c^2\delta\mu_a$ looks just like $\delta\mu_a$. β_2 is not written because $\delta\mu_c\delta\mu_a = 0$. We have not written the sea quark parameters either, in this analysis because we are only using one ensemble for each volume.

As the charm quark is the non-symmetric quark in the spin $\frac{1}{2}$ wave function and because e_c is unchanging and $\delta\mu_c = 0$, the S_3 breaking terms are absorbed into other terms. This reduces the singly charmed spin $\frac{1}{2}$ expansion to the same form as the spin $\frac{3}{2}$. At the same order of the expansion as Eq. (6.3), we write the expressions for the spin $\frac{1}{2}$ and spin $\frac{3}{2}$ singly charmed baryons:

$$\begin{aligned} M(abc) &= M_0 + \alpha_1(\delta\mu_a + \delta\mu_b) + \beta_1(\delta\mu_a^2 + \delta\mu_b^2) + \beta_2(\delta\mu_a\delta\mu_b) \\ &+ \beta_1^{EM}(e_a^2 + e_b^2 + e_c^2) + \beta_2^{EM}(e_a e_b + e_c e_a + e_c e_b) + \beta_3^{EM}(e_a + e_b)e_c \\ &+ \gamma_1^{EM}(e_a^2\delta\mu_a + e_b^2\delta\mu_b) + \gamma_2^{EM}(\delta\mu_a e_a(e_b + e_c) + \delta\mu_b e_b(e_a + e_c)) \\ &+ \gamma_3^{EM}(\delta\mu_a e_b e_c + \delta\mu_b e_a e_c) + \gamma_4^{EM}(\delta\mu_a(e_b^2 + e_c^2) + \delta\mu_b(e_a^2 + e_c^2)) \\ &+ \gamma_5^{EM}(\delta\mu_a e_a + \delta\mu_b e_b)e_c. \end{aligned} \tag{6.4}$$

The additional terms β_3^{EM} and γ_5^{EM} are included because the mass of the particle is not considered to respond in the same way to changes in charm quark parameters. This is because we expand around the physical charm point, where the charm quark is not considered symmetric with the other quarks.

For doubly charmed baryons the expansion simplifies in a similar way to the

meson expansion,

$$\begin{aligned} M(acc) &= M_0 + \alpha_1(\delta\mu_a) + \beta_1(\delta\mu_a^2) \\ &+ \beta_1^{EM}(e_a^2 + 2e_c^2) + \beta_2^{EM}(2e_a e_c + e_c^2) + \gamma_1^{EM}(e_a^2\delta\mu_a) \\ &+ \gamma_2^{EM}\delta\mu_a e_a(2e_c), \end{aligned} \tag{6.5}$$

β_2, γ_3^{EM} and γ_4^{EM} coefficients have been absorbed into other coefficients.

6.3 Charm quark implementation

The lattice spacing of our lattice provides a maximum cut-off energy for propagating modes. Higher energy static particles can still be implemented, however these particles do not propagate in the same way. The charm quark has a mass of approximately ~ 1.3 GeV, which is less than half our lattice spacing $\sim 2.8\,\text{GeV}$. This means the charm quark can be implemented in a similar way to the light quarks, and no special action is required. The charm quark kinetic energy is limited to $\sim 100\,\text{MeV}$, and given we study the zero momentum charmed baryons, we do not expect states with above 100 MeV energy difference to significantly contribute. The charmed quark mass was tuned to reproduce the physical $\eta_{c\bar{c}}$ mass. To ensure this mass was determined correctly, the physical charm charge was used. To obtain this charge the charm quark charge was scaled by $1/\sqrt{10}$ to counter the greater than physical α_{QED}. The charge used was $+0.632/3$. The $\eta_{c\bar{c}}$ has mass $(aM^*_{c\bar{c}})^2 = 1.0522$ in lattice units, and we use the κ value corresponding to this number in Fig. 6.2 which we obtain from fitting the points, $\kappa_c = 0.113026$.

In treating the charm quark in this way we have accurately obtained the κ value, but we have reduced its charge to the physical value. It is conceivable that this may effect the precision of the charge splittings. The charge distribution will be different than the real hadrons, with interactions being typically weaker. Due to the extrapolation method we use, we do not suspect this has had a significant effect. Though it would be interesting to see how the results would change if we had applied a different strategy. We will explain some different strategies to address this problem, and then explain why our extrapolation method is resistant to these effects.

Firstly, we could sacrifice some accuracy in determining the κ value by determining the $\eta_{c\bar{c}}$ with unphysical charm charge of $+2/3$ and $\alpha_{QED} \approx 10\alpha^*_{QED}$. The κ value obtained is then compatible with the higher charge quark, and we can use the higher charge when partially quenching. Using this approach, total mass of the charmed hadrons would be slightly incorrect, however we would have the additional resolution on the QED effects. Secondly, we could use a more complicated approach, which provides the best of both cases. Produce $c\bar{c}$ plot for $+0.632/3$

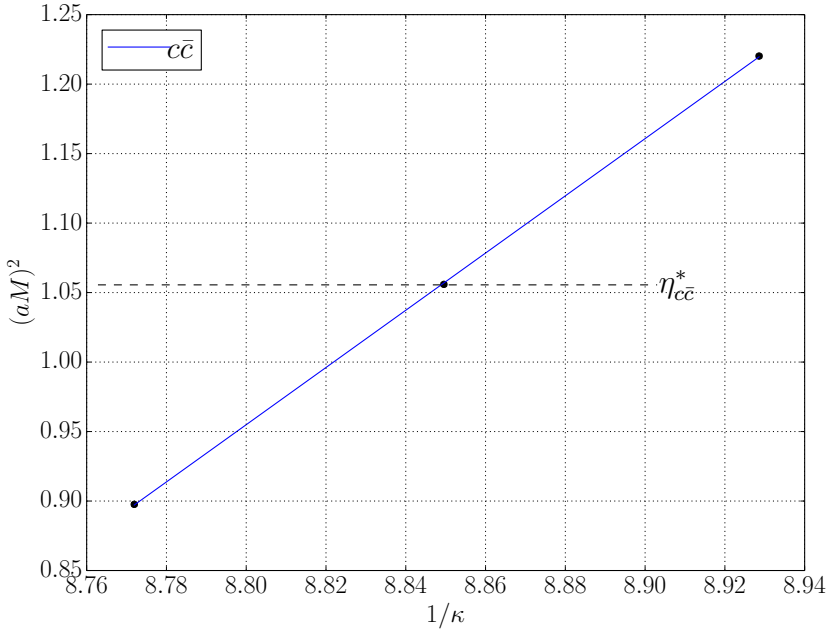


Figure 6.2: The pseudoscalar $c\bar{c}$ meson mass as a function of $\frac{1}{\kappa}$.

and determine m_c for the physical charge $0.632/3$, then use this quark mass in the $+2/3$ charged case.

The method we use to extrapolate to the physical point provides a different expansion for each sector of charm quark in a hadron. With the current method for determining quark masses, the fitting algorithm still has access to the high charge region through highly charged quarks u, d. For each charm sector, the fitting algorithm sees the same changes in charge irrespective of the charmed charge, because the level of charm does not change within each set fit. The total charge of the hadrons is of course lower, and this may lead to some lack of sensitivity to QED effects.

If we were to fit all charm sectors of charm together using flavour breaking expansions about an SU(4) or SU(3) point (expansions including the charm quark), such an analysis would be sensitive to the charm charge. However, we do not use either of these extrapolation methods in this work.

One of the mysteries we are trying to solve involves the absolute masses as well as the splittings of the two different spin types Ξ_{cc}^{++} and Ξ_{cc}^{+}, and our current method is best suited for this because we have obtained the most accurate estimate of κ for the charm quark possible, while not effecting the splittings.

Treating the charm quark as we have does not introduce additional difficulty when determining the coefficients of the expansions. The hadrons in the spectrum we produce from the lattice are not the real hadrons, rather they are points in

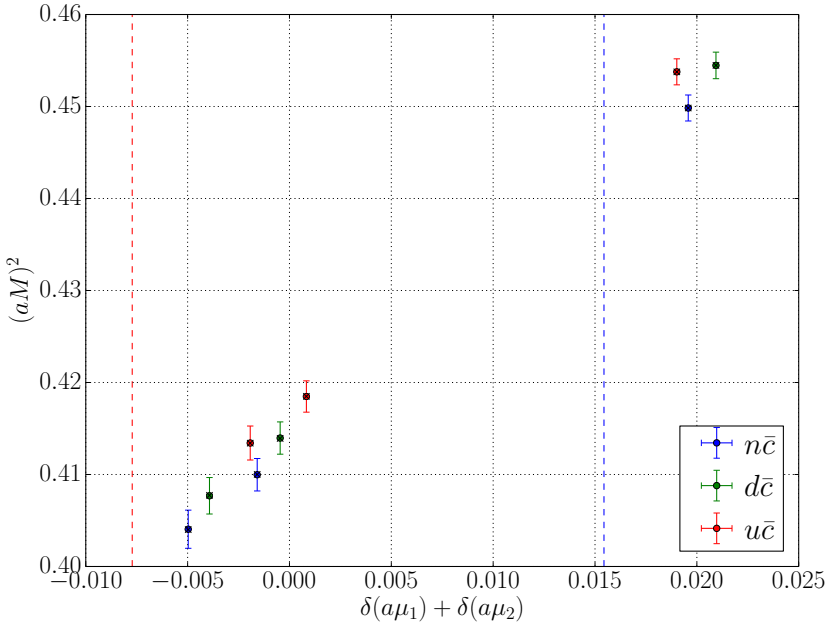


Figure 6.3: The mass spectrum of the D mesons in lattice units as a function of the displacement of the quark masses from the symmetric point value in the Dashen Scheme. The black crosses show what the expansion predicts.

parameter space which we use to fit the coefficients of our expansion. This is why it is acceptable to use an unphysical α_{QED} at all. While fitting the expansion coefficients, we use the charge $+0.632/3$ for the charm quark, however once the coefficients are determined we wish to make contact with the physical charm particle, and once the electromagnetic coefficients are corrected for α_{QED} this charge is $+2/3$.

6.4 Fitting procedure

The quark masses at the physical point are determined using un-charmed mesons, the values are in Tab. 5.3 and the method is as described in section 5.5. Once these are determined we can plot the charmed mass spectrum against the Dashen scheme $\delta\mu$. Fig. 6.3 and Fig. 6.4 show the lattice mass spectrum for the singly charmed pseudoscalar meson spectrum and spin $\frac{1}{2}^+$ Λ_c and Ξ_c baryon spectrum. Fig. 6.5 and Fig. 6.6 show the singly and doubly charmed the spin $\frac{1}{2}^+$ Σ_c, Ξ_c and Ξ_{cc} baryon spectrum. Fig. 6.7 and Fig. 6.8 show the spin $\frac{3}{2}^+$ Σ_c, Ξ_c and Ξ_{cc} baryons spectrum.

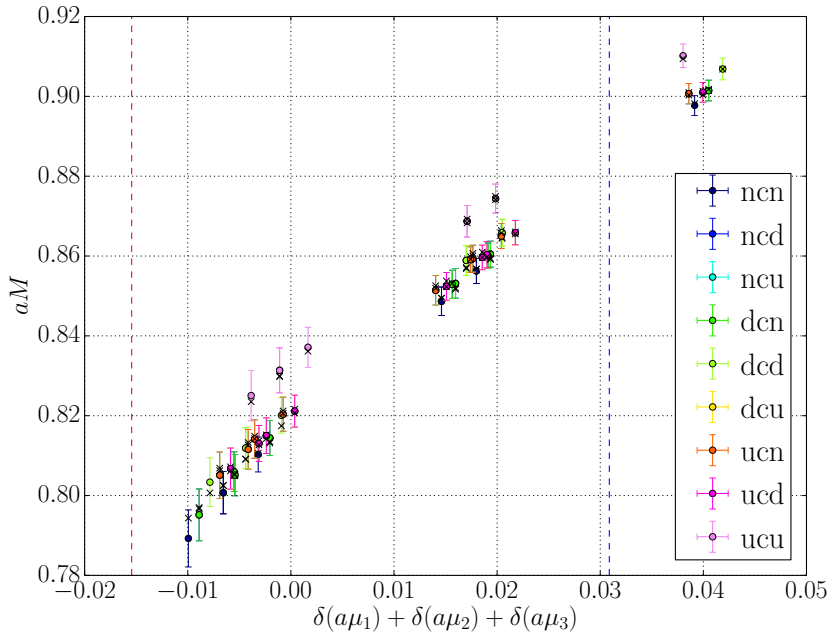


Figure 6.4: The mass spectrum of the spin $\frac{1}{2}^+$ singly charmed Λ_c and Ξ_c baryons in lattice units as a function of the displacement of the quark masses from the symmetric point value in the Dashen scheme. The black crosses show what the expansion predicts.

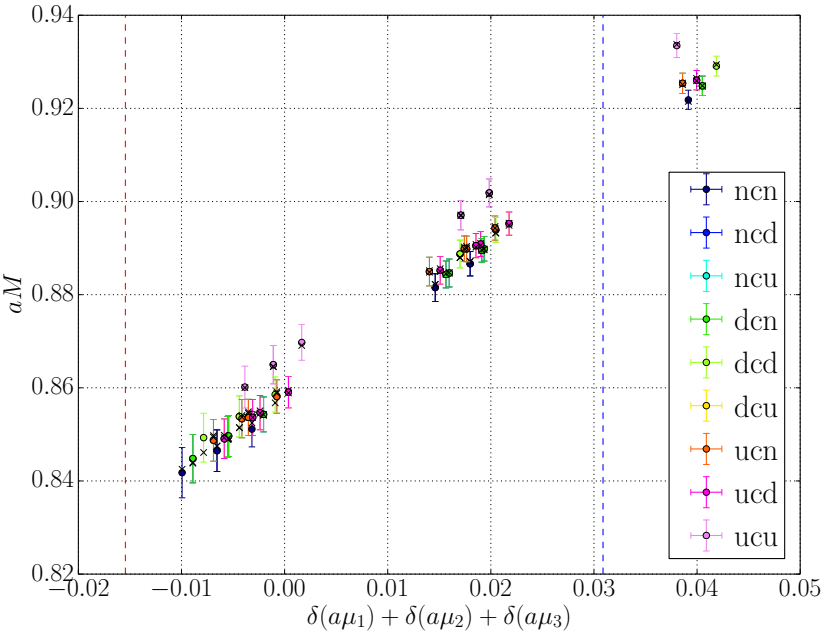


Figure 6.5: The mass spectrum of the spin $\frac{1}{2}^+$ singly charmed Σ_c, Ξ_c and Ω_c baryons in lattice units as a function of the displacement of the quark masses from the symmetric point value in the Dashen scheme.

J	D meson 0	Λ_c, Ξ_c $\frac{1}{2}$	$\Sigma_c, \Xi_c', \Omega_c$ $\frac{1}{2}$	Ξ_{cc}, Ω_{cc} $\frac{1}{2}$	$\Sigma_c^*, \Xi_c^*, \Omega_c^*$ $\frac{3}{2}$	$\Xi_{cc}^*, \Omega_{cc}^*$ $\frac{3}{2}$
M_0	0.4123(16)	0.8176(41)	0.8570(36)	1.2247(22)	0.8926(63)	1.2497(44)
α	1.8(13)	2.30(24)	1.51(18)	1.62(15)	1.37(29)	1.69(29)
β_1	4.5(53)	-8.1(83)	6.8(61)	-1.3(60)	5.0(10)	-9(12)
β_2	–	-10.5(40)	-0.7(22)	–	1.8(48)	–
β_1^{EM}	0.000965(71)	0.00061(27))	0.00037(17)	0.000777(81)	0.00024(34)	0.00096(14)
β_2^{EM}	-0.00067(42)	0.00057(30))	0.00041(20)	0.00012(18)	-0.00019(38)	0.0005(47)
β_3^{EM}	–	0.00036(14)	0.0002(11)	–	0.00059(22)	–
γ_1^{EM}	0.0065(16)	-0.0082(78))	0.0085(53)	-0.0007(30)	0.004(10)	-0.0031(51)
γ_2^{EM}	0.0081(77)	0.0018(47)	0.0063(30)	0.0116(65)	0.0094(63)	-0.002(15)
γ_3^{EM}	–	-0.048(15)	-0.019(10)	–	-0.015(27)	–
γ_4^{EM}	–	0.0114(93)	0.0133(63)	–	0.027(12)	–
γ_5^{EM}	–	0.0082(27)	0.0012(26)	–	0.0093(47)	–
χ^2	0.00409	5.65	2.80	0.00234	5.49	0.00768
DOF	11	69	69	2	69	2
χ^2/DOF	0.000372	0.0819	0.0417	0.00117	0.0796	0.00384
χ^2_{corr}/DOF[1]	84.01	8.94	9.07	149.91	3.54	52.00

Table 6.4: Coefficients obtained from fitting the mass spectrum of the various particle types.

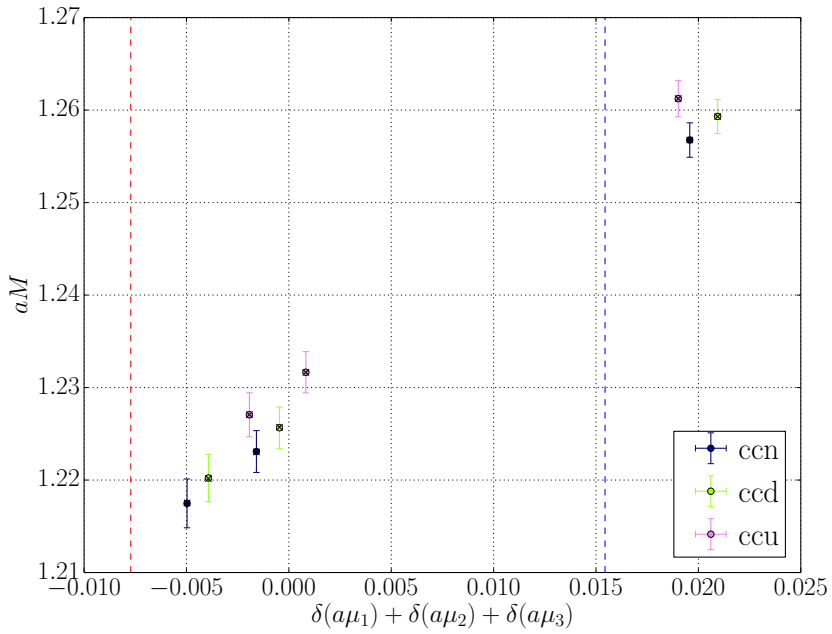


Figure 6.6: The mass spectrum of the spin $\frac{1}{2}^+$ Ξ_{cc} and Ω_{cc} baryons in lattice units as a function of the displacement of the quark masses from the symmetric point value. The black crosses show what the expansion predicts.

Using the expansion Eq. (6.3) we can fit to the charmed meson spectrum. The black crosses show our estimates for these masses using the symmetry breaking expansion. We can repeat this for the baryons using Eq. (6.4) to the singly-charmed spectra and Eq. (6.5) fit the doubly-charmed spectra. The coefficients obtained from these fits are tabulated in Tab. 6.4. The electromagnetic coefficients have been corrected for unphysical α_{QED} using Eq. (5.12). Using the physical quark masses determined from Tab. 5.3, we can then produce estimates for the physical hadron masses.

We have also included in Tab. 6.4 the correlated and uncorrelated χ^2/DOF. As discussed in section 5.5, the χ^2 is the cost function we minimise on to produce the fit parameters. We may choose to consider the correlations between the data points or minimise excluding these correlations. To be clear, we did not use the correlated χ^2_{corr} to produce the parameters, it is only included here to give additional information on whether we have over or under parameterised our model.

As the correlated χ^2_{corr} values are greater than one, it suggests that (in some respect) we have under parameterised our model. However, in this case, excluding

[1]Correlated χ^2 is not used to determine parameters, but is included for reference to provide a guide to whether our method is over or under parameterised.

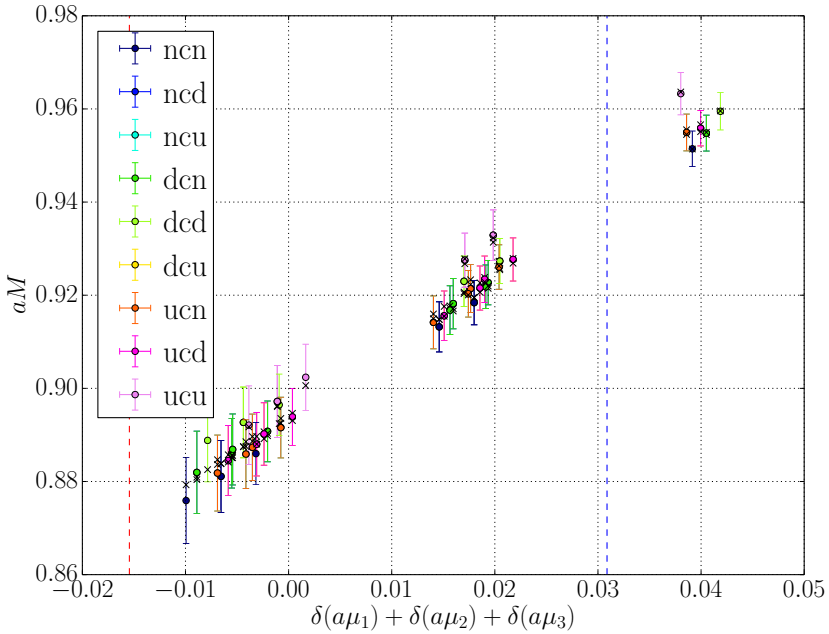


Figure 6.7: The mass spectrum of the spin $\frac{3}{2}^+$ singly charmed Σ_c, Ξ_c and Ω_c baryons in lattice units as a function of the displacement of the quark masses from the symmetric point value.

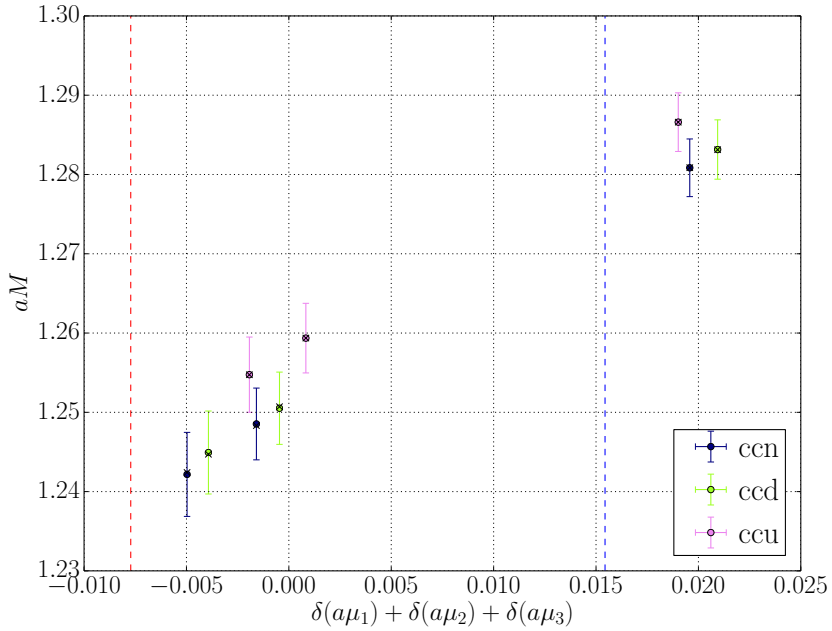


Figure 6.8: The mass spectrum of the spin $\frac{3}{2}^+$ Ξ_{cc} and Ω_{cc} baryons in lattice units as a function of the displacement of the quark masses from the symmetric point value. The black crosses show what the expansion predicts.

M_0	1.00	-0.07	0.02	-0.05	-0.48	0.58	0.36
α	-0.07	1.00	-0.98	0.11	-0.38	-0.20	0.07
β_1	0.02	-0.98	1.00	-0.11	0.48	0.17	-0.13
β_1^{EM}	-0.05	0.11	-0.11	1.00	-0.57	-0.68	0.66
β_2^{EM}	-0.48	-0.38	0.48	-0.57	1.00	-0.02	-0.86
γ_1^{EM}	0.58	-0.20	0.17	-0.68	-0.02	1.00	-0.03
γ_2^{EM}	0.36	0.07	-0.13	0.66	-0.86	-0.03	1.00

Figure 6.9: Correlation matrix of the parameters in the fitting model for the D mesons. The label names along the left hand side are also valid along the top. The matrices are colour coded to guide the eye. Red is close to -1, green is close to 1 and the colour scales through to white at 0.

	M_0	α	β_1	β_2	β_1^{EM}	β_2^{EM}	β_3^{EM}	γ_1^{EM}	γ_2^{EM}	γ_3^{EM}	γ_4^{EM}	γ_5^{EM}
M_0	1.00	-0.89	0.87	0.72	-0.79	-0.25	0.32	0.79	0.48	-0.29	0.66	-0.61
α	-0.89	1.00	-0.99	-0.90	0.62	0.15	-0.15	-0.62	-0.39	0.05	-0.57	0.60
β_1	0.87	-0.99	1.00	0.86	-0.61	-0.18	0.19	0.62	0.42	-0.10	0.59	-0.56
β_2	0.72	-0.90	0.86	1.00	-0.40	0.06	-0.08	0.36	0.13	0.14	0.32	-0.46
β_1^{EM}	-0.79	0.62	-0.61	-0.40	1.00	0.75	-0.78	-0.89	-0.86	0.68	-0.94	0.29
β_2^{EM}	-0.25	0.15	-0.18	0.06	0.75	1.00	-0.97	-0.54	-0.90	0.80	-0.87	-0.27
β_3^{EM}	0.32	-0.15	0.19	-0.08	-0.78	-0.97	1.00	0.56	0.89	-0.84	0.87	0.25
γ_1^{EM}	0.79	-0.62	0.62	0.36	-0.89	-0.54	0.56	1.00	0.79	-0.45	0.77	-0.58
γ_2^{EM}	0.48	-0.39	0.42	0.13	-0.86	-0.90	0.89	0.79	1.00	-0.69	0.89	-0.12
γ_3^{EM}	-0.29	0.05	-0.10	0.14	0.68	0.80	-0.84	-0.45	-0.69	1.00	-0.74	-0.29
γ_4^{EM}	0.66	-0.57	0.59	0.32	-0.94	-0.87	0.87	0.77	0.89	-0.74	1.00	-0.06
γ_5^{EM}	-0.61	0.60	-0.56	-0.46	0.29	-0.27	0.25	-0.58	-0.12	-0.29	-0.06	1.00

Figure 6.10: Correlation matrix of the parameters in the fitting model for the $J^P = \frac{1}{2}^+$ Λ_c, Ξ_c baryons. The label names along the left hand side are also valid along the top. The matrices are colour coded to guide the eye. Red is close to -1, green is close to 1 and the colour scales through to white at 0.

	M_0	α	β_1	β_2	β_1^{EM}	β_2^{EM}	β_3^{EM}	γ_1^{EM}	γ_2^{EM}	γ_3^{EM}	γ_4^{EM}	γ_5^{EM}
M_0	1.00	-0.80	0.71	0.70	-0.35	-0.27	0.38	0.64	0.64	-0.10	0.39	-0.08
α	-0.80	1.00	-0.98	-0.78	0.60	0.39	-0.43	-0.82	-0.70	0.43	-0.51	0.28
β_1	0.71	-0.98	1.00	0.65	-0.60	-0.36	0.40	0.79	0.67	-0.46	0.48	-0.30
β_2	0.70	-0.78	0.65	1.00	-0.43	-0.31	0.30	0.67	0.51	-0.25	0.42	-0.26
β_1^{EM}	-0.35	0.60	-0.60	-0.43	1.00	0.85	-0.91	-0.89	-0.55	0.80	-0.95	-0.32
β_2^{EM}	-0.27	0.39	-0.36	-0.31	0.85	1.00	-0.88	-0.70	-0.69	0.86	-0.91	-0.52
β_3^{EM}	0.38	-0.43	0.40	0.30	-0.91	-0.88	1.00	0.75	0.47	-0.67	0.98	0.62
γ_1^{EM}	0.64	-0.82	0.79	0.67	-0.89	-0.70	0.75	1.00	0.71	-0.65	0.80	0.06
γ_2^{EM}	0.64	-0.70	0.67	0.51	-0.55	-0.69	0.47	0.71	1.00	-0.70	0.55	-0.08
γ_3^{EM}	-0.10	0.43	-0.46	-0.25	0.80	0.86	-0.67	-0.65	-0.70	1.00	-0.76	-0.12
γ_4^{EM}	0.39	-0.51	0.48	0.42	-0.95	-0.91	0.98	0.80	0.55	-0.76	1.00	0.49
γ_5^{EM}	-0.08	0.28	-0.30	-0.26	-0.32	-0.52	0.62	0.06	-0.08	-0.12	0.49	1.00

Figure 6.11: Correlation matrix of the parameters in the fitting model for the $J^P = \frac{1}{2}^+$ $\Sigma_c, \Xi_c', \Omega_c$ baryons. The label names along the left hand side are also valid along the top. The matrices are colour coded to guide the eye. Red is close to -1, green is close to 1 and the colour scales through to white at 0.

	M_0	α	β_1	β_2	β_1^{EM}	β_2^{EM}	β_3^{EM}	γ_1^{EM}	γ_2^{EM}	γ_3^{EM}	γ_4^{EM}	γ_5^{EM}
M_0	1.00	-0.34	0.05	0.22	-0.24	-0.09	0.11	0.24	-0.01	0.27	0.20	-0.20
α	-0.34	1.00	-0.89	-0.36	0.64	0.29	-0.45	-0.69	-0.47	0.15	-0.60	0.28
β_1	0.05	-0.89	1.00	0.02	-0.68	-0.37	0.63	0.57	0.55	-0.47	0.72	-0.12
β_2	0.22	-0.36	0.02	1.00	0.12	0.13	-0.34	0.10	0.06	0.48	-0.21	-0.31
β_1^{EM}	-0.24	0.64	-0.68	0.12	1.00	0.75	-0.88	-0.78	-0.69	0.60	-0.96	-0.21
β_2^{EM}	-0.09	0.29	-0.37	0.13	0.75	1.00	-0.87	-0.33	-0.82	0.36	-0.75	-0.63
β_3^{EM}	0.11	-0.45	0.63	-0.34	-0.88	-0.87	1.00	0.46	0.70	-0.59	0.94	0.52
γ_1^{EM}	0.24	-0.69	0.57	0.10	-0.78	-0.33	0.46	1.00	0.48	-0.27	0.64	-0.29
γ_2^{EM}	-0.01	-0.47	0.55	0.06	-0.69	-0.82	0.70	0.48	1.00	-0.40	0.71	0.14
γ_3^{EM}	0.27	0.15	-0.47	0.48	0.60	0.36	-0.59	-0.27	-0.40	1.00	-0.68	-0.13
γ_4^{EM}	0.20	-0.60	0.72	-0.21	-0.96	-0.75	0.94	0.64	0.71	-0.68	1.00	0.25
γ_5^{EM}	-0.20	0.28	-0.12	-0.31	-0.21	-0.63	0.52	-0.29	0.14	-0.13	0.25	1.00

Figure 6.12: Correlation matrix of the parameters in the fitting model $J^P = \frac{3}{2}^+$ $\Sigma_c^*, \Xi_c^*, \Omega_c^*$ baryons. The label names along the left hand side are also valid along the top. The matrices are colour coded to guide the eye. Red is close to -1, green is close to 1 and the colour scales through to white at 0.

	M_0	α	β_1	β_1^{EM}	β_2^{EM}	γ_1^{EM}	γ_2^{EM}
M_0	1.00	-0.65	0.66	0.51	-0.36	-0.56	0.60
α	-0.65	1.00	-0.99	-0.18	0.03	0.51	-0.32
β_1	0.66	-0.99	1.00	0.15	-0.06	-0.50	0.35
β_1^{EM}	0.51	-0.18	0.15	1.00	0.12	-0.58	0.04
β_2^{EM}	-0.36	0.03	-0.06	0.12	1.00	0.31	-0.91
γ_1^{EM}	-0.56	0.51	-0.50	-0.58	0.31	1.00	-0.46
γ_2^{EM}	0.60	-0.32	0.35	0.04	-0.91	-0.46	1.00

Figure 6.13: Correlation matrix of the parameters in the fitting model $J^P = \frac{1}{2}^+$ Ξ_{cc}, Ω_{cc} baryons. The label names along the left hand side are also valid along the top. The matrices are colour coded to guide the eye. Red is close to -1, green is close to 1 and the colour scales through to white at 0.

M_0	1.00	-0.13	0.09	0.01	0.10	-0.18	-0.17
α	-0.13	1.00	-0.97	-0.64	0.15	0.89	-0.56
β_1	0.09	-0.97	1.00	0.58	-0.15	-0.90	0.57
β_1^{EM}	0.01	-0.64	0.58	1.00	0.50	-0.77	-0.16
β_2^{EM}	0.10	0.15	-0.15	0.50	1.00	-0.06	-0.84
γ_1^{EM}	-0.18	0.89	-0.90	-0.77	-0.06	1.00	-0.26
γ_2^{EM}	-0.17	-0.56	0.57	-0.16	-0.84	-0.26	1.00

Figure 6.14: Correlation matrix of the parameters in the fitting model of the $J^P = \frac{3}{2}^+$ $\Xi^*_{cc}, \Omega^*_{cc}$ baryons. The label names along the left hand side are also valid along the top. The matrices are colour coded to guide the eye. Red is close to -1, green is close to 1 and the colour scales through to white at 0.

the D mesons, none of the fits seemed to converge to the points. I think the main issue is that, because correlations between the points were included, the significance of changes in points as a result of a change in independent variable (charge for instance) was reduced in value when using the correlated metric. This lead to the fitting algorithm not finding the correct minimum because it did not have enough unique data points. Given that the fit did not converge, the validity of the χ^2_{corr}/DOF as a measure of the goodness of fit is reduced. To remedy this in the future we may need to use less correlated data points, or consider an SU(4) or modified SU(3) extrapolation[2]to better constrain the parameters in the correlated case .

The errors on the parameters in Tab. 6.4 are determined from conducting the fit for each bootstrap ensemble and then taking the standard deviation. We also provide normalised correlation matrices for the parameters of the functions fit to the particle mass spectra of the D mesons (Fig. 6.9), Λ_c, Ξ_c (Fig. 6.10), $\Sigma_c, \Xi'_c, \Omega_c$ (Fig. 6.11), $\Sigma^*_c, \Xi^*_c, \Omega^*_c$ (Fig. 6.12) , Ξ_{cc}, Ω_{cc} (Fig. 6.13), $\Xi^*_{cc}, \Omega^*_{cc}$ (Fig. 6.14).

We can see from the correlation matrices that certain parameters are typically not determined well by the data set. For instance the data set is very linear, and so it is difficult to determine coefficients that are proportional to $(\delta\mu)^2$. As such, there is high correlation between the α coefficient and the β_i coefficients.

The cross correlations tend to form two groups, terms with electromagnetic effects and terms without. Providing data that the fitting algorithm can break down into parameters easily reduces the uncertainty on the value of the parameter. The only way to improve the fit is to obtain data that allows one to distinguish between the different degrees of freedom, alternatively one could reduce the degrees of freedom. In the future the analysis could be optimised to reduce covariance in the parameters.

[2]As discussed in section 6.2.

6.5 Results

Having determined expansion coefficients and the physical quark mass parameters, we can then obtain estimates of the charmed hadron spectrum at the physical point. The masses of the charmed particles are tabulated in Tab. 6.5. The first uncertainty in the mass splitting is statistical, while the second provides an estimate of the systematic uncertainties, which are obtained by the mass difference of the $48^3 \times 96$ results compared with the $32^3 \times 64$. The central value is obtained from the larger volume. We note that since our simulations are performed at only a single value of the lattice spacing, no continuum extrapolation is possible. In the literature these effects seem to be below statistical and systematic errors. Collaborations which include continuum extrapolation [67, 62, 63] and those that do not, have results that tend to agree [60], which suggests to us that this correction will have only a minor impact on the absolute value. The effect of continuum extrapolation on isospin splittings is mostly unknown because previous studies that include continuum extrapolation have had difficulty resolving isospin splittings [67, 61].

We can see in Tab. 6.5 that our results agree with experiment within one standard deviation when both systematic and statistical error is taken into account. A portion of this systematic uncertainty can be attributed to uncertainty in the scale, which is approximately 2%. For the doubly charmed particles we have not tabulated the experimental values because the states proposed by SELEX [16] and the LHCb [17] do not have a spin associated with them. We suggest from our results that the SELEX value $\Xi_{cc}^{+}(3520)$ fits well with the Ξ_{cc}^{+} $J^P = \frac{1}{2}^+$ while the LHCb estimate $\Xi_{cc}^{++}(3621)$ is closest to Ξ_{cc}^{++*} $J^P = \frac{3}{2}^+$. We suspect that our results tabulated in Tab. 6.5 for the $48^3 \times 96$ volume are systematically low because the $32^3 \times 64$ values were all higher than their $48^3 \times 96$ counterparts, and this is reflected in the second error on the values. Additionally, the centre values of the $48^3 \times 96$ are low compared to other collaborations [59], [60], [61], [62],[63], while the centre values of the $32^3 \times 96$ closely agree with those collaborations. Hence, the systematic errors give an indication of how our values compare to other collaborations.

The splitting of the charmed hadron masses is shown in Fig. 6.15. We see in this figure that after removing the average mass, the splitting between the different charmed SU(3) multiplets is statistically consistent with the experimental values, where there are experimental estimates. With the exception of the $\Xi_{cc}^{*} - \Omega_{cc}^{*}$, the two volumes agree on predictions within one standard deviation which is what should be expected if the finite volume effects are under control. Considering statistical errors only, the $\Xi_{cc}^{*} - \Omega_{cc}^{*}$ splittings of the two volumes do not seem to agree, even within two standard deviations.

The splitting of the isospin multiplets is shown in Fig. 6.16, with the average

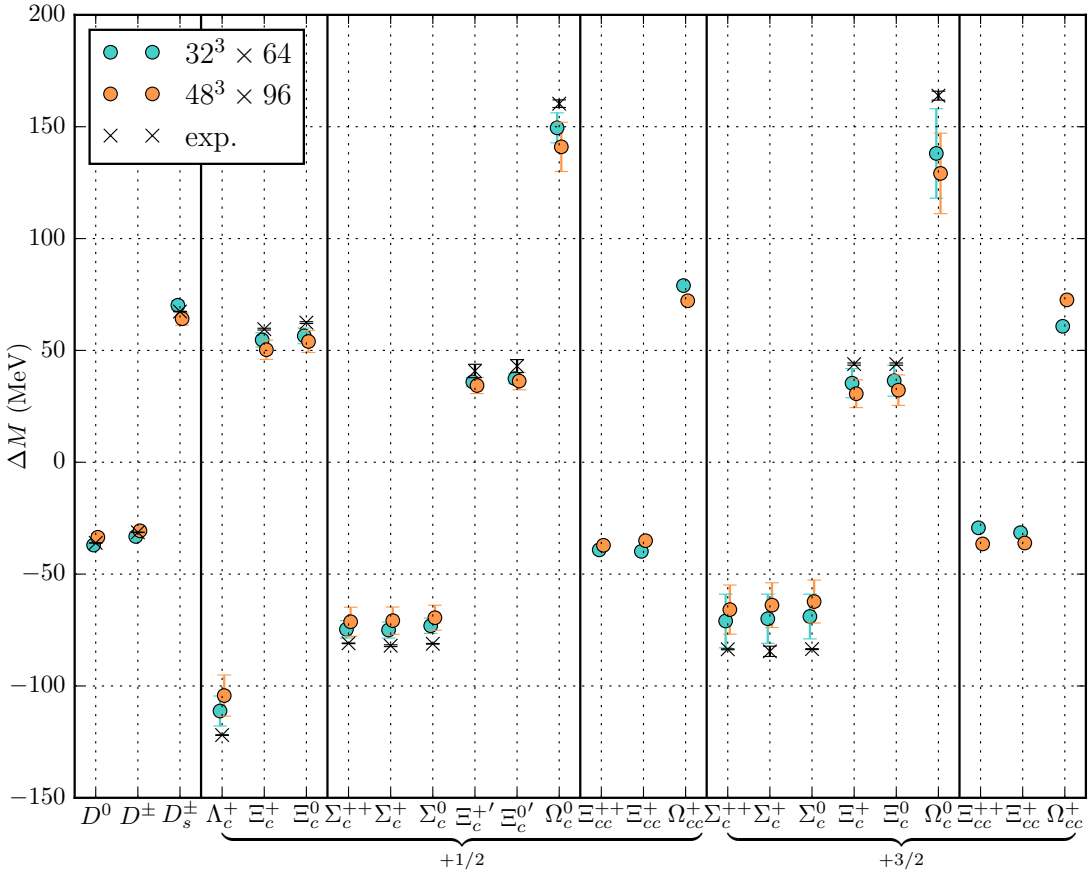


Figure 6.15: Charmed hadron mass splitting with the average mass within each SU(3) multiplet subtracted. The black crosses are experimental data. The coloured points are estimates generated from our lattice analysis.

of the isospin multiplet subtracted. We see agreement between the $32^3 \times 64$ and $48^3 \times 96$ for all but the Ξ_{cc} of both spin types, and only in the spin $\frac{1}{2}$ case is the disagreement statistically significant. Where experimental values are provided, the lattice analysis agrees well. The issues arise when QED effects become significant.

In Fig. 6.17 we show the strong isospin and charge splitting from the average of isospin multiplets. QCD effects typically dominate these splittings, with QED only playing a significant role when +2 charge particles are involved. The statistical error on the QED effects is quite high, and increasing the accuracy of the QED splittings will be the subject of future work. In any case there are no signs of unexpectedly large QCD isospin or charge splittings. We see no evidence of exceptional QCD isospin ($\sim 60\,\text{MeV}$) or charge ($\sim 20\,\text{MeV}$) generated splittings, and so if the other resonances detected by SELEX exist, $\Xi_{cc}^{++}(3460), \Xi_{cc}^{+}(3443), \Xi_{cc}^{++}(3541), \Xi_{cc}^{++}(3780)$,

we do not suspect they are ground-state positive-parity charmed baryons.

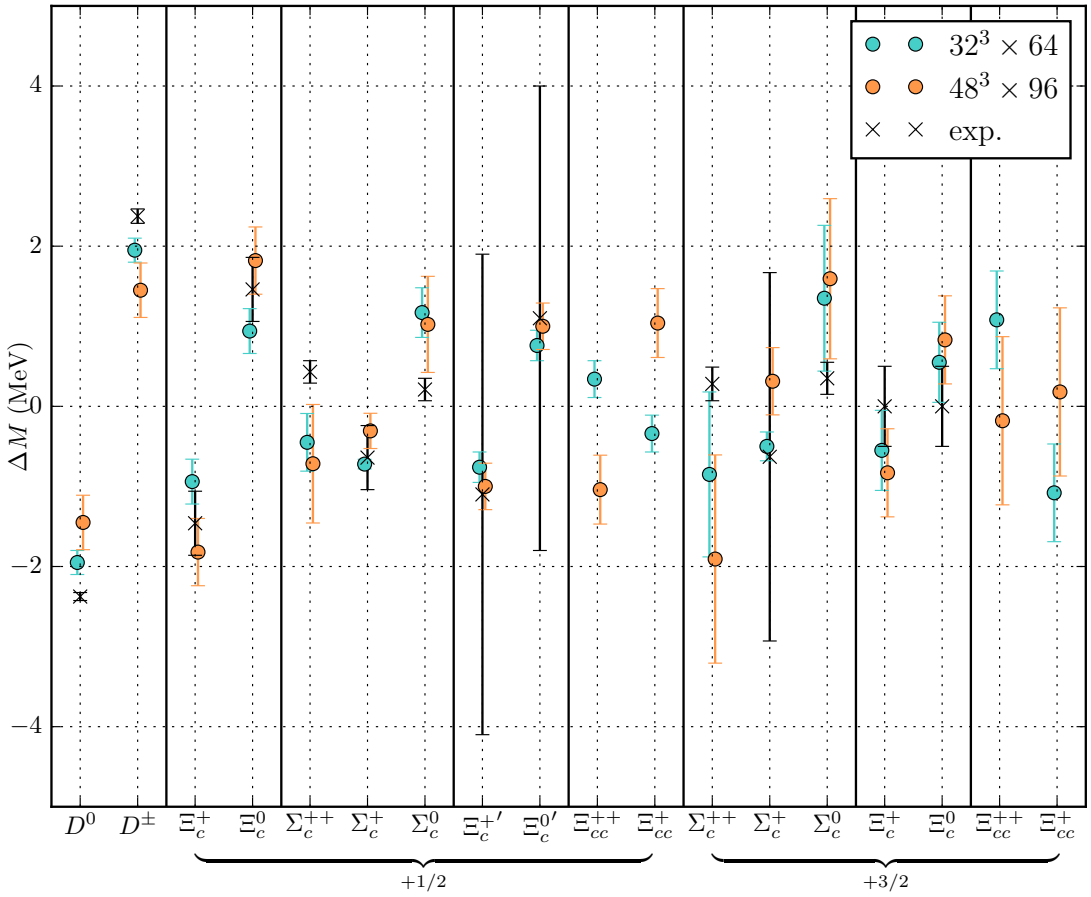


Figure 6.16: Charmed hadron mass splitting with the average mass within each isospin multiplet subtracted. The black crosses are experimental data. The coloured points are estimates generated from our lattice analysis.

Lastly, in Tab. 6.6 we show various hyperfine mass splittings. The values that are displayed are for the $48^3 \times 96$ volume. These values were systematically lower than the $32^3 \times 64$, the second error (which is an estimate of the systematic error) represents this uncertainty. The first value in the table is the mass splitting between the $J^P = \frac{3}{2}^+$ and the $J^P = \frac{1}{2}^+$ Ξ_{cc}^{++} and is consistent with the mass splitting between the measurements given by SELEX and LHCb $\sim 100\,\text{MeV}$. This value and the other values are compared with Brown [62] and experiment. We see good agreement with experiment when both systematic and statistical errors are included, and reasonable agreement with Brown.

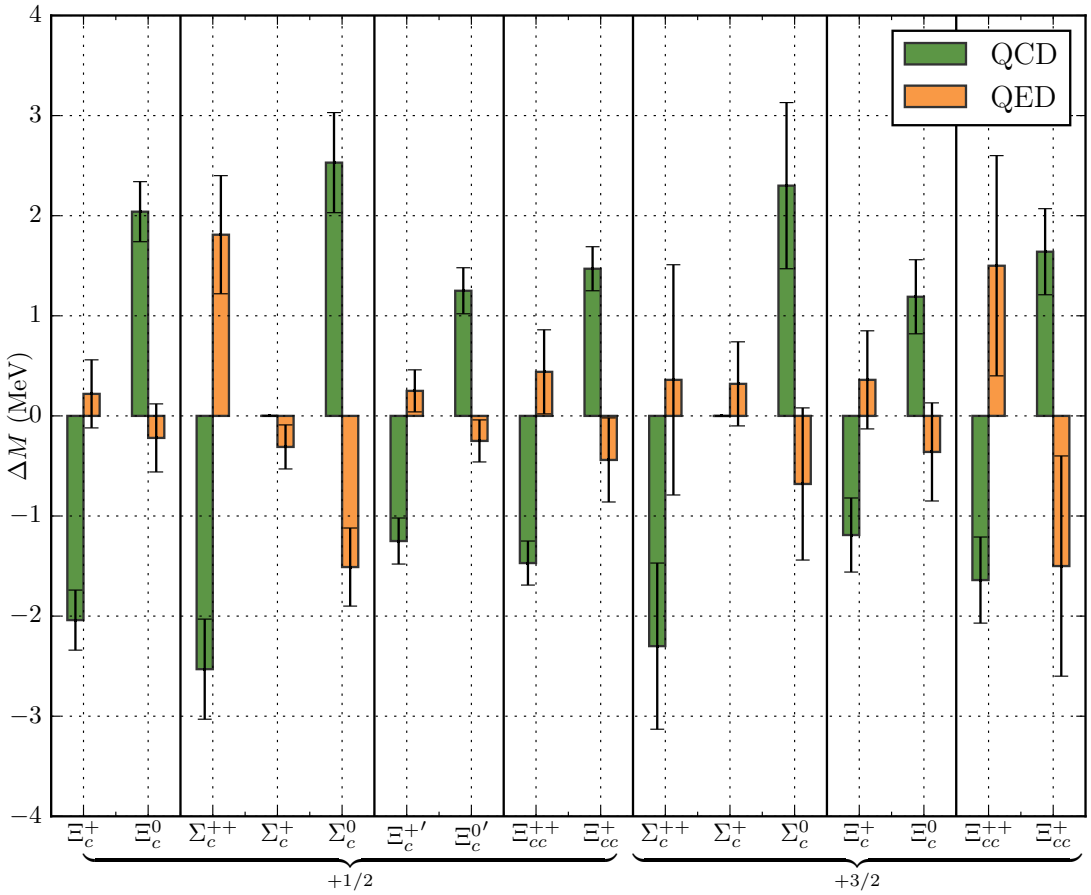


Figure 6.17: QCD and QED components of the charmed baryon mass splittings with the average mass within each isospin multiplet subtracted, for the $48^3 \times 96$ ensemble. The error bars represent symmetric statistical errors on the centre values.

6.6 Summary

We produced estimates for the masses and mass splittings of the lowest lying charmed baryons states using Lattice QCD+QED. This included estimates for the full spectrum's absolute mass, as well as more accurate estimates of the splittings between the SU(3) and isospin multiplets and hyperfine splittings. We also include a breakdown of the isospin splittings in terms of QCD and QED components. These estimates agreed with experiment where experiment was available.

In terms of the mass spectrum we calculated, the Ξ_{cc} found by LHCb [17] is closest to Ξ_{cc} with $J^P = \frac{3}{2}^+$, while the particle confirmed by SELEX [16] was

closest to the Ξ_{cc} with $J^P = \frac{1}{2}^+$. This picture was constructed from considering quantum numbers, approximate mass and approximate hyperfine splitting within our errors. We did not find evidence for the other states found by SELEX. In particular the strong and electromagnetic isospin splittings were of a similar order to lower mass particles $\sim 1 - 6\,\mathrm{MeV}$.

$J^{PC} = 0^{-+}$	D^0	$D^{\pm}$	$D_s^{\pm}$
This work	1819.0(63)(30)	1821.9(61)(30)	1916.8(46)(30)
Experiment	1864.83	1869.58	1968.27
$J^P = \frac{1}{2}^+$	Λ_c	Ξ_c^0	$\Xi_c^{\pm}$
This work	2250(23)(30)	2404(12)(30)	2408(12)(30)
Experiment	2286.46(14)	2467.93(40)	2470.85(4)
$\frac{1}{2}^+$	Σ_c^{++}	Σ_c^{+}	Σ_c^{0}
This work	2405(18)(30)	2405(18)(30)	2407(10)(30)
Experiment	2453.97(14)	2452.9(4)	2453.75(14)
$\frac{1}{2}^+$	$\Xi_c^{+\prime}$	$\Xi_c^{0\prime}$	Ω_c^0
This work	2510(10)(40)	2512(10)(40)	2617(69)(40)
Experiment	2575.7(30)	2577.9(29)	2695.2(17)
$\frac{3}{2}^+$	Σ_c^{++*}	Σ_c^{+*}	Σ_c^{0*}
This work	2511(32)(40)	2513(31)(40)	2515(31)(40)
Experiment	2518.41(21)	2517.5(23)	2518.48(20)
$\frac{3}{2}^+$	Ξ_c^{+*}	Ξ_c^{0*}	Ω_c^{0*}
This work	2607(18)(40)	2609(17)(40)	2706(12)(40)
Experiment	2645.9(5)	2645.9(5)	2765.9(20)
$\frac{1}{2}^+$	Ξ_{cc}^{++}	Ξ_{cc}^{+}	Ω_{cc}^{+}
This work	3495(10)(100)	3497(10)(100)	3604(7)(100)
Alexandrou[61]	3606(11)(8)	3606(11)(8)	3711(5)(30)
$\frac{3}{2}^+$	Ξ_{cc}^{++*}	Ξ_{cc}^{+*}	Ω_{cc}^{+*}
This work	3566(10)(120)	3567(10)(120)	3675(11)(120)
Alexandrou[61]	3682(10)(26)	3682(10)(26)	3770(6)(30)

Table 6.5: Predicted masses for D mesons and singly and doubly charmed baryons in MeV obtained from the $48^3 \times 96$ ensemble. The first error is statistical, while the second error is an estimate of the systematic error, which is obtained by the difference in the $32^3 \times 64$ results to those of the $48^3 \times 96$.

	This work	Brown [62]	Experiment
$\Xi_{cc}^{++*} - \Xi_{cc}^{++}$	71(16)(30)	82.8(72)(58)	$--$
$\Xi_c^{+*} - \Xi_c^{+\prime}$	97(17)(10)	73.7(50)(87)	69.1(22)
$\Xi_c^{+\prime} - \Xi_c^{+}$	106(9)(30)	140(16)(38)	107.9(22)
$\Sigma_c^{++*} - \Sigma_c^{++}$	108(23)(9)	78(7)(11)	64.53(43)
$\Sigma_c^{++} - \Lambda_c^{+}$	155(19)(2)	219(36)(43)	167.33(18)

Table 6.6: The hyperfine mass splittings. Values quoted in MeV. The first error is statistical, while the second error is an estimate of the systematic error, which is obtained by the difference in the $32^3 \times 64$ results to those of the $48^3 \times 96$.

CHAPTER 7

Flavour-neutral pseudoscalar mesons using Lattice QCD+QED

In this work, we explore the spectrum of the flavour-neutral pseudoscalar mesons on dynamical QCD+QED lattices. To reduce the statistical noise associated with annihilation diagrams we utilise exact colour and spin dilution with a spatial interlacing for our $\mathbb{Z}_2$ noise sources. While modern calculations using lattice QCD give accurate estimates of the η–η' splitting, we make first estimates of the contribution of electromagnetic effects in the π^0–η splitting. We will also consider the state composition of the π^0, η and η' around points of exact degeneracy (the SU(3) point) in an attempt to predict state behaviour near the physical point.

Recent work studying the mass and flavour composition of flavour-neutral pseudoscalar mesons on the lattice include Refs. [68, 69, 70, 71, 72, 73, 74, 75, 76], all of which are 2+1 or 2+1+1 flavour analysis in pure QCD (no QED). In this simpler case more recent papers obtain very reliable estimates of the η' mass, state composition of the η, η' system, and hence good estimates of mixing angle between the η and η'.

An essential difficulty in determining the η' mass originates from the axial current anomaly, which I discuss in section 7.1. The anomaly generates mass through topological charge density, and the η' couples to the charge density through disconnected quark loop diagrams. A key feature of all of the above analysis is the need to calculate the contribution of these disconnected quark loop diagrams, and diagonalise the correlation functions from a flavour basis into the energy eigenstate basis of the π^0, η, η' system.

To determine the disconnected quark loop diagrams, one must calculate either the all-to-all propagator, which is very computationally expensive, or (alternatively) obtain an estimate of the self-to-self components of the propagator. The

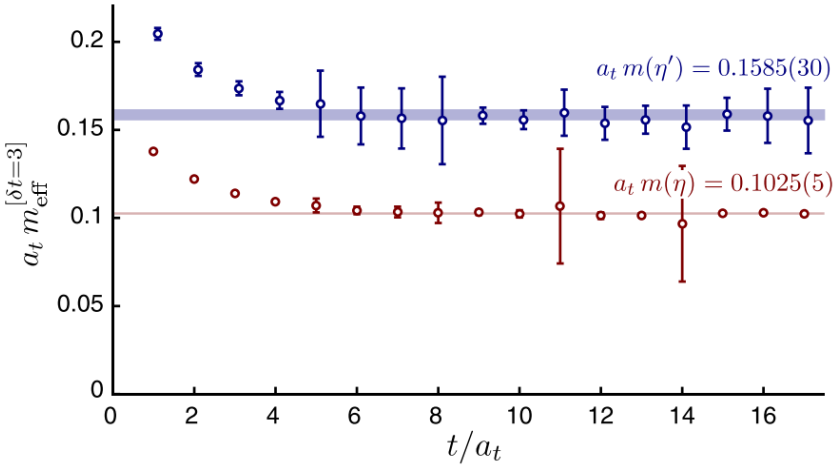


Figure 7.1: Effective mass plot of the η and η' from [70].

latter is the tact used in most of the above analyses. The self-to-self components are calculated using noisy sources.

The results of a very nice analysis by Ref. [70] are show in Fig. 7.1 from which they calculated an effective mass of the η' of approximately 900 MeV and an η mass of approximately 600 MeV at a pion mass of 396 MeV. They gave an estimate of the mixing angle of the η and η' of $42(1)^\circ$[1](light-strange angle). A more recent study by [74] produced similar results, the effective mass plot is show in Fig. 7.2. They give an estimate of the η' mass of 878(35) MeV and the η mass of 504(7) MeV at a pion mass of 282 MeV.

This chapter is organised as follows. In section 7.1 I explain some of the important background regarding flavour-neutral pseudoscalar mesons. I explain why the η' in particular has been of interest to the physics community for so long, and describe some ideas that have been presented as to how and why disconnected quark line contributions generate significant mass in the η'. In section 7.2 I show how disconnected quark line contributions occur in the correlation functions for flavour-neutral pseudoscalar mesons. I then formalise this in a flavour diagonal basis.

In section 7.3 I explain the difficulties in calculating disconnected quark line contributions, and how this relates to self-to-self propagator elements. I then describe how noisy sources can be used to improve calculations of the self-to-self components of the quark propagators. In section 7.4 and section 7.5 I describe the two methods we explored to increase dilution of the noisy sources, interlacing and

[1]The mixing angle is a useful parameter in the two state system. In our analysis, however, the state vectors sit on the surface of a sphere rather than a circle, and hence the interpretation using angles is not as useful.

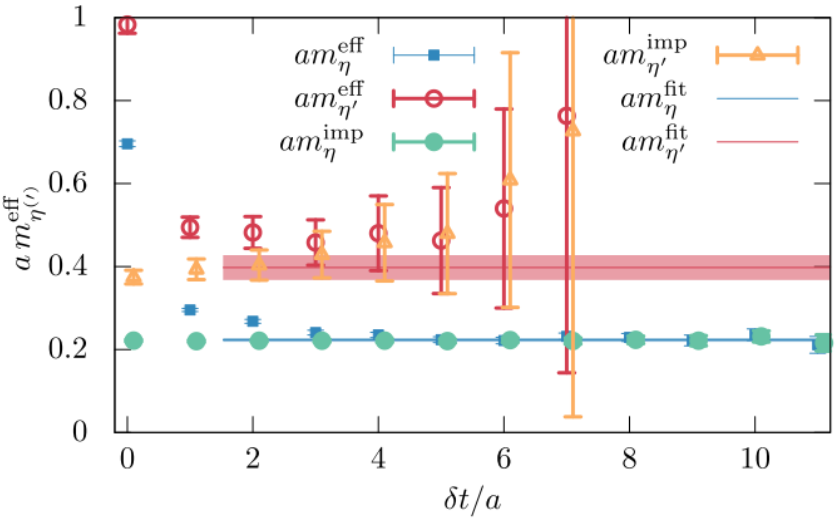


Figure 7.2: Effective mass plot of the η and η' from [74].

hadamard vectors. These both reduce the error associated with calculation of the self-to-self propagator. In section 7.6 I describe the specifics concerning the lattice configurations. I explain why we selected certain quark dynamical mass points; these points were chosen to determine the electromagnetic splitting of the π^0 and η masses. I also explain how the signal of the (difficult to calculate) disconnected quark line components was improved in our analysis.

In section 7.7 I present results from our analysis, including effective mass plots of the π^0–η–η' system, mass splittings of the two lowest states and determination of the QED part of this splitting under the Dashen scheme definition near the SU(3) symmetric point.

We more generally explore the state composition under symmetry breaking. We show that the quark flavour diagonal basis, $M_{q\bar{q}}$ (connected quark line components only) is a useful way to measure symmetry breaking in hadronic QCD+QED systems that have symmetries broken in both charge and quark mass. Lastly, we use the analysis on these ensembles to predict the state changing behaviour of the π^0–η system near the isospin point, when the strange quark mass is held fixed $\delta m_s = 0$, but the u, d quark masses vary and satisfy $\delta m_u = -\delta m_d$. Finally, I summarise our results in section 7.8.

7.1 The U(1) Problem

The pseudoscalar mesons are the lightest observed hadrons within the QCD mass spectrum. These mesons are significantly lighter than the spin one vector mesons;

for example the π^+ is approximately 600 MeV lighter than the ρ^+, but both have the same quark content. The mass splitting is not typically this large for a change in spin. The lightness of these mesons is believed to occur because the pseudoscalar mesons are approximate Goldstone bosons. I will explain what this means.

In the zero mass limit of the three lightest quark flavours, the QCD Lagrangian is invariant under global SU(3) flavour symmetries of quark fermion fields. The QCD Lagrangian is invariant under the transformation $\psi \to e^{iT^a\epsilon^a}\psi$, for Gell-Mann matrices T^a and index $a \in (1, 2...8)$. This symmetry remains when the quark masses are equal but non-zero. The corresponding conserved current is called the vector current,

$$V^a_\mu = \bar{\psi}\gamma_\mu T^a \psi. \tag{7.1}$$

The massless QCD Lagrangian is also invariant under the transformation $\psi \to e^{i\gamma_5 T^a\epsilon^a}\psi$. This symmetry only exists in the massless limit $m_q \to 0$, as the term $m\bar{\psi}\psi \to m\bar{\psi}\psi - 2im\epsilon^a(\bar{\psi}\gamma_5 T^a\psi)$ is not invariant. The conserved current associated with this symmetry is called the axial-vector current,

$$A^a_\mu = \bar{\psi}\gamma_\mu\gamma_5 T^a \psi. \tag{7.2}$$

If these axial-vector symmetry is spontaneously broken at the quantum level in this zero quark mass limit, it implies the existence of eight massless bosons, called Goldstone bosons. As we move away from the zero quark mass limit these Goldstone bosons obtain mass proportional to the quark mass ($M^2_{q\bar{q}} \propto m_q$). The approximate conservation of these currents is observed in their divergences,

$$\partial_\mu V^\mu_a = i\bar{\psi}\left[M, T_a\right]\psi \tag{7.3}$$

$$\partial_\mu A^\mu_a = i\bar{\psi}\gamma_5\left\{T_a, M\right\}\psi, \tag{7.4}$$

for diagonal quark mass matrix $M = \text{diag}\begin{pmatrix} m_u & m_d & m_s \end{pmatrix}$. The pseudoscalar mesons are believed to be these approximate Goldstone bosons.

There are two additional symmetries of the massless QCD Lagrangian. The first is the $U(1)_V$ symmetry corresponding to the transformation $\psi \to e^{i\epsilon}\psi$, which remains even if quark masses are non-zero and not equal. The conserved current is,

$$V^0_\mu = \bar{\psi}\gamma_\mu\psi. \tag{7.5}$$

This current is associated with baryon number conservation, and the symmetry exists irrespective of the quark mass values. The divergence of this current is, $\partial_\mu V^\mu_0 = 0$. The second is the $U(1)_A$ symmetry corresponding to the transformation $\psi \to e^{i\gamma_5\epsilon}\psi$ with conserved current,

$$A^0_\mu = \bar{\psi}\gamma_5\gamma_\mu\psi. \tag{7.6}$$

which is only conserved in the $m \to 0$ limit. A^0_μ current should be associated with either a conserved quantum number or an extra Goldstone boson. In nature, however, we do not observe either $U(1)_A$ quantum number or an extra Goldstone boson, and this is called the "U(1) problem". The lightest SU(3) singlet pseudoscalar is the η' which has a mass of almost 1 GeV; too heavy to be a Goldstone boson. Adler, Bell and Jackiw [77] showed that there is an anomaly associated with this current. A modern derivation was given by Fujikawa [78, 79]. The divergence of the $U(1)_A$ current is [80],

$$\partial^\mu A^0_\mu(x) = 2i\bar{\psi}(x)\gamma_5 M\psi(x) + 2LQ \tag{7.7}$$

where the first term on the right hand side is associated with the breaking with respect to quark mass, M is the quark mass matrix, and the second term is the anomalous term which is non-zero even at quark mass $m = 0$. Within the anomalous term, L is the number of light quark flavours while Q is called the topological charge density,

$$Q = \frac{g^2}{32\pi} F^{\alpha\beta} F^{\gamma\delta} \epsilon_{\alpha\beta\gamma\delta}. \tag{7.8}$$

g is the gauge field coupling coefficient, $F^{\mu\nu}$ is the gauge field strength tensor and $\tilde{F}_{\alpha\beta} = F^{\gamma\delta}\epsilon_{\alpha\beta\gamma\delta}$ is the dual field strength tensor.

The topological charge density represents vacuum charge in QCD, which is not unlike a background electric field in QED. Q is an allowed term in the QCD lagrangian which is CP violating. This CP violating term can be added to the QCD Lagrangian using the term $2\theta Q$, for constant θ called the vacuum angle. However a non-zero vacuum angle has not been detected in QCD, with an experimental upper bound $|\theta| < 1.61(51)\times 10^{-10}$ [81] which is determined by the neutron electric dipole moment [82]. Due to the anomaly in the A^μ_0 current the η' mass does not have to vanish in the limit quark mass $m \to 0$.

The anomaly by itself is not enough to explain the large mass difference between the η' and the other octet baryons. We would also expect a non-zero matrix element of the η' with the vacuum,

$$\langle 0 \,|Q|\, \eta' \rangle \neq 0 \tag{7.9}$$

however the topological charge density Eq. (7.8) is a total divergence ∂, hence to finite orders in perturbation theory at zero momentum, the matrix element vanishes. This indicates it is a non-perturbative effect, but the matrix element may still be non-zero.

In pure gauge QCD, physical θ dependence is allowed, however it is not allowed when there is a $U(1)_A$ anomaly and the quark masses are zero [83]. Witten argued

[1] Using neutron electron dipole moment.

[84] that the η' meson mass cancels the θ dependence in the Lagrangian in the zero quark mass limit causing the η' to gain mass. This leads to the conclusion that the mass of the η' is of order $\mathcal{O}(1/N_c)$ (N_c being the number of quark colours), which I will explain following [84, 80].

The pure gauge partition function with θ dependence is,

$$Z = \int DA \exp\left(i \int d^4x \ \mathrm{Tr}\left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + 2\theta Q\right]\right). \tag{7.10}$$

If we take the derivative with respect to θ twice of Z, we produce a two point function of the topological charge density,

$$\left(\frac{\partial^2 Z}{\partial\theta^2}\right)_{\theta=0} = \left(\frac{g^2}{16\pi^2}\right)^2 \int d^4y \int d^4x \ \left\langle T(F\tilde{F}(y)F\tilde{F}(x))\right\rangle. \tag{7.11}$$

Consider the function,

$$U(k) = \int d^4x \ e^{ik\cdot x} \left\langle T(F\tilde{F}(x)F\tilde{F}(0))\right\rangle. \tag{7.12}$$

In the $1/N_c$ expansion formalism g is held constant while the number of colours is increased, $g = g'\sqrt{N_c}$. The ratio N_f/N_c is kept fixed, where N_f is the number of quark flavours. In the limit $N_c \to \infty$, then the ratio $L/N_c \to 0$, where L is the number of light flavours. Each successive quark loop is suppressed by an extra factor of $1/N_c$ because each loop of quarks is of order $\mathcal{O}(N_c)$ while gluon interactions are of order $\mathcal{O}(N_c^2)$. If a process occurs as a result of a quark loop, it could have also occurred as a gluon, hence is reduced in magnitude compared to the pure gauge theory.

Break $U(k)$ down into an expansion in $1/N_c$ so that $U(k) = U_0(k) + U_1(k) + ...$ where $U_0(k)$ is the sum of all diagrams without quark loops, $U_1(k)$ is the sum of all diagrams with one quark loop, etc. Suppose θ dependence occurs at leading order. There are no quark loop effects at this order, but as quarks become massless the θ dependence needs to disappear at leading order. If one of the mesons is massive when $m_q = 0$ and the meson has mass squared $m_{\eta'}^2 \propto 1/N_c$ then the meson propagators occur at leading order in the $1/N_c$ expansion (they occur in $U_0(k)$) which removes any contradiction. With these two assumptions we arrive at,

$$\begin{aligned} U_0(k) &= U_{YM}(k)\left(1 + \frac{L}{N_c}\frac{\lambda_{\eta'}^2}{k^2 - m_{\eta'}^2}\right) \\ &= U_{YM}(k)\left(\frac{k^2 - m_{\eta'}^2 + \frac{L}{N_c}\lambda_{\eta'}^2}{k^2 - m_{\eta'}^2}\right) \end{aligned} \tag{7.13}$$

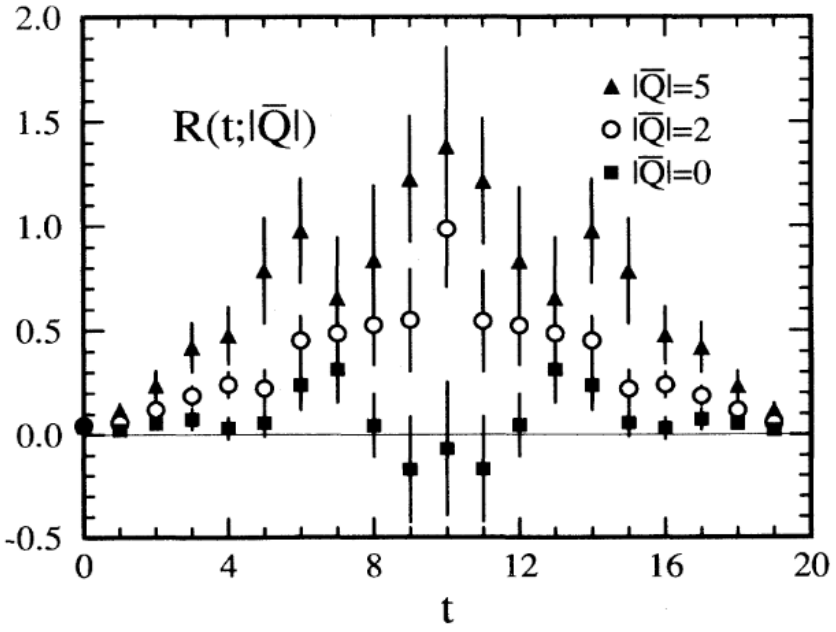


Figure 7.3: Representative values of the ratio of the disconnected amplitude to the connected amplitude $R(t;|\bar{Q}|) = \frac{\langle\eta'(t)\eta'(0)\rangle_{dis}}{\langle\eta'(t)\eta'(0)\rangle_{con}}$ of the η' propagator for several values of $|\bar{Q}|$ at $a = 0.14\,\text{fm}$ and $m_\pi \approx 458\,\text{MeV}$ on a $12^3 \times 20$ lattice. Taken from [85].

where the coefficient $\frac{L}{N_c}\lambda^2_{\eta'} = \frac{\left(\langle 0|F\tilde{F}|\eta'\rangle\right)^2}{U_{YM}(0)}$. When $m_q = 0$ and $k = 0$ in Eq. (7.13), $U_0(k) = 0$, hence $m^2_{\eta'} = \lambda^2_{\eta'}\frac{L}{N_c}$. Veneziano used these ideas to produce an estimate for the η' mass away from this ideal limit [80]. When $m_q > 0$ we can use the identity Eq. (7.7) to suggest the η' gains mass in the same way as the other pseudoscalars, $m^2_{\eta'} = m^2_{NS} + \lambda^2_{\eta'}\frac{L}{N_c}$, where m_{NS} is the mass of non-singlet pseudocalars and increases proportional the quark mass. For the predicted SU(3) quark composition of the η', $m^2_{NS} \sim \frac{2}{3}m^2_{K^0} + \frac{1}{3}m^2_{\pi^0}$ at the physical quark masses, where this quark composition assumes isospin breaking effects to be insignificant and either η–η' mixing is small or the light quark masses are the same.

Without either experimental or theoretical estimates of these quantities there is little more we can do to advance this argument, which is why lattice QCD is an essential tool for studying the η'. Lattice studies of the η' have produced estimates for the η' meson mass and mixing angles with the η [68, 69, 70, 71, 72, 73, 74, 75, 76]. The masses are very close to the physical value of η' ($m_{\eta'} \sim 958\,\text{MeV}$) in more recent studies. There is also evidence for the η' mass dependence on the topological charge Q [85, 86], a plot of this dependence in quenched QCD is shown in Fig. 7.3. This shows, based on the previous work by Smit [87], that

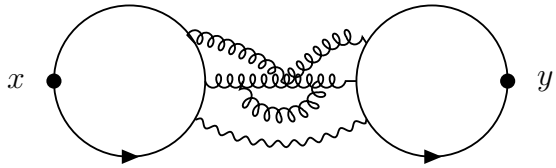


Figure 7.4: Disconnected quark line Feynman diagram.

topological charge and disconnected quark line contributions are directly related $\mathrm{Tr}(\gamma_5 K_q^{-1}(x,x)) = \frac{Q}{m_q}$ for quark flavour q and quark mass m_q. This topological charge dependence occurs when performing Wick contractions of flavour-neutral mesons. Normally these components are not important because they are OZI suppressed [88, 89, 90]. OZI suppressed means that the QCD gauge coupling is small when the gluons have high energy, where high energy corresponds to the mass of mesons. The importance of these disconnected quark line contributions within the flavour-neutral pseudoscalar mesons, when such contributions should be OZI suppressed, is explained through Witten's and Veneziano's arguments. In this way the anomaly generates the η' mass difference through quark—anti-quark annihilation diagrams, which is in agreement with quark model arguments.

7.2 Quark Annihilation diagrams

Most pseudoscalar mesons can be classified uniquely in terms of their quark content. However this is more complicated for the flavour diagonal pseudoscalar mesons π^0, η and η' because they have quark annihilation diagrams, also called disconnected quark line diagrams (or disconnected contributions), contributing to their mass. These disconnected quark line diagrams are shown in Fig. 7.4. A disconnected contribution is a process by which the flavour diagonal meson propagates only via gluon and (or) photon lines; the quark inside the meson propagates back to its starting point and annihilates. This allows mesons of different flavours to interact. The disconnected contributions break the S_3 symmetry of the meson states and lead to mixing of states. The composition of the states vary with flavour symmetry breaking.

Flavour neutral mesons in flavour basis for example $\pi_{\bar{u}u} = \bar{u}\gamma_5 u$ propagate via quark lines but also via gluons. The correlation function in flavour basis for flavour-neutral pseudoscalar meson two point functions is,

$$C_{q';q}(\vec{p},t) = \sum_{\vec{x},\vec{y}} e^{-i\vec{p}(\vec{y}-\vec{x})} \left\langle \Omega \left| \mathrm{T}\{O_{\pi_{\bar{q}q}}(y) O^{\dagger}_{\pi_{\bar{q}'q'}}(x)\} \right| \Omega \right\rangle \tag{7.14}$$

$$= \sum_{\vec{x},\vec{y}} e^{-i\vec{p}(\vec{y}-\vec{x})} \left\langle \Omega \left| \mathrm{T}\{\bar{q}\gamma_5 q(y)[-\bar{q}'\gamma_5 q'(x)]\} \right| \Omega \right\rangle . \tag{7.15}$$

By performing Wick contractions we produce propagators. The contractions only occur between quarks of the same flavour but in this case, as the meson is flavour-neutral, the quark can propagate back to its source,

$$C_{q';q}(\vec{p},t) \approx \frac{1}{N}\sum_{n=1}^{N}\sum_{\vec{x},\vec{y}} e^{-i\vec{p}(\vec{y}-\vec{x})}\left(\delta_{qq'}\mathrm{tr}[\gamma_5 K_q^{-1}(x-y)\gamma_5 K_q^{-1}(y-x)][U_n]\right.$$
$$\left.-\mathrm{tr}[\gamma_5 K_q^{-1}(y-y)][U_n]\mathrm{tr}[\gamma_5 K_{q'}^{-1}(x-x)][U_n]\right). \quad (7.16)$$

The components of $C_{q';q}(0,t)$ form a correlation matrix $C(t)$ in the flavour basis. In Eq. (7.16) the first term is the connected part and the second term in the disconnected part.

With exact SU(3) symmetry, the decomposition of the eigenstates of π^0, η and η' is known in flavour operator basis $\mathcal{O}_q = \bar{q}\gamma_5 q$. We can determine these eigenstates that appear in the QCD vacuum at the SU(3) point using the fact that the correlation matrix is proportional to the Hamiltonian of the system $C(t) \propto e^{-Ht}$. In flavour basis (considering only light flavours $\{u,d,s\}$), the connected and disconnected parts of the correlation matrix are the same for each flavour. The correlation matrix in this case takes the form, $C(t) = cI + d[1]$ where I is the 3×3 identity matrix, [1] is a 3×3 matrix of ones '1', c is the connected and d is the disconnected part of the correlation matrix. If we determine eigenvectors of this matrix at large times[2]we produce the eigenstates for the π^0, η and η',

$$\pi^0 = \frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})$$
$$\eta = \frac{1}{\sqrt{6}}(u\bar{u} + d\bar{d} - 2s\bar{s})$$
$$\eta' = \frac{1}{\sqrt{3}}(u\bar{u} + d\bar{d} + s\bar{s}). \quad (7.17)$$

However these states are not necessarily energy eigenstates when SU(3) flavour symmetry is broken (the quark masses and charges are not degenerate).

If there were no disconnected contributions the eigenstates would simply be of quark composition $\pi^0 = u\bar{u}$, $\eta = d\bar{d}$ and $\eta' = s\bar{s}$, labelled in mass order assuming $m_u < m_d < m_s$. Instead each particle is a mixture of these three flavour diagonal states, as well as existing as propagating photons and gluons. The state composition at the SU(3) point is considered to be fully mixed in flavour space. At this SU(3) symmetry point only the η' has a non-zero disconnected contribution to its mass and its mass is different from the lowest two states. These lowest states are energy degenerate, and hence the eigenstates listed above are not unique, rather any orthogonal linear combination of these two states are possible.

[2]Note in this special case the states are the same irrespective of the time of diagonalisation.

We do not know the state composition of the energy eigenstates π^0, η, and η' in the flavour basis at arbitrary quark mass points, however they can be determined in a similar way as above by finding the eigenvectors of the flavour basis correlation matrix at large times. The flavour basis is the practical basis for calculation, because it is these correlation matrix components we can calculate on the lattice.

We can formalise correlation functions for the energy eigenstates π^0, η and η' in terms of the flavour basis correlation functions. In the limit of large time on the lattice, only the lowest energy neutral flavour state remains, for instance for η we find $\mathcal{O}_\eta^\dagger |\Omega\rangle \to |\mathcal{O}_\eta\rangle$[3]. If we let $\pi, \pi' \in \{\pi^0, \eta, \eta'\}$, then the elements of the correlation matrix for flavour-neutral mesons π^0, η, η' (which are eigenstates of the Hamiltonian) are given by,

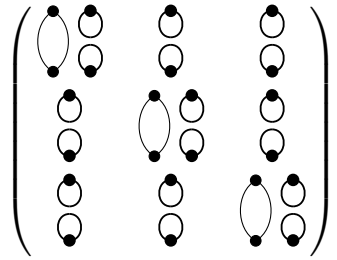
$$C_{\pi';\pi}(t'-t_0;\vec{p}) = \sum_{\vec{x},\vec{y}} e^{-i\vec{p}(\vec{y}-\vec{x})} \left\langle \mathcal{O}_{\pi'}(y) \left| e^{-H(t'-t_0)} \right| \mathcal{O}_\pi(x) \right\rangle \tag{7.18}$$

where $y = (t', \vec{y})$ and $x = (t_0, \vec{x})$. These elements form a matrix which we call $C(t'-t_0;\vec{p})$ and this matrix is diagonal at large times $t = t'-t_0$. The states $|\mathcal{O}_\pi\rangle$ are those states that diagonalise the Hamiltonian at large lattice times, however we do not explicitly know them yet in any basis. Let's consider only $\vec{p}=0$ and implicitly include the sums over the $\vec{x}$, $\vec{y}$ components into the brackets $|\mathcal{O}_\pi(t_0)\rangle = \left|\sum_{\vec{x}} e^{i\vec{p}\cdot\vec{x}} \mathcal{O}_\pi(t_0,\vec{x})\right\rangle$. We want to change this expression into something that includes the flavour basis correlation matrix, which is a basis we can calculate correlation functions in. The expression becomes,

$$\begin{aligned} C_{\pi';\pi}(t) &= \left\langle \mathcal{O}_\pi(t') \left| e^{-Ht} \right| \mathcal{O}_{\pi'}(t_0) \right\rangle \\ &= \sum_{q,q'} \langle \mathcal{O}_\pi(t') \, |\mathcal{O}_q(t')\rangle \underbrace{\sum_k \langle \mathcal{O}_q(t') \, |k\rangle \, \langle k \, |\mathcal{O}_{q'}(t_0)\rangle \, e^{-E_k t}}_{C_{q';q}(t)} \langle \mathcal{O}_{q'}(t_0) \, |\mathcal{O}_{\pi'}(t_0)\rangle \end{aligned} \tag{7.19}$$

where $q, q' \in \{u, d, s\}$ and the states k are energy eigenstates. In matrix notation this becomes,

$$C(t) = P^\dagger \underbrace{\left(\quad \right)}_{C_O(t)} P. \tag{7.20}$$



[3] $\mathcal{O}_\eta^\dagger$ is defined as the interpolator which produces η as its lowest state.

The matrix $C_O(t)$ is the correlation matrix in the flavour operator basis, where the elements are given by $C_{q';q}(t)$ in Eq. (7.16). The matrices P and $P^\dagger$, diagonalise $C_O(t)$ at large $t = t' - t_0$ to produce $C(t)$. P is composed of the eigenvectors of the matrix $C_O(t)$. Note that each new t_0 will give an estimate of $C(t)$ for $t = t' - t_0$. The eigenvectors only represent π^0, η and η' at large times and using these states we can determine the energy of the particles using $E = -\log\left(C^{-1}(t)C(t+\delta t)\right)/\delta t$. By applying this form we assume that the vectors in flavour basis that diagonalise $C(t+\delta t)$ are the same as those that diagonalise $C(t)$. The assumption is valid at large lattice times, where the state composition of the correlation matrix is static.

7.3 Stochastic Noise sources

Disconnected contributions appear in many quantities of interest within QCD, including baryon form factors, axial operators, hadronic coupling constants and of course masses and states of flavour-neutral mesons. In section 7.2 we showed that the disconnected quark line components appear from Wick contractions of elements of the correlation matrix of flavour-neutral mesons. After Wick contracting, the correlation matrix Eq. (7.16) has terms that look like $\mathrm{tr}(\Gamma G(y;x)\Gamma G(x;y))$ which correspond to quark line connected diagrams, while $\mathrm{tr}\left(\Gamma G(y;y)\right)\mathrm{tr}\left(\Gamma G(x;x)\right)$ correspond to quark line disconnected diagrams, where $G(x;y)$ is the quark propagator. The propagator $G(x;y)$ is called the all-to-all propagator and contains the amplitude of the propagation of quarks from any one point to any other point. The calculation of this object is very computationally expensive if each inversion of the propagator were to be computed exactly. Additionally, there may be no advantage in such a calculation, as the estimate of observables is statistical. If the error in the estimate of $G(x;y)$ is small with respect to statistical errors associated with the gauge configurations, the estimate is good enough.

A good estimate of observables relating to the connected part of the all-to-all propagator $G(x;y)$ can be determined from the point to all propagator $G(0;y)$, as translational invariance allows us to consider only a single source point as a representation of the full calculation. Note in this calculation there is a lattice volume worth of information.

A similar tactic can not be employed for the disconnected parts of the propagator, as the result is only $G(0;0)$, a single point. The statistics on this point are a volume smaller than that of point-to-all propagator. To get a good estimate on the disconnected parts we need a great number of statistics to obtain a correlation amplitude above gauge noise. Hence, we need all, or most of the information that we can obtain from a gauge configuration. One method that has been used in the literature to great success is the distillation method [70, 73], which produces an estimate of the all-to-all propagator in a systematic way. We use a slightly simpler

method that instead determines all self-to-self propagator components $G(y;y)$.

A different estimation method than the one used for the connected part, is used to determine $G(y;y)$ components of the propagator. Sources are placed at all, or many of the lattice sites, these are called wall sources. This estimation method introduces noise into the calculation of $G(y;y)$. Originally wall source techniques relied on differences in the gauge configurations to cancel wall source noise in the propagator components [87, 85]. More recently these errors in the propagator estimate have been reduced by producing random wall sources (or stochastic noise sources) and by dilution [91, 92, 93, 94], which are explained in more detail below.

Dilution of the noise source is an important consideration when applying wall source techniques. Dilution means making some parts of the noise vector zero and calculating each of the parts separately to reconstruct the full propagator. A complete dilution results in an exact all-to-all propagator. Noise in the estimate of the propagator is proportional to the proximity of non-diluted elements. Dilution in spin, colour and time significantly reduces error, because each of these elements is related by their spatial location, hence is a nearest neighbour. Typically the higher the level of dilution the more accurate the estimate of the propagator is, however more inversions are required. Finding the right balance is dependent on the problem that one is trying to solve. Using the insights of previous studies of the pseudoscalar mesons in the literature [92, 93], we will confine our discussion to Z_2 noise only on the spatial components of the lattice $\xi(\vec{x})$, the noise is diluted in time, spin and colour. Further dilution is applied using interlacing, which will be explained in more detail in section 7.5.

For a given spatial coordinate $\vec{x}_i$ and time slice t, the Z_2 noise source vector is given by $\xi(\vec{x}_i) \in \{-1, +1\}$, with equal probability and zero on all other time slices. The spin and colour components are diluted, hence calculated separately. The scalar product of an element of a noisy source vector $\xi(\vec{x})$ with another element $\xi(\vec{y})$ satisfies the following conditions,

$$\xi(\vec{x})\xi^*(\vec{y}) = \begin{cases} 1 & \text{if } \vec{x} = \vec{y} \\ \xi'(\vec{x}) \in \{1, -1\} & \text{if } \vec{x} \neq \vec{y} \end{cases} \tag{7.21}$$

where the $*$ denotes the complex conjugate (it does nothing for the real case we are looking at) and ξ' is some new noisy vector.

Consider two ordered sets of lattice coordinates X and Y, where in each set every coordinate appears only once $X = (\vec{x}_1, \vec{x}_2, ..., \vec{x}_{N_v})$, $Y = (\vec{y}_1, \vec{y}_2, ..., \vec{y}_{N_v})$. Using these ordered sets of coordinates, we can write the product of a noisy source vector ξ with its conjugate as,

$$\sum_{i=1}^{N_v} \xi(\vec{x}_i)\xi^*(\vec{y}_i) = \begin{cases} N_v & \text{if } \vec{y}_i = \vec{x}_i \forall i \\ \sum_{i=1}^{N_v} \xi'(\vec{x}_i) & \text{if } \vec{y}_i \neq \vec{x}_i \forall i \end{cases} \tag{7.22}$$

for N_v spatial volume and where ξ' is a new noisy source vector (with the same properties of the original one) generated from the misaligned product of ξ with itself.

The value associated with $\langle \xi' \rangle = \frac{1}{N_v} \sum_i^{N_v} \xi'(\vec{x}_i)$ is binomially distributed about zero with probability function

$$P(\sum_i^{N_v} \xi'(\vec{x}_i) = \pm 2k) = \frac{N_v!}{(N_v/2+k)!(N_v/2-k)!}(0.5)^{N_v}, \tag{7.23}$$

for $k \in \mathbb{N}$. The factorials make this probability difficult to calculate, even on a computer, however if one considers a small volume of $N_v = 100$, the probability of k having the lowest 10% of values, meaning $k \in [-10, 10]$ is $P \approx 96.5\%$, if we increase the volume to $N_v = 150$ the probability that it is in the lower 10% of values, $k \in [-15, 15]$, is $P \approx 98.9\%$. As we increase N_v the probability for ξ' to be less than a certain percentage of N_v increases. Hence $\langle \xi' \rangle \to 0$ as the volume increases. Large factorials can be accurately estimated using Stirling's formula. For the full case, $N_v = 24^3$ and $k \in [-150, 150]$, where here we satisfy $2k/N_v \leq 2.18\%$, the probability is $P \approx 99\%$. Put into words, the noise is reduced by a factor of fifty or more, when compared to the signal, in 99% of cases.

Unless a noisy vector is multiplied in the right way, the value of the product will be binomially distributed about 0. For large volumes, this makes the scalar product of a noisy vector with its conjugate act (approximately) like a Kronecker-Delta function.

We can use these sources to obtain the disconnected quark line component. Noting that

$$\sum_{\vec{y}} \mathrm{tr}\left[\Gamma G(t, \vec{y}; t, \vec{y})\right] = \mathrm{tr}\left[\Gamma \sum_{\vec{y}} G(t, \vec{y}; t, \vec{y})\right], \tag{7.24}$$

then it is sufficient to determine only an estimate for the sum of components of the propagator, $\sum_{\vec{y}} G(t, \vec{y}; t, \vec{y})$. This sum can be determined approximately using the below noisy source technique,

$$\sum_{\vec{y}} G(t, \vec{y}; t, \vec{y}) \approx \sum_{\vec{y}} \frac{1}{N_r} \sum_{r=1}^{N_r} \left[\sum_{\vec{x}} G(t, \vec{y}; t, \vec{x}) \xi_r(\vec{x})\right] \xi_r^*(\vec{y}) \tag{7.25}$$

$$= \left[\sum_{\vec{y}} G(t, \vec{y}; t, \vec{y})\right] + \frac{1}{N_r} \sum_{r=1}^{N_r} \sum_{\vec{y}} \left[\sum_{\vec{x}|\vec{x}\neq\vec{y}} G(t, \vec{y}; t, \vec{x}) \xi_r(\vec{x}) \xi_r^*(\vec{y})\right] \tag{7.26}$$

$$\approx \left[\sum_{\vec{y}} G(t, \vec{y}; t, \vec{y})\right] + N_v^2 \underbrace{\overline{|G(t, \vec{y}; t, \vec{x})|}}_{\vec{x}\neq\vec{y}} \frac{1}{N_r} \sum_r \langle \xi' \rangle_r \tag{7.27}$$

the bar on the second term of the last line denotes the average of the propagator elements at time t under the condition $\vec{x} \neq \vec{y}$. The second term on the last line gives an estimate of the error in the propagator $G(t,\vec{y};t,\vec{y})$. As the value of $G(t,\vec{x};t,\vec{y})|_{\vec{x}\neq\vec{y}}$ is largest for nearest neighbours[95], nearest neighbours contribute most to the error in our prediction. The multiplication can be represented using matrices provided there is a prescription for converting the hypercubic lattice to a one dimensional vector, $\sum_{\vec{y}} G(t,\vec{y};t,\vec{y}) \approx \frac{1}{N_r}\sum_{r=1}^{N_r} \xi_r^\dagger(\vec{y}) \times G(t,\vec{x};t,\vec{y}) \times \xi_r(\vec{x})$.

Noisy sources can also be used to determine the quark line connected part. This technique is called the 'one end trick'. For Dirac matrix K the noisy-source solution vector ϕ is defined as $K\phi = \xi$. Hence the solution vector can be written as $\phi(t,\vec{x};t') = \sum_{\vec{y}} G(t,\vec{x};t',\vec{y})\xi(\vec{y})$ which we use to produce an estimate for the correlation function,

$$\frac{1}{N_r}\sum_{\vec{x},r}\phi(t,\vec{x};t')_r\Gamma\phi^\dagger(t';t,\vec{x})_r\bar{\Gamma} = \frac{1}{N_r}\sum_{\vec{x},r,\vec{y},\vec{z}} G(x;t',\vec{y})\xi(\vec{y})_r\Gamma\xi^\dagger(\vec{z})_r G^\dagger(x;t',\vec{z})\bar{\Gamma} \tag{7.28}$$

$$= \left[\sum_{\vec{x},\vec{y}} G(x;y)\Gamma G(y;x)\bar{\Gamma}\right] + \frac{1}{N_r}\sum_{\vec{x},r,\vec{y},\vec{z}} \underbrace{G(x;t',\vec{y})\xi(\vec{y})_r\Gamma\xi^\dagger(\vec{z})_r G^\dagger(x;t',\vec{z})\bar{\Gamma}}_{\vec{z}\neq\vec{y}} \tag{7.29}$$

$$\approx \left[\sum_{\vec{x},\vec{y}} G(x;y)\Gamma G(y;x)\bar{\Gamma}\right] + N_v^3 \underbrace{\overline{|G(x;t',\vec{y})||G(x;t',\vec{z})|}}_{\vec{z}\neq\vec{y}} \frac{1}{N_r}\sum_r \langle\xi'\rangle_r \tag{7.30}$$

the bar on the second term of the last line denotes the average of the product of propagator elements under the condition $\vec{z} \neq \vec{y}$. The second term on the last line gives an estimate of the error in the quantity $G(x;y)\Gamma G(y;x)\bar{\Gamma}$.

7.4 Improvements using Hadamard vectors

Hadamard vectors allow for incremental improvement of the quark line connected and disconnected contributions to the correlation function. Though this was not the method that was used to produce our final results, the Hadamard technique was explored and we did attempt to implement it. I outline some details here including the advantages of the technique and difficulties.

We define the set of N_v Hadamard vectors to be orthogonal vectors of length N_v, defined by $v_i^\dagger \cdot v_j = 0$ if $i \neq j$ and $v_i^\dagger \cdot v_i = N_v$, where $\cdot$ is the inner product. The set can be constructed by taking the outer product of Z_n roots of unity, where n is a prime factor of N_v. For instance, if $N_v = 24 = 2^3 \cdot 3$, then one representation

of the Hadamard vectors is,

$$\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\otimes\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\otimes\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\otimes\begin{bmatrix}1 & 1 & 1\\ 1 & e^{i2\pi/3} & e^{i4\pi/3}\\ 1 & e^{i4\pi/3} & e^{i8\pi/3}\end{bmatrix}$$
$$=\begin{bmatrix}1 & 1 & 1 & 1 & \cdots\\ 1 & -1 & 1 & -1 & \cdots\\ 1 & 1 & -1 & -1 & \cdots\\ 1 & -1 & -1 & 1 & \cdots\\ \vdots & & & & \end{bmatrix}=\begin{bmatrix}h_0\\ h_1\\ h_2\\ h_3\\ \vdots\end{bmatrix} \tag{7.31}$$

where h_i are Hadamard vectors. In fact if $N_v = 24^3$, then the first 24 vectors which we have developed above will be the same, just repeated 24^2 times to fill the vector. We label these vectors from the top row down as h_0, h_1, h_2.... There is no unique way to do this mapping onto a three dimensional lattice. A nice prescription for mapping these vectors is given by $h_{ijk} = h_i \otimes h_j \otimes h_k$, where $\otimes$ is the outer product. The effect on the $G(y, y)$ propagator when including (for example) h_{111} is,

$$\sum_{\vec{y}} G(\vec{y},\vec{y}) \approx \frac{1}{2}\left[\xi_0^\dagger \times G \times \xi_0 + \xi_0^\dagger \cdot h_{111}^\dagger \times G \times h_{111}\cdot\xi_0\right]$$
$$= \xi_0^\dagger \begin{pmatrix} G_{11} & 0 & G_{13} & \cdots\\ 0 & G_{22} & 0 & \\ G_{31} & 0 & G_{33} & \\ \vdots & & & \end{pmatrix}\xi_0 \tag{7.32}$$

where we have represented the multiplication using matrices. The diagonal components are the $G(y, y)$ and the off diagonal components in its column (or row) are the noise due to $G(x, y)$. If we use the mapping of the outer product to apply the Hadamard vectors, the above solutions only remove nearest neighbour noise elements along the $(0, 0, 1)$, $(1, 0, 0)$, $(0, 1, 0)$ and $(1, 1, 1)$ directions, this is in fact the best we can do with only two Hadamard noise sources. To get full removal of nearest neighbours we need all eight vectors, h_{ijk} with $i, j, k \in \{0, 1\}$, the same as the interlacing dilution, which I will speak about in the next section. The intermediate steps in the number of Hadamard vectors add additional noise, so only when a full set of hadamard vectors are used does the noise cancellation produce good noise reduction.

The advantage of the Hadamard technique is that it can be applied incrementally to remove the noise and it can be used to remove all nearest neighbours. The solutions already obtained can be used in the improvement processes. The disadvantage is it only achieves the same level of error reduction (or better) as a new

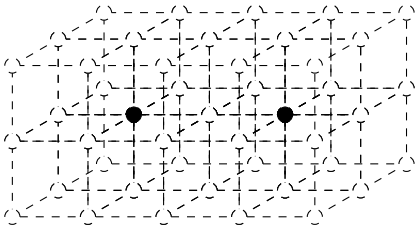


Figure 7.5: Interlacing dilution applied to the noisy source. The unfilled points correspond to masked values (set to zero) while the filled points are kept at their original values.

random source if a certain number and combination of vectors is used. This means we are required to calculate (for a 24^3 lattice) $2, 8, 27, 56...$ propagator inversions per time-slice per configuration to achieve incremental improvement of the propagator. Typically the time cost of previous calculations is almost nothing compared with the next order of calculations hence there is only a small time saving in using them.

7.5 Improvements using Interlacing

The interlacing technique was utilised in producing our results and was chosen for two reasons. The simplicity of its application and that it was already working correctly in existing code. I outline the specifics below.

If we aim to calculate the $G(y, y)$ components of the propagator with a noisy source, our estimate of $G(y, y)$ is corrupted by noise caused by the other all-to-all components $G(x, y)$ for $x \neq y$, or if we dilute in time, spin and colour, the nearest spatial components $G(\vec{x}, \vec{y})$ for $\vec{x} \neq \vec{y}$. If however we calculate the propagator with a noisy source where nearest neighbours are set to zero, we rapidly reduce the error associated with our estimate of $G(y, y)$ because the associated propagator value at a nearest neighbour point will be zero. The interlacing dilution is shown graphically in Fig. 7.5.

The error in our estimate for $G(y, y)$ is associated with the average value of $G(\vec{x}, \vec{y})$ $\vec{x} \neq \vec{y}$ and this will be reduced significantly because the nearest neighbours have the highest values in $G(x, y)$ $x \neq y$. An estimate of the error on the disconnected component is shown as part of Eq. (7.27). In masking nearest neighbours we lose the estimate for $G(y, y)$ at that point, so our propagator is incomplete, however if we calculate the solution multiple times shifting the masking to uncover all of the orthogonal components produced by the mask $G_i(y, y)$, we can fully reconstruct an error reduced propagator $G(y, y) = \sum_i G_i(y, y)$, by stitching the solution together. This technique is called interlacing.

For a lattice which has spatial dimension divisible by 2, the simplest spatial interlacing can be constructed using propagator solutions from a masked noisy source. We will assume time, spin and colour dilution. The masking function for the source is given by,

$$P(n_x, n_y, n_z) = \begin{cases} 1 & \text{if } (n_x, n_y, n_z) \bmod 2 = (0,0,0) \\ 0 & \text{else} \end{cases}$$

The offset vectors for the mask are $s = \{(0,0,0), (1,0,0), (0,1,0), (0,0,1), (1,1,0), (1,0,1), (0,1,1), (1,1,1)\}$, $P_i = P(\vec{n} + \vec{s}_i)$. To calculate the orthogonal propagator components $G_i(y,y)$, we need to calculate the $G(y,y)$ components of the propagator with the source $P_i \cdot \xi$ for random source ξ. Hence we solve $K\phi_i = P_i \cdot \xi$ for ϕ_i, where K is the Dirac matrix. Then we can produce an estimate by $G_i(y,y) \approx \phi_i \cdot (P_i \cdot \xi^*)$. The completed estimate for the $G(y,y)$ components is just the sum of these orthogonal G_i parts; the sum is done element by element, preserving the structure of the propagator.

7.6 Lattice details

The action we are using is given in Eq. (5.1). This is a small modification to the standard clover action defined in Eq. (3.10). The clover coefficient c_{SW} has been computed non-perturbatively for pure QCD, we do not include the QED clover term.

Simulations are carried out on a lattice volume of size $24^3 \times 48$, with one noise source per configuration. The sea quarks have κ values shown in Tab. 7.1 and charges of $e_u = +2/3$, $e_d = e_s = -1/3$. The strong coupling was chosen to be $\beta = 5.50$ and the electromagnetic coupling was chosen to be $e^2 = 1.25$, ten times greater than physical. These choices lead to a lattice spacing of $a = 0.068(1)$fm [34]. Further details can be found in [34] and [36]. The pion mass is $\sim 435\,\text{MeV}$.

No electromagnetic finite volume corrections have been applied to the results of this chapter, as all hadrons have zero net charge. QCD finite volume effects aren't considered in this exploratory calculation. Only gauge field stout smearing has been employed, no source or sink smearing was applied. We chose this because it is difficult to understand what smearing means when all the spatial elements of the lattice already have sources on them. Further analysis needs to be undertaken to ensure a random noisy source makes sense (does what it is supposed to do) after smearing is applied. We look forward to comparing results to this slightly more complicated case.

#	# of conf.	β	e^2	V	$\kappa_u, \ +2/3$	$\kappa_d, \ -1/3$	$\kappa_s, \ -1/3$
1	2000	5.50	1.25	$24^3 \times 48$	0.124362	0.121713	0.121713
2	1000	5.50	1.25	$24^3 \times 48$	0.124374	0.121713	0.121701
3	1000	5.50	1.25	$24^3 \times 48$	0.124400	0.121713	0.121677

Table 7.1: Summary of lattice details.

7.6.1 Ensemble selection

The chosen κ values provide an insight into the π^0–η–η' system near the SU(3) symmetric point. Ensemble 1 was tuned to correspond to the SU(3) symmetric point, where the point was defined as the set of κ which produced, $M^2_{u\bar{u}} = M^2_{d\bar{d}} = M^2_{s\bar{s}} = X^2_\pi$ (connected quark line only). This was achieved with reasonable accuracy, where the $M^2_{u\bar{u}}$ is actually slightly higher than the X^2_π while $M^2_{d\bar{d}}$ is slightly below X^2_π. Ensemble 1 has exact U-spin symmetry; the d and s quarks are identical. In this situation we would expect three distinct mass eigenvalues from the system, with the lowest state having composition $\pi^0_U = \frac{1}{\sqrt{2}}(d\bar{d} - s\bar{s})$. The difference in energy between the lowest state (π^0_U) and the next highest state (η_U) is generated by both the quark mass difference between the u and the d, s quarks and the charge difference. More pragmatically, the mass difference is generated by the difference in $M_{u\bar{u}}$ and $M_{d\bar{d}}$. If we had obtained the symmetric point using definition $M^2_{u\bar{u}} = M^2_{d\bar{d}} = M^2_{s\bar{s}} = X^2_\pi$ exactly, then there should be no mass difference between the bottom two states π^0 and η.

The κ values of the other two ensembles were chosen to try and determine the electromagnetic part of the mass difference between these two lowest states. This mass splitting can be considered the QED component of the anomaly. If we had achieved the Dashen scheme symmetric point, then ensemble 1 should correspond to this electomagnetic part, however it isn't quite at the right point. To obtain an estimate of this splitting, ensemble 2 and 3 were chosen by reducing the u quark mass by δm_u and increasing the s quark mass by the same amount, $\delta m_s = -\delta m_u$, while the d quark mass is held fixed, $\delta m_d = 0$. This ensures we remain on the $\delta m_u + \delta m_d + \delta m_s = 0$ line. The position of these points are shown in Fig. 7.6.

Assuming the lattice ensembles are at or very close to the symmetric point in the Dashen definition (refer to section 4.6) we can use this path to get a good estimate of the π^0–η mass splitting generated by the electromagnetic symmetry breaking. At leading order in SU(3) symmetry breaking, the effective field theory Lagrangian including only symmetry breaking terms has the form,

$$\mathcal{L}_{SB} = M_0^2 \mathrm{tr}(\phi\phi) + b_1 \mathrm{tr}(\phi\phi\Delta m) + c_4^{EM} \mathrm{tr}(\phi\phi Q^2) + c_5^{EM} \mathrm{tr}(\phi Q \phi Q) + c_6^{EM} (\mathrm{tr}(\phi Q))^2 \tag{7.33}$$

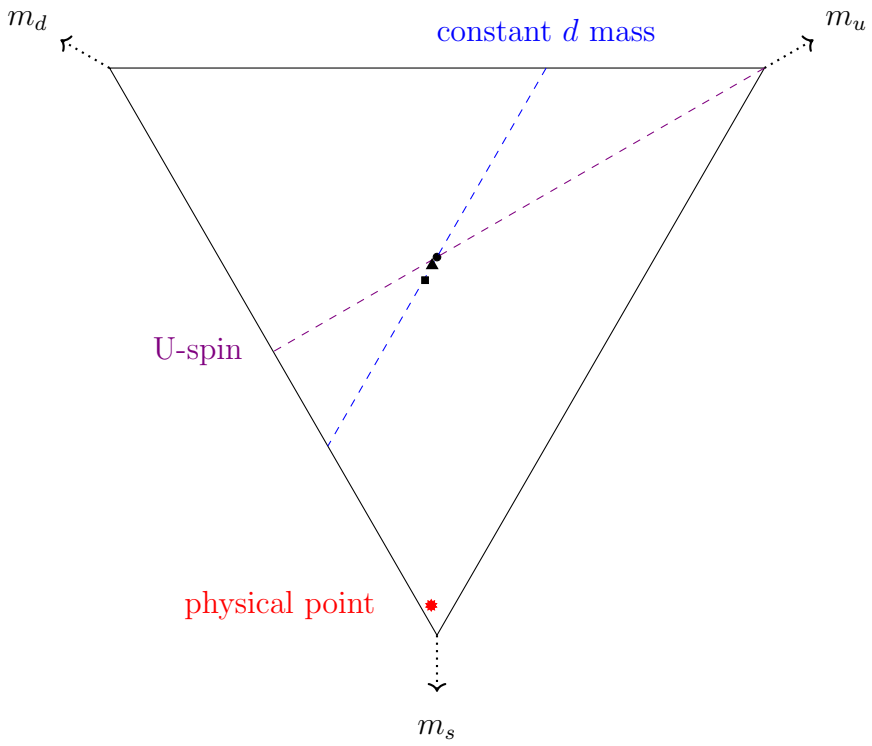


Figure 7.6: The position of the ensemble quark masses, represented graphically using $\delta m_u + \delta m_d + \delta m_s = 0$. Ensemble 1 is the circular point at the centre of the triangle where the U-spin and constant d lines intersect. Ensemble 2 and ensemble 3 are the triangle and square points respectively along the constant d line.

where $\Delta m = \text{diag}\begin{pmatrix}\delta m_u & 0 & -\delta m_u\end{pmatrix}$ and $Q = \text{diag}\begin{pmatrix}+2/3 & -1/3 & -1/3\end{pmatrix}$ and $\phi = \phi^a \lambda^a$ for Gell-Mann matrices λ^a, and pseudoscalar fields ϕ^a for $a \in \{1, 2, ..., 8\}$.

Under our Dashen scheme definition in chapter 5, $M^2_{u\bar{u}} - M^2_{d\bar{d}} = 0.0001317(37)$ at the symmetric point ($\delta m = 0$). In this calculation the coefficients have been corrected for unphysical α_{QED} using Eq. (5.12), hence on our lattice we should expect $M^2_{u\bar{u}} - M^2_{d\bar{d}} = 0.001615(45)$ at ensemble 1, however we have $M^2_{u\bar{u}} - M^2_{d\bar{d}} = 0.00243(14)$, approximately one and a half as large as the Dashen Scheme definition. Using $\alpha = 1.1703(47)$ from Tab. 5.2, we estimate that the correction of our simulation point is $4\delta\mu_u = (0.00243 - 0.001615)/1.1703 = 0.0006964$ if it was implemented under the prescription $\Delta m = \text{diag}\begin{pmatrix}2\delta\mu_u & -\delta\mu_u & -\delta\mu_u\end{pmatrix}$.

Including this correction in the above Lagrangian Eq. (7.33), the mass terms of the pseudoscalars can be determined by taking two derivatives with respect to the fields (one of the particle and one of the anti-particle) which produces,

$$\begin{aligned}
M^2_{\pi^+} &= M_0^2 + b_1(\delta m_u + \delta\mu_u) + c_4^{EM} \\
M^2_{K^+} &= M_0^2 + b_1\delta\mu_u + c_4^{EM} \\
M^2_{K^0} &= M_0^2 - b_1(\delta m_u + 2\delta\mu_u) \\
M^2_{\eta_\pm} &= M_0^2 + \frac{2}{3}c_6^{EM} \pm \frac{2}{3}\sqrt{3b_1^2\delta m_u^2 + 3b_1 c_6^{EM}\delta m_u + (c_6^{EM})^2} \\
&\quad + \delta\mu_u(9b_1^2\delta\mu_u + 9b_1^2\delta m_u + 6b_1 c_6^{EM})
\end{aligned} \tag{7.34}$$

where we have defined K^0 to have no electromagnetic component. Along the path $\delta m_d = 0$, the $M_{u\bar{u}} = M_{s\bar{s}}$ point should correspond to a minimum in the mass curves of η_+ and η_-. While the electromagnetic component of the splitting is defined at a point when $\delta\mu_u = \delta m_u = 0$, hence $M^2_{\eta_+} - M^2_{\eta_-} = \frac{4}{3}c_6^{EM}$.

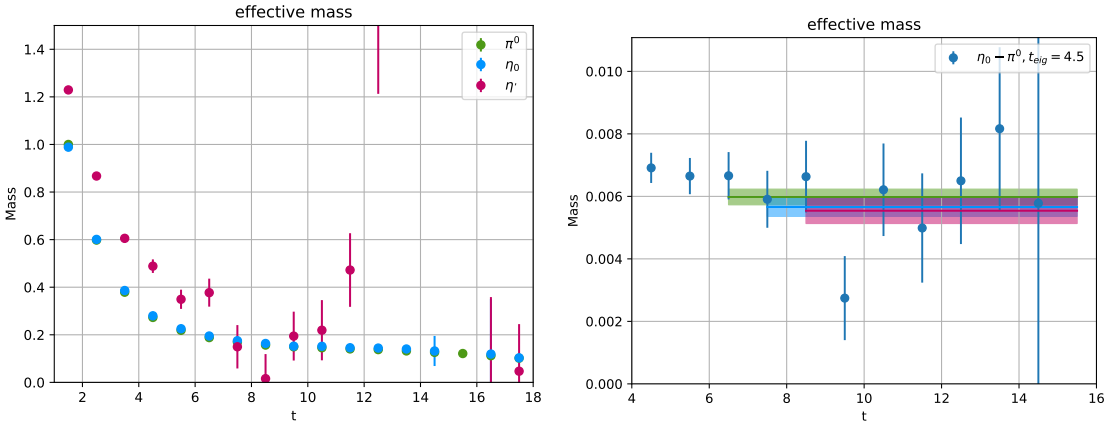
7.6.2 Improving the signal

Noise in the signal of the π^0–η–η' system primarily originates in the disconnected component of the correlation function. To obtain an accurate estimate the disconnected component we need a significant increase in statistics. This is normally achieved by determining an estimate for the self-to-self propagator $G(x;x)$, to get the most out of each configuration. In this study we use noisy $\mathbb{Z}_2$ wall sources to determine this estimate, as wall sources are the fastest method in producing an estimate of $G(x;x)$ and have been proven to be successful [91, 92, 93, 94, 85, 86, 96, 69]. Noisy sources produce only an estimate of the self-to-self propagator $G(x;x)$, and introduce an error in $G(x;x)$. These errors can be reduced using dilution and in this study we employ spin, colour and time dilution, as well as a spatial dilution of the nearest neighbour elements using interlacing, as described above. While the disconnected part is calculated on each time-slice, the connected part is only

calculated (using the one-end trick) once per configuration from $t = 0$ to all other time-slices. This saves on inversions, as the connected part is not the component of the analysis that contributes the most noise.

7.7 Results

Once we have obtained estimates for the propagators, we calculate the correlation functions of the flavour diagonal states, which populate the correlation matrix, $C_O(t)$ of Eq. (7.20). Using the symmetry of the forward and backwards propagating states in the time dimension we double the statistics of $C_O(t)$. We diagonalise the matrix $C_O(t)^{-1}C_O(t+\delta t)$, where $\delta t = 1$, to produce $C(t)^{-1}C(t+\delta t)$[4]. The eigenvectors for the diagonalisation are only determined at a single time, t_{eig} and these vectors are assumed to diagonalise the correlation matrix, $C_O(t)$ for all later time slices. We chose the time $t_{eig} = 4.5$ because it is the latest time where the η' state is well resolved. We take the latest time that gives good resolution (to at least three significant figures) of the state vector, which is approximately the resolution required to determine the mixing. This choice is a compromise between noise in the disconnected part and excited state contamination.

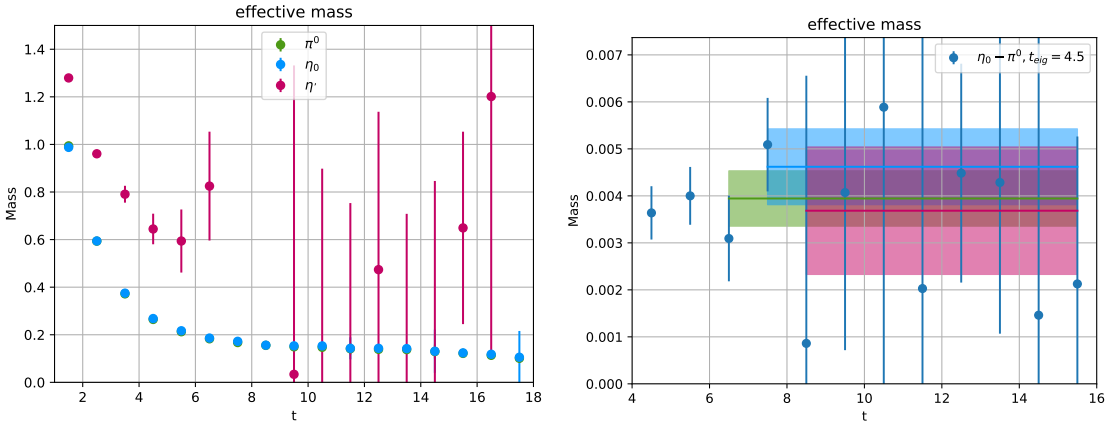


(a) (b)

Figure 7.7: Ensemble 1: (a) Effective mass plot of the neutral psuedoscalar mesons. (b) Effective mass plot of the energy difference between the two lowest states of the neutral pseudoscalar mesons.

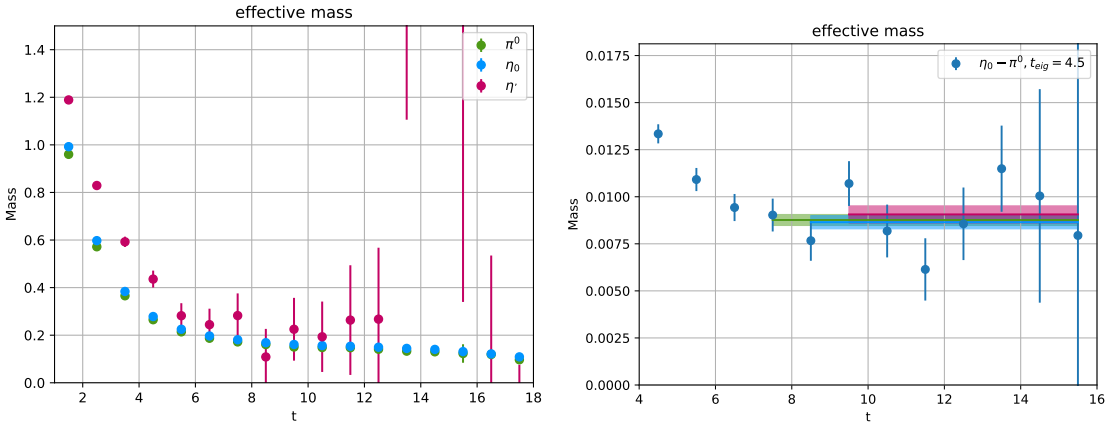
Taking the negative log of this diagonalised matrix $C(t)^{-1}C(t+\delta t)$, we produce effective mass plots. The eigenvectors from the diagonalisation, v give the flavour

[4]Note this defines an effective mass point at t+0.5 when $\delta t = 1$.



(a) (b)

Figure 7.8: Ensemble 2: (a) Effective mass plot of the neutral psuedoscalar mesons. (b) Effective mass plot of the energy difference between the two lowest states of the neutral pseudoscalar mesons.



(a) (b)

Figure 7.9: Ensemble 3: (a) Effective mass plot of the neutral psuedoscalar mesons. (b) Effective mass plot of the energy difference between the two lowest states of the neutral pseudoscalar mesons.

composition of each state. The effective mass plots are shown in Fig. 7.7a, Fig. 7.8a and Fig. 7.9a showing all three states on each ensemble. We always name the energy eigenstates in these plots π^0, η and η' in energy order irrespective of the state composition. We see in the figures that we have not obtained a good estimate of the η', but that at the time of diagonalisation $t_{eig} = 4.5$ the states are well

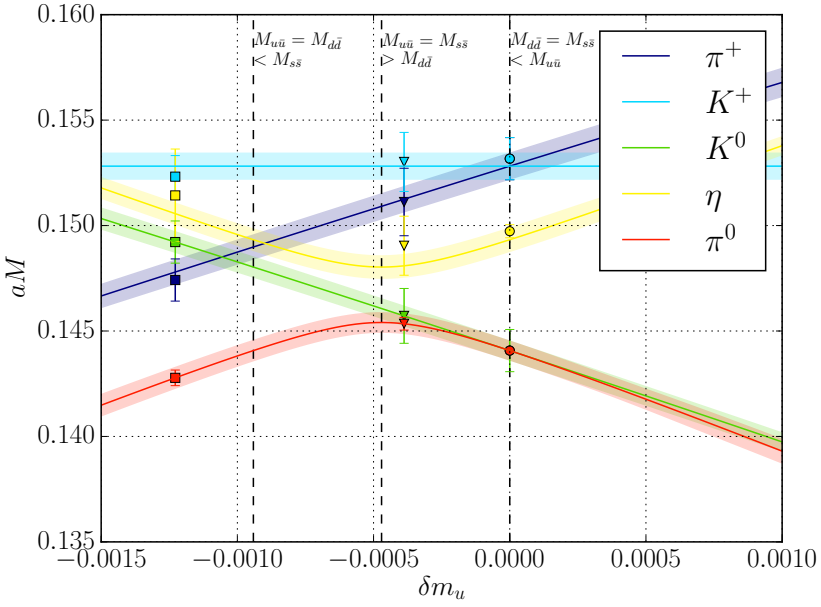


Figure 7.10: The π^0–η system at the symmetric point, fit using Eq. (7.34). The fit was obtained from mass splittings of the particles using K^0 to fix the mass constant M_0^2. With this in mind, the π^0 data points and the K^0 points are fixed to the fitting lines, and the other masses are shifted to ensure the splittings are properly represented.

resolved. The states are not completely free of excited state contamination at this time.

As we are at the SU(3) symmetric point, we would expect the connected parts of the η' (under the assumption it is a singlet) to add up approximately to the same number as at the physical point. The anomaly should also be generated in a similar way to the physical point. Hence, we expect the η' to have approximately (within two standard deviations) the physical η' mass on all three ensembles. Using the lattice spacing $a = 0.068(1)\,\mathrm{fm}$, we can determine that our estimates for the η' at $t = 5.5$ are in agreement[5]with the physical η' mass (within two standard deviations) and that the π^0 and η are close to the π^+ mass of 442 MeV.

The ratio of the two lowest energy states of the diagonalised matrix $C(t)^{-1}C(t+\delta t)$ were used to produce an estimate for the mass splitting between these states. Fig. 7.7b, Fig. 7.8b and Fig. 7.9b show the mass splitting between the lower two states. The uncertainty on the points is significantly smaller than the for the η'

[5]Though we admit that the signal is not of sufficient quality to give an estimate of the η' mass, as a reality check, the result does not seem to be completely erroneous (does not disagrees with the physical mass).

in the corresponding figure (a). This is because the π^0 and η receive only a small contribution to their mass from the disconnected term and this partially cancels when we take the difference of their masses. These mass differences, together with estimates for π^+, K^+, K^0, are then used to fit the coefficients in Eq. (7.34). The fit to these points is shown in Fig. 7.10.

In this plot the fit is obtained from the mass splittings while the overall mass M_0^2 in Eq. (7.34) is obtained from the K^0 mass. When we plot the data points, the π^0 and K^0 have been fixed to their respective fit lines, and the remaining data points are determined by their mass splittings. This allows us to compare data points from the different ensembles constructively with each other and the fit.

The dotted vertical lines correspond to the centre of the three regimes on the plot. The right most dashed line is the point where the $M_{d\bar{d}} = M_{s\bar{s}}$. It corresponds to the data points of ensemble 1. Here the $M_{u\bar{u}}$ is heavier than the other flavour diagonal states. The eigenvectors near this line correspond most closely to U-spin states. The centre dashed line corresponds to $M_{u\bar{u}} = M_{s\bar{s}}$ degeneracy, where the $M_{d\bar{d}}$ is lighter than the other flavour basis states. Near this line, the states correspond most closely to V-spin, however because the degeneracy is between two heavier quarks, the lightest state is the $\eta_V = \frac{1}{\sqrt{6}}(u\bar{u} + s\bar{s} - 2d\bar{d})$. This line corresponds to the minimum, which is located at $\delta m_u = -0.0004706$, and the difference between π^0 and η at this point is 0.00263(31) in lattice units. The left most dotted line corresponds to the $M_{u\bar{u}} = M_{d\bar{d}}$ degeneracy, where $M_{s\bar{s}}$ is heavier. States near this line most closely correspond to isospin states. The electromagnetic splitting is determined when $\delta m_u = \delta\mu_u = 0$ and is given by $M_\eta - M_{\pi^0} = \frac{4}{3}\frac{c_6^{EM}}{M_\eta + M_{\pi^0}} = 0.554(78)\,\text{MeV}$ after correcting for unphysical α_{QED} using Eq. (5.12).

The eigenstates of each of the ensembles are shown in Tab. 7.2 in flavour basis and in Tab. 7.3 in the three SU(2) symmetry bases, U-spin, V-spin and isospin. These tables support our interpretation, where most of the states in Tab. 7.3 have a probability of greater than 98% to be as prescribed above. This shows that the flavour diagonal meson masses $M_{q\bar{q}}$ (connected quark line only) can be useful in determining the state behaviour, and hence are a good measure of flavour symmetry breaking within the system.

As you can see in Fig. 7.10 the masses of the π^0 and η, rather than being straight lines, turn a corner (called an avoided level crossing) at the $M_{u\bar{u}} = M_{s\bar{s}}$ point. This sort of behaviour is only possible for flavour-neutral states because their flavour composition is not fixed. The change in the mass trajectory of the two lowest eigenstates corresponds to a reorganisation of state composition. The states change along our path; the lowest energy state transitions from $\pi_U^0 = \frac{1}{\sqrt{2}}(d\bar{d} - s\bar{s})$ to $\eta_V = \frac{1}{\sqrt{6}}(u\bar{u} + s\bar{s} - 2d\bar{d})$ to $\pi^0 = \frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})$. The higher energy state changes from $\eta_U = \frac{1}{\sqrt{6}}(d\bar{d} + s\bar{s} - 2u\bar{u})$ to $\pi_V^0 = \frac{1}{\sqrt{2}}(u\bar{u} - s\bar{s})$ to $\eta = \frac{1}{\sqrt{6}}(u\bar{u} + d\bar{d} - 2s\bar{s})$. Along

	Basis	$\langle q\bar{q} \left\vert \pi^0 \right\rangle^2$	$\langle q\bar{q} \left\vert \eta \right\rangle^2$	$\langle q\bar{q} \left\vert \eta' \right\rangle^2$
Ensemble 1	$u\bar{u}$	0.000(1)	0.6849(24)	0.3151(24)
	$d\bar{d}$	0.500(1)	0.1576(12)	0.3424(12)
	$s\bar{s}$	0.500(1)	0.1576(12)	0.3424(12)
Ensemble 2	$u\bar{u}$	0.167(76)	0.537(76)	0.2962(30)
	$d\bar{d}$	0.646(18)	0.000(18)	0.3540(17)
	$s\bar{s}$	0.187(62)	0.463(62)	0.3498(19)
Ensemble 3	$u\bar{u}$	0.5994(95)	0.1142(84)	0.2864(49)
	$d\bar{d}$	0.3950(12)	0.2600(12)	0.3448(21)
	$s\bar{s}$	0.0055(27)	0.6256(54)	0.3688(46)

Table 7.2: The energy eigenstates in the flavour basis. Note that in the top row, energy eigenstates are labelled π^0, η and η' based solely on their energy not their state composition, with π^0 corresponding to the lowest energy, then η then η'.

our $\delta m_d = 0$ trajectory, the $M_{u\bar{u}} = M_{s\bar{s}}$ point is the closest to SU(3) symmetry. We know this because at the SU(3) point, the mass difference between the two lowest states is zero. As the states head towards the SU(3) point the energy difference between them must reduce. Additionally, we would expect the u and s quark to be equally present in these two lowest states when $M_{u\bar{u}} = M_{s\bar{s}}$, hence the derivative of the π^0 and η masses with respect to δm_u along the $\delta m_d = 0$ path should be zero at this point.

A similar reorganisation is expected near the physical point, when u and d quarks are approximately mass degenerate. Consider the trajectory $\delta m_s = 0$, $\delta m_u = -\delta m_d$ around a point of exact isospin symmetry $M_{u\bar{u}} = M_{d\bar{d}}$ [6]where $M_{s\bar{s}} \gg M_{u\bar{u}}$. At an exact isospin symmetry point, both states have an equal amount of u and d quark, hence along the path the masses of the π^0 and η to leading order do not change. The exact isospin point corresponds to an extrema in the π^0 and η mass trajectories. With this in mind, the analysis in Fig. 7.10 can be viewed from a different perspective. This plot can also characterise the $\delta m_s = 0$ path, as at our starting point (ensemble 1) the d and s quarks are degenerate. Under this assignment of names, the minimum and turning point in the plot occurs when $M_{u\bar{u}} = M_{d\bar{d}}$. Though the size of the level repulsion may change, the general shape remains the same near the physical isospin point.

Near the physical point, as we move away from the exact isospin symmetry, the reorganisation of states will change the lower state, π^0, to have either more u or d quark component. The higher energy state η will be effected in the opposite way.

[6]This definition of isospin symmetry is only 'exact' for the π^0–η–η' system, and not generally for all hadrons because of the symmetry breaking generated by the charges. Exact isospin symmetry for all hadrons can only be guarrenteed in the charge degenerate limit.

	Basis, ϕ	$\langle\phi \left\vert\pi^0\right\rangle^2$	$\langle\phi \left\vert\eta\right\rangle^2$	$\langle\phi \left\vert\eta'\right\rangle^2$
Ensemble 1	π^0_U	1.0000(1)	0.0000(1)	0.0000(1)
	η_U	0.0000(1)	0.999623(83)	0.000377(83)
	η'_U	0.0000(1)	0.000377(83)	0.999623(83)
Ensemble 2	η_V	0.999(23)	0.000(23)	0.000489(84)
	π^0_V	0.000(23)	0.999(23)	0.00112(20)
	η'_V	0.00046(23)	0.00114(28)	0.9984(27)
Ensemble 3	π^0_T	0.9838(53)	0.0148(53)	0.00136(32)
	η_T	0.0145(52)	0.9841(53)	0.00142(42)
	η'_T	0.00169(42)	0.00108(31)	0.99722(69)

Table 7.3: The energy eigenstates in the exact SU(3) basis Eq. (7.17), where the basis states are aligned with U-spin, V-spin or isospin "T". Note that in the top row, energy eigenstates are labelled π^0, η and η' based solely on their energy not their state composition, with π^0 corresponding to the lowest energy, then η then η'.

The π^0 would have a higher d quark component when $M_{u\bar{u}} > M_{d\bar{d}}$. This would occur if the quark masses are the same, but the u quark has double the electric charge of the d quark. In the case $M_{u\bar{u}} < M_{d\bar{d}}$, one would expect the π^0 state to have a slightly higher u quark component. This case applies to the physical point.

7.8 Summary

We were able to determine the energy difference between π^0 and η at three points in quark mass parameter space (ensembles 1,2 and 3) near the SU(3) flavour degenerate point. We modelled this system using effective field theory to produce hadron mass expansions in terms of quark mass and charge. By fitting the model to the lattice data we were able to determine an estimate for the mass difference generated by electromagnetic isospin violation under the Dashen scheme definition.

Using the eigenvectors near the SU(3) symmetric point and our model we were able to infer the behaviour of the states at a point close to the physical point, where isospin is almost an exact symmetry. In this case the state that mixes most strongly with the heavy η' state will shift (continuously) from the up quark dominated state to the down quark dominated state as we move through the quark mass degenerate point towards the physical point.

We have also explored the nature of SU(3) flavour breaking within this system. Definitions of flavour degeneracy in terms of purely quark mass and quark charge begin to become difficult when both are included. It is interesting to note that the

degeneracy of quark flavours can be measured by the degeneracy of the flavour-diagonal mesons $M_{q\bar{q}}$ (connected quark line only). The masses of the flavour-diagonal mesons provide a useful definition of the SU(3) flavour symmetry breaking in QCD+QED.

CHAPTER 8

Conclusion

Isospin violation is an important effect when considering various physical phenomena, some examples include the flavour decomposition of nucleon structure, neutrino-nucleus interactions, precision constraints on the CKM matrix and quark masses. In this thesis I described how we used lattice QCD+QED to determine estimates of the mass splittings generated by isospin violation. The novel aspect of these studies being the dynamical inclusion of QED on the lattice, and hence the inclusion of these effects in our estimates. The QED contribution to the masses is important when determining isospin violating effects.

We presented isospin violating mass splitting for the octet and decuplet baryons. The estimates were extrapolated to the physical quark mass using a simplified analytic expansion from the SU(3) point. These are the first lattice estimates of isospin violating effects on the decuplet baryons which include QED dynamically. We found that the electromagnetic effects are of a similar size to the strong isospin breaking effects, producing contributions of approximately $\sim 5\,\mathrm{MeV}$. These electromagnetic contributions play an important role in the mass splitting of isospin multiplets.

The major uncertainty in the mass splittings of the light hadron spectrum was from the electromagnetic effects. In future work on the light hadron spectrum, improved resolution can be obtained by including more points and conducting a detailed analysis on the effects of quark source smearing to obtain a better signal on the correlation functions. Improvements in the strong component would most likely improve the resolution overall, and this can be obtained by the previously mentioned methods and by having data closer to the physical point. This reduces the extrapolation distance of the quark masses, reducing uncertainty especially in the strong components of the mass splittings. It would also be interesting to

perform partially-quenched simulations on a lattice ensemble with quark masses near the physical point. In this case the sea quark parameters would not have to be included. The difference between an analysis at the physical point and our current analysis may reveal short comings in the current analysis.

We determined the isospin violating mass splitting of the charmed baryon spectrum including QED. Ours is the first calculation to include QED dynamically. We found that the electromagnetic effects were comparable to the strong isospin breaking effects, contributing at a level of $\sim 5\,\text{MeV}$. Another motivation in calculating the isospin violating mass splittings was to investigate the discrepancy between the SELEX $\Xi_{cc}^{+}(3520)$ and the LHCb estimate of $\Xi_{cc}^{++}(3621)$. The mass difference is quite large if it is to be attributed solely to an isospin violating effect. We calculated that the isospin violating effects are quite small, on the order of $5\,\text{MeV}$, and so this discrepancy could not be described through isospin violation. We investigated if the discrepancy could be explained by hyperfine mass splitting. Our results showed that the discrepancy could be explained in this way within our uncertainties. One resolution to the discrepancy may be assigning these two measurements with different spin.

As future work for charmed spectroscopy, the main goal is to reduce the uncertainty on the electromagnetic components of the isospin violating splittings, as these are significant. This would involve using more charge types and expanding the number of data points we fit to. More generally, the fits to the spectra easily fall within the large error bars, and reducing the uncertainty in the lattice data would help constrain these fits better. Reducing these uncertainties may involve improving interpolator overlap by applying the variational method to quark source smearing, as well as increasing statistics.

Now that we have produced some estimates of what the charge splittings should be in the charm spectrum using a fairly conservative method, namely the expansion of the light quark masses about the physical charm point, it would be interesting to apply a combined fit using SU(3) or SU(4) points. This can very quickly increase the number of points to fit to, hence constraining fit coefficients, while (however) adding uncertainty as to the expansion's convergence.

We studied the neutral flavour pseudoscalar mesons π^0, η and η' near the SU(3) symmetric point. We determined how large the mass splitting between π^0 and η would be if it was purely generated by QED effects at the SU(3) symmetric point. More generally we investigated the state changing behaviour of the π^0–η–η' system near a point of exact degeneracy. We used the knowledge obtained at the SU(3) point to infer the behaviour of the states at a point close to the physical point. We also showed that the diagonal flavour meson masses $M_{q\bar{q}}$, though purely lattice constructions, are a useful way to define SU(3) flavour symmetry breaking for hadronic systems when including both QCD and QED.

It would be very interesting to extend this work in lattice QCD+QED to include the vector mesons ρ, ω and ϕ at the SU(3) point (where they are stable). The disconnected contributions of the vector mesons are much smaller than that of the pseudoscalar mesons, which complies with the OZI rule [88, 89, 90] in the absence of the axial anomaly and seen in the small mixing angles in lattice QCD [70, 73]. We would need a significant increase in the signal to attempt to extract the mass eigenstates. One way to do this is to calculate on a larger volume, so that our determination of the disconnected part is obtained at more points, and hence higher accuracy. Using larger volumes increases the computation time significantly, while also improving the efficacy of the noisy source technique. The strategy for dilution may need to be reconsidered to help reduce the computation time. Alternative distillation methods and their variants may prove beneficial, for example Ref. [73].

Our attempt here to understand state changing behaviour in the π^0–η–η' system only goes so far in increasing our understanding of these states at the physical point. As the neutral flavour mesons do not have fixed flavour composition as quark masses change, the most straightforward way to calculate the masses and flavour composition is at physical quark masses and charges. This is a rather long term goal, however, as the time required to calculate propagators is already significant.

In summary, we produced some of the first results to include QED in calculation of isospin violating mass splittings, considering octet and decuplet baryons, charmed baryons and neutral flavour pseudoscalar mesons. While these results have provided a foundation, further work is needed to more precisely isolate these splittings.

Bibliography

[1] Gerald A. Miller, Allena K. Opper, and Edward J. Stephenson. Charge symmetry breaking and QCD. *Ann. Rev. Nucl. Part. Sci.*, 56:253–292, 2006.

[2] J. T. Londergan, J. C. Peng, and A. W. Thomas. Charge Symmetry at the Partonic Level. *Rev. Mod. Phys.*, 82:2009–2052, 2010.

[3] Michael Wagman and Gerald A. Miller. Charge Symmetry Breaking and Parity Violating Electron-Proton Scattering. *Phys. Rev.*, C89(6):065206, 2014. [Erratum: Phys. Rev.C91,no.1,019903(2015)].

[4] P. E. Shanahan, R. Horsley, Y. Nakamura, D. Pleiter, P. E. L. Rakow, G. Schierholz, H. Stben, A. W. Thomas, R. D. Young, and J. M. Zanotti. Charge symmetry violation in the electromagnetic form factors of the nucleon. *Phys. Rev.*, D91(11):113006, 2015.

[5] G. P. Zeller et al. A Precise determination of electroweak parameters in neutrino nucleon scattering. *Phys. Rev. Lett.*, 88:091802, 2002. [Erratum: Phys. Rev. Lett.90,239902(2003)].

[6] W. Bentz, I. C. Cloet, J. T. Londergan, and A. W. Thomas. Reassessment of the NuTeV determination of the weak mixing angle. *Phys. Lett.*, B693:462–466, 2010.

[7] Nicola Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–533, 1963. [,648(1963)].

[8] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[9] Vincenzo Cirigliano and Helmut Neufeld. A note on isospin violation in Pl2(gamma) decays. *Phys. Lett.*, B700:7–10, 2011.

[10] Wolfgang Lucha, Dmitri Melikhov, and Silvano Simula. Isospin breaking in the decay constants of heavy mesons from QCD sum rules. *Phys. Lett.*, B765:365–370, 2017.

[11] V. Cirigliano, H. Neufeld, and H. Pichl. K(e3) decays and CKM unitarity. *Eur. Phys. J.*, C35:53–65, 2004.

[12] J. Gasser and H. Leutwyler. Quark Masses. *Phys. Rept.*, 87:77–169, 1982.

[13] H. Leutwyler. Bounds on the light quark masses. *Phys. Lett.*, B374:163–168, 1996.

[14] J. Gasser, A. Rusetsky, and I. Scimemi. Electromagnetic corrections in hadronic processes. *Eur. Phys. J.*, C32:97–114, 2003.

[15] Gilberto Colangelo, Stefan Lanz, Heinrich Leutwyler, and Emilie Passemar. $\eta \to 3\pi$: Study of the Dalitz plot and extraction of the quark mass ratio Q. *Phys. Rev. Lett.*, 118(2):022001, 2017.

[16] M. Mattson et al. First observation of the doubly charmed baryon Xi+(cc). *Phys. Rev. Lett.*, 89:112001, 2002.

[17] Roel Aaij et al. Observation of the doubly charmed baryon Ξ_{cc}^{++}. *Phys. Rev. Lett.*, 119(11):112001, 2017.

[18] W. Rindler. *Relativity.* Oxford University Press, 2nd edition, 2006.

[19] J. Aitchison and J. Hey. *Gauge Theories in Particle Physics*, volume 1. Institute of Physics Publishing, 3nd edition, 2003.

[20] M. Peskin and D. Schroeder. *An Introduction to Quantum Field Theory.* Westview Press, 1st edition, 1995.

[21] J. D. Bjorken and E. A. Paschos. Inelastic electron-proton and γ-proton scattering and the structure of the nucleon. *Phys. Rev.*, 185:1975–1982, Sep 1969.

[22] J. D. Bjorken. Inequality for backward electron- and muon-nucleon scattering at high momentum transfer. *Phys. Rev.*, 163:1767–1769, Nov 1967.

[23] Richard P. Feynman. Very high-energy collisions of hadrons. *Phys. Rev. Lett.*, 23:1415–1417, Dec 1969.

[24] F. E. Close. *An Introduction to Quark and Partons.* Academic Press Inc. (London), 1st edition, 1979.

[25] A. Zee. *Quantum Field Theory in a Nutshell*. Princeton University Press, 2nd edition, 2010.

[26] H. Rothe. *Lattice Gauge Theories*. Springer, 3rd edition, 2005.

[27] C. Lang and C. Gattringer. *Quantum Chromodynamics on the Lattice*. World Scientific Publishing, 1st edition, 2010.

[28] B. Sheikholeslami and R. Wohlert. Improved Continuum Limit Lattice Action for QCD with Wilson Fermions. *Nucl. Phys.*, B259:572, 1985.

[29] N. Cundy et al. Non-perturbative improvement of stout-smeared three flavour clover fermions. *Phys. Rev.*, D79:094507, 2009.

[30] Colin Morningstar and Mike J. Peardon. Analytic smearing of SU(3) link variables in lattice QCD. *Phys. Rev.*, D69:054501, 2004.

[31] N. Cundy et al. Clover improvement for stout-smeared 2+1 flavour SLiNC fermions: Non-perturbative results. *PoS*, LATTICE2008:132, 2008.

[32] K. Symanzik. Continuum limit and improved action in lattice theories: (i). principles and 4 theory. *Nuclear Physics B*, 226(1):187 – 204, 1983.

[33] W. Bietenholz et al. Flavour blindness and patterns of flavour symmetry breaking in lattice simulations of up, down and strange quarks. *Phys. Rev.*, D84:054509, 2011.

[34] R. Horsley et al. QED effects in the pseudoscalar meson sector. *JHEP*, 04:093, 2016.

[35] A. Duncan, E. Eichten, and H. Thacker. Electromagnetic splittings and light quark masses in lattice QCD. *Phys. Rev. Lett.*, 76:3894–3897, 1996.

[36] R. Horsley et al. Isospin splittings of meson and baryon masses from three-flavor lattice QCD + QED. 2015.

[37] Roger F. Dashen. Chiral SU(3) x SU(3) as a symmetry of the strong interactions. *Phys. Rev.*, 183:1245–1260, 1969.

[38] R. Horsley, Y. Nakamura, D. Pleiter, P. E. L. Rakow, G. Schierholz, H. Stben, R. D. Young, and J. M. Zanotti. Electromagnetic splitting of quark and pseudoscalar meson masses from dynamical QCD + QED. *PoS*, Lattice2013:499, 2014.

[39] Masashi Hayakawa and Shunpei Uno. QED in finite volume and finite size scaling effect on electromagnetic properties of hadrons. *Prog. Theor. Phys.*, 120:413–441, 2008.

[40] Sz. Borsanyi et al. Ab initio calculation of the neutron-proton mass difference. *Science*, 347:1452–1455, 2015.

[41] Jong-Wan Lee and Brian C. Tiburzi. Finite Volume Corrections to the Electromagnetic Mass of Composite Particles. *Phys. Rev.*, D93(3):034012, 2016.

[42] Zohreh Davoudi and Martin J. Savage. Finite-Volume Electromagnetic Corrections to the Masses of Mesons, Baryons and Nuclei. *Phys. Rev.*, D90(5):054503, 2014.

[43] R. D. Young. Infrared features of dynamical QED + QCD simulations. *Lattice 2016, 34th International Symposium on Lattice Field Theory*, 2016.

[44] Andre Walker-Loud, Carl E. Carlson, and Gerald A. Miller. The Electromagnetic Self-Energy Contribution to $M_p - M_n$ and the Isovector Nucleon Magnetic Polarizability. *Phys. Rev. Lett.*, 108:232301, 2012.

[45] Sz. Borsanyi et al. Isospin splittings in the light baryon octet from lattice QCD and QED. *Phys. Rev. Lett.*, 111(25):252001, 2013.

[46] F. B. Erben, P. E. Shanahan, A. W. Thomas, and R. D. Young. Dispersive estimate of the electromagnetic charge symmetry violation in the octet baryon masses. *Phys. Rev.*, C90(6):065205, 2014.

[47] Ross Daniel Young, Derek Bruce Leinweber, Anthony William Thomas, and Stewart Victor Wright. Chiral analysis of quenched baryon masses. *Phys. Rev.*, D66:094507, 2002.

[48] Vladimir Pascalutsa and Marc Vanderhaeghen. The Nucleon and delta-resonance masses in relativistic chiral effective-field theory. *Phys. Lett.*, B636:31–39, 2006.

[49] R. E. Cutkosky. Isospin splitting in the baryon octet and decuplet. *Phys. Rev.*, C47:367–371, 1993.

[50] A. B. Gridnev, I. Horn, W. J. Briscoe, and I. I. Strakovsky. The K-matrix approach to the Delta - resonance mass splitting and isospin violation in low-energy pi-N scattering. *Phys. Atom. Nucl.*, 69:1542–1551, 2006.

[51] E. Pedroni et al. A Study of Charge Independence and Symmetry from pi+ and pi- Total Cross-Sections on Hydrogen and Deuterium Near the 3,3 Resonance. *Nucl. Phys.*, A300:321–347, 1978.

[52] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.

[53] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev.*, D98(3):030001, 2018.

[54] R. E. Cutkosky. Isospin splitting in the baryon octet and decuplet. 1992.

[55] A. Ocherashvili et al. Confirmation of the double charm baryon Xi+(cc)(3520) via its decay to p D+ K-. *Phys. Lett.*, B628:18–24, 2005.

[56] Bernard Aubert et al. Search for doubly charmed baryons Xi(cc)+ and Xi(cc)++ in BABAR. *Phys. Rev.*, D74:011103, 2006.

[57] R. Chistov et al. Observation of new states decaying into Lambda(c)+ K- pi+ and Lambda(c)+ K0(S) pi-. *Phys. Rev. Lett.*, 97:162001, 2006.

[58] R Aaij et al. Search for the doubly charmed baryon Ξ_{cc}^{+}. *JHEP*, 12:090, 2013.

[59] Y. Namekawa et al. Charmed baryons at the physical point in 2+1 flavor lattice QCD. *Phys. Rev.*, D87(9):094512, 2013.

[60] Paula Prez-Rubio, Sara Collins, and Gunnar S. Bali. Charmed baryon spectroscopy and light flavor symmetry from lattice QCD. *Phys. Rev.*, D92(3):034504, 2015.

[61] Constantia Alexandrou and Christos Kallidonis. Low-lying baryon masses using $N_f = 2$ twisted mass clover-improved fermions directly at the physical pion mass. *Phys. Rev.*, D96(3):034511, 2017.

[62] Zachary S. Brown, William Detmold, Stefan Meinel, and Kostas Orginos. Charmed bottom baryon spectroscopy from lattice QCD. *Phys. Rev.*, D90(9):094507, 2014.

[63] Raul A. Briceno, Huey-Wen Lin, and Daniel R. Bolton. Charmed-Baryon Spectroscopy from Lattice QCD with $N_f = 2 + 1 + 1$ Flavors. *Phys. Rev.*, D86:094504, 2012.

[64] C. Alexandrou, J. Carbonell, D. Christaras, V. Drach, M. Gravina, and M. Papinutto. Strange and charm baryon masses with two flavors of dynamical twisted mass fermions. *Phys. Rev.*, D86:114501, 2012.

[65] Liuming Liu, Huey-Wen Lin, Kostas Orginos, and Andre Walker-Loud. Singly and Doubly Charmed J=1/2 Baryon Spectrum from Lattice QCD. *Phys. Rev.*, D81:094505, 2010.

[66] R. Horsley, Z. Koumi, Y. Nakamura, H. Perlt, P. E. L. Rakow, G. Schierholz, A. Schiller, H. Stuben, R. D. Young, and J. M. Zanotti. Charmed states and flavour symmetry breaking. *EPJ Web Conf.*, 175:06017, 2018.

[67] C. Alexandrou, V. Drach, K. Jansen, C. Kallidonis, and G. Koutsou. Baryon spectrum with $N_f = 2 + 1 + 1$ twisted mass fermions. *Phys. Rev.*, D90(7):074501, 2014.

[68] Y. Kuramashi, M. Fukugita, H. Mino, M. Okawa, and A. Ukawa. eta-prime meson mass in lattice QCD. *Phys. Rev. Lett.*, 72:3448–3451, 1994.

[69] N. H. Christ, C. Dawson, T. Izubuchi, C. Jung, Q. Liu, R. D. Mawhinney, C. T. Sachrajda, A. Soni, and R. Zhou. The η and η' mesons from Lattice QCD. *Phys. Rev. Lett.*, 105:241601, 2010.

[70] Jozef J. Dudek, Robert G. Edwards, Balint Joo, Michael J. Peardon, David G. Richards, and Christopher E. Thomas. Isoscalar meson spectroscopy from lattice QCD. *Phys. Rev.*, D83:111502, 2011.

[71] Eric B. Gregory, Alan C. Irving, Christopher M. Richards, and Craig McNeile. Study of the η and η' mesons with improved staggered fermions. *Phys. Rev. D*, 86:014504, Jul 2012.

[72] Chris Michael, Konstantin Ottnad, and Carsten Urbach. η and η' mixing from Lattice QCD. *Phys. Rev. Lett.*, 111(18):181602, 2013.

[73] Jozef J. Dudek, Robert G. Edwards, Peng Guo, and Christopher E. Thomas. Toward the excited isoscalar meson spectrum from lattice qcd. *Phys. Rev. D*, 88:094505, Nov 2013.

[74] Gunnar Bali, Sara Collins, and Jakob Simeth. η and η' masses and decay constants. *EPJ Web Conf.*, 175:05028, 2018.

[75] Andrey Yu. Kotov, Maria Paola Lombardo, and Anton M. Trunin. Fate of the in the quark gluon plasma. *Physics Letters B*, 794:83 – 88, 2019.

[76] Konstantin Ottnad. Pseudoscalar flavor-singlet mesons from lattice QCD. In *9th International Workshop on Chiral Dynamics (CD18) Durham, NC, USA, September 17-21, 2018*, 2019.

[77] Stephen L. Adler. Axial vector vertex in spinor electrodynamics. *Phys. Rev.*, 177:2426–2438, 1969. [,241(1969)].

[78] Kazuo Fujikawa. Path Integral Measure for Gauge Invariant Fermion Theories. *Phys. Rev. Lett.*, 42:1195–1198, 1979.

[79] Kazuo Fujikawa. Path Integral for Gauge Theories with Fermions. *Phys. Rev.*, D21:2848, 1980. [Erratum: Phys. Rev.D22,1499(1980)].

[80] G. Veneziano. U(1) Without Instantons. *Nucl. Phys.*, B159:213–224, 1979.

[81] Jack Dragos, Thomas Luu, Andrea Shindler, Jordy de Vries, and Ahmed Yousif. Confirming the Existence of the strong CP Problem in Lattice QCD with the Gradient Flow. 2019.

[82] C. A. Baker, D. D. Doyle, P. Geltenbort, K. Green, M. G. D. van der Grinten, P. G. Harris, P. Iaydjiev, S. N. Ivanov, D. J. R. May, J. M. Pendlebury, J. D. Richardson, D. Shiers, and K. F. Smith. Improved experimental limit on the electric dipole moment of the neutron. *Phys. Rev. Lett.*, 97:131801, Sep 2006.

[83] Edward Witten. Instantons, the Quark Model, and the 1/n Expansion. *Nucl. Phys.*, B149:285–320, 1979.

[84] Edward Witten. Current Algebra Theorems for the U(1) Goldstone Boson. *Nucl. Phys.*, B156:269–283, 1979.

[85] M. Fukugita, Y. Kuramashi, M. Okawa, and A. Ukawa. Lattice QCD solution to the U(1) problem. *Phys. Rev.*, D51:3952–3954, 1995.

[86] L. Venkataraman and G. Kilcup. The eta-prime meson with staggered fermions. *Submitted to: Phys. Rev. D*, 1997.

[87] J. Smit and J. C. Vink. Neutral Pseudoscalar Masses in Lattice QCD. *Nucl. Phys.*, B284:234–252, 1987.

[88] S. Okubo. Phi meson and unitary symmetry model. *Phys. Lett.*, 5:165–168, 1963.

[89] Susumu Okubo. A Survey of Quark Line Rule. *Prog. Theor. Phys. Suppl.*, 63:1–48, 1978.

[90] Jugoro Iizuka. A Systematics and Phenomenology of Meson Family*. *Progress of Theoretical Physics Supplement*, 37-38:21–34, 03 1966.

[91] N. Eicker et al. Evaluating sea quark contributions to flavor singlet operators in lattice QCD. *Phys. Lett.*, B389:720–726, 1996.

[92] Walter Wilcox. Noise methods for flavor singlet quantities. In *Numerical challenges in lattice quantum chromodynamics. Proceedings, Joint Interdisciplinary Workshop, Wuppertal, Germany, August 22-24, 1999*, pages 127–141, 1999.

[93] Justin Foley, K. Jimmy Juge, Alan O'Cais, Mike Peardon, Sinead M. Ryan, and Jon-Ivar Skullerud. Practical all-to-all propagators for lattice QCD. *Comput. Phys. Commun.*, 172:145–162, 2005.

[94] P. A. Boyle, A. Juttner, C. Kelly, and R. D. Kenway. Use of stochastic sources for the lattice determination of light quark physics. *JHEP*, 08:086, 2008.

[95] Andreas Stathopoulos, Jesse Laeuchli, and Kostas Orginos. Hierarchical probing for estimating the trace of the matrix inverse on toroidal lattices. 2013.

[96] S. Aoki et al. An Estimate of the eta and eta-prime meson masses in N(f) = 2+1 lattice QCD. *PoS*, LAT2006:204, 2006.

[97] H. Goldstein, J. Poole, and J. Safko. *Classical Mechanics*. Addison Wesley, 3rd edition, 2002.

[98] M. Roughan. *Variational Methods & Optimal Control*. University of Adelaide, 2012.

[99] F. Larusson. *Topology and Anaysis*. University of Adelaide, 2015.

[100] F. Larusson. *Differential Geometry*. University of Adelaide, 2008.

[101] R. Feynman and A. Hibbs. *Quantum Mechanics and Path Integrals*. McGraw-Hill Inc., 1st edition, 1965.

APPENDIX A

Minimisation in Nature

The processes of change in Nature are observed to centre around paths that minimise or maximise certain quantities. Simple examples of this include a ball rolling down a hill. The process of change seems to be related to a minimisation of potential energy. The process of change has been more rigorously developed by the likes of d'Alembert and others to be the result of local forces, which can be captured in the framework of Lagrangian mechanics. As the processes of change in Nature seem to follow the predictions from this area of mechanics and are compliant with relativity, Lagrangian mechanics is useful in predicting the motions of the particles, which are the subject of this thesis.

In section A.1 I explain how these local forces can be understood in terms of Legrangian mechanics, through d'Alembert's principle. In section A.2 after an introduction to vector spaces I introduce the concept of a functional. In section A.3 I derive the Euler-Lagrange equations, which minimise the functional, and show how this relates to d'Alembert's principle. This allows us to understand the concept of the action of a system, S. In section A.4 I generalise d'Alembert's principle to special relativity. In section A.5 and A.6 I introduce the concept of the path integral and how it can be used to predict the states of quantum systems. Finally, in section A.7 I talk about identical particles and how these act as indistinguishable alternatives, much like the slits in the double slit experiment.

A.1 D'Alembert's Principle

The workings of nature are most apparent when we consider forces that result directly in the motion of particles. Certainly, Newton's theories of motion are correct in the regime he conceived them. Using their transparency, I will explain

their connection to Lagrangian mechanics and expand these classical ideas to a relativistic regime.

The following proof is derived from [97]. A virtual displacement is defined as a thought experiment that tests the effect of a change in the configuration of a system. Such a change is instantaneous. Consider a system in equilibrium. Each particle in this system has a zero net force applied to it, $\mathbf{F_i} = \mathbf{0}$ where the index denotes a particle label. By applying virtual displacements to the particles we find we produce no virtual work and so $\sum_i \mathbf{F_i} \cdot \delta\mathbf{r_i} = \mathbf{0}$. If we separate the forces into two types, those defined by constraints f_i and applied forces F_i^a we find, $\sum_i \mathbf{F_i^a} \cdot \delta\mathbf{r_i} + \sum_i \mathbf{f_i} \cdot \delta\mathbf{r_i} = \mathbf{0}$. If we restrict the allowable constraints to those that can do no work (those constraints that act perpendicular to allowed motion), then we have that the virtual work of applied forces vanish as well, $\sum_i \mathbf{F_i^a} \cdot \delta\mathbf{r_i} = \mathbf{0}$.

Consider the equation of motion, $\mathbf{F_i} = \dot{\mathbf{p}}_\mathbf{i}$. That is, a force applied to a particle over time generates a change in the momentum of the particle. Applying this to our previous analysis, we find

$$\sum_i (\mathbf{F_i^a} - \dot{\mathbf{p}}_\mathbf{i}) \cdot \delta\mathbf{r_i} = \mathbf{0} \tag{A.1}$$

which is called D'Alembert's principle. The principle is satisfied irrespective of the coordinate system. In fact, if we can show that the equations decouple, then the path of each particle can be solved separately. That is $(\mathbf{F_i^a} - \dot{\mathbf{p}}_\mathbf{i}) \cdot \delta\mathbf{r_i} = \mathbf{0}$ in some new coordinate system. If we let $\mathbf{r_i}(\mathbf{q_1}, \mathbf{q_2}...\mathbf{q_n}, t)$ then this change of variables can be achieved via the following transformation,

$$\delta\mathbf{r_i} = \sum_\mathbf{j} \frac{\partial \mathbf{r_i}}{\partial \mathbf{q_j}} \delta\mathbf{q_j} \tag{A.2}$$

$$\sum_\mathbf{i} \mathbf{F_i} \cdot \delta\mathbf{r_i} = \sum_\mathbf{i,j} \mathbf{F_i} \cdot \frac{\partial \mathbf{r_i}}{\partial \mathbf{q_j}} \delta\mathbf{q_j} = \sum_\mathbf{j} \mathbf{Q_j} \cdot \delta\mathbf{q_j}. \tag{A.3}$$

Where Q_j are the components of the generalised force. If we assume Newton's second law then,

$$\sum_\mathbf{i} \dot{\mathbf{p}}_\mathbf{i} \cdot \delta\mathbf{r_i} = \sum_i m_i \ddot{\mathbf{r}}_\mathbf{i} \cdot \delta\mathbf{r_i} \tag{A.4}$$

$$= \sum_{i,j} m_i \ddot{r}_i \cdot \frac{\partial r_i}{\partial q_j} \delta q_j \tag{A.5}$$

$$= \sum_{i,j} \frac{d}{dt}\left(m_i \dot{r}_i \cdot \frac{\partial r_i}{\partial q_j} \right) \delta q_j - m_i \dot{r}_i \cdot \frac{d}{dt}\left(\frac{\partial r_i}{\partial q_j} \right) \delta q_j \tag{A.6}$$

$$= \sum_{i,j} \frac{d}{dt}\left(m_i v_i \cdot \frac{\partial r_i}{\partial q_j} \right) \delta q_j - m_i v_i \cdot \frac{\partial v_i}{\partial q_j} \delta q_j \tag{A.7}$$

Where $v_i = \frac{dr_i}{dt} = \sum_j \frac{\partial r_i}{\partial q_j}\frac{\partial q_j}{\partial t} + \frac{\partial r_i}{\partial t}$. Using this definition we also find $\frac{\partial v_i}{\partial \dot{q}_k} = \frac{\partial r_i}{\partial q_k}$. Hence,

$$\sum_{\mathbf{i}} \dot{\mathbf{p}}_\mathbf{i} \cdot \delta \mathbf{r}_\mathbf{i} = \sum_{i,j} \frac{d}{dt}\left(m_i v_i \cdot \frac{\partial v_i}{\partial \dot{q}_j}\right)\delta q_j - m_i v_i \cdot \frac{\partial v_i}{\partial q_j}\delta q_j \tag{A.8}$$

$$= \sum_{i,j} \frac{d}{dt}\left[\frac{\partial}{\partial \dot{q}_j}\left(\frac{1}{2}m_i v_i^2\right)\right]\delta q_j - \frac{\partial}{\partial q_j}\left(\frac{1}{2}m_i v_i^2\right)\delta q_j \tag{A.9}$$

Substituting Eq. (A.3) and Eq. (A.9) into d'Alembert's principle Eq. (A.1) and letting $T = \sum_i \frac{1}{2}m v_i^2$,

$$\sum_j \left\{\left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j}\right] - Q_j\right\}\delta q_j = 0. \tag{A.10}$$

If the constraints are holonomic, meaning they can be put in the form $f(\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}, ..., t) = \mathbf{0}$, then the equations of constraint define a solution space that is a submanifold of the original solution space. If we choose a basis that spans this submanifold and use these to parameterise our problem, then the only forces that survive are the projections of the applied forces onto the subspace, so $\mathbf{Q_i^a} - \dot{\mathbf{p}}_\mathbf{i} = \mathbf{0}$, by definition. As such we can find a parameterisation in which the equations decouple. So the equation simplifies to,

$$\left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j}\right] - Q_j = 0. \tag{A.11}$$

If the applied forces are derivable from a scalar potential function V, then $\mathbf{F_i} = -\nabla V_i$, so $Q_j = \sum_j \mathbf{F_i} \cdot \frac{\partial \mathbf{r_i}}{\partial \mathbf{q_j}} = -\sum_i \nabla_i V \cdot \frac{\partial r_i}{\partial q_j} = \sum_j -\frac{\partial V_i}{\partial q_j}$. So,

$$\left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial (T-V)}{\partial q_j}\right] = 0. \tag{A.12}$$

If the potential is not dependent on the first time derivative of the generalised coordinates $V(q_1, ..., q_n, t)$ then

$$\left[\frac{d}{dt}\left(\frac{\partial (T-V)}{\partial \dot{q}_j}\right) - \frac{\partial (T-V)}{\partial q_j}\right] = 0 \tag{A.13}$$

which can be written in terms of the Lagrangian $L = T - V$,

$$\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j}\right] = 0, \tag{A.14}$$

this expression is referred to as Lagrange's equations.

If not all the forces acting on the system are derivable from a potential, Lagrange's equations can always be written,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = Q_j \tag{A.15}$$

where L contains the potential of the conservative forces and Q_j represents the forces not arising from a potential.

A.2 Vector Spaces and Functionals

A vector space is a collection of objects (vectors) along with two operators, addition and scalar multiplication. A vector space is closed under both of these operations [98]. Let $\mathbb{R}^n$ denote the set of ordered n-tuples $(x_1, ..., x_n)$ of real numbers. This set is a vector space over the field $\mathbb{R}$ when equipped with operators for addition and scalar multiplication. The set of continuous functions $\mathcal{C}([a, b])$ on interval $[a, b] \to \mathbb{R}$ is an example of an infinite dimensional vector space [99].

Further topology can be induced on this space by defining a metric, a function that measures the distance between two vectors, and hence defines neighbourhoods [99]. A metric space is a set X with a function $d : X \times X \to [0, \infty)$ such that,

$$\begin{aligned}
&d(x, y) = 0 \text{ if and only if } x = y,\\
&d(x, y) = d(y, x)\ \forall x, y \in X,\\
&d(x, z) \leq d(x, y) + d(y, z)\ \forall x, y, z \in X\ (\textit{triangle inequality}).
\end{aligned}$$

A metric can automatically define a norm for the space. For finite dimensional vector spaces, the choice of metric or norm does not effect the topology because each norm or metric can be related to the others via scalar multiplication [98]. That is, for any two norms $||\cdot||_a, ||\cdot||_b$ where $||x||_a \leq ||x||_b$, $\exists n$ such that $||x||_a \leq ||x||_b \leq n||x||_a$. This is not necessarily the case for infinite dimensional vector spaces.

An inner product is a function that maps two elements in a vector space to the field on that space [99]. An inner product space is a set X with a map $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$ called an inner product, such that,

$$\begin{aligned}
\langle x, x \rangle &> 0\ \forall x \in X,\ x \neq 0,\\
\langle x, y \rangle &= \langle y, x \rangle\ \ \forall x, y \in X,\\
\langle ax, y \rangle &= a\langle x, y \rangle\ \ \forall x, y \in X \text{ and } a \in \mathbb{R},\\
\langle x + y, z \rangle &= \langle x, z \rangle + \langle y, z \rangle\ \ \forall x, y, z \in X.
\end{aligned}$$

An inner product induces a norm $||\cdot||$ on X and hence a metric d and a topology induced by d [99].

A functional maps elements of a space X to the field defined on that space, hence for some vector space X, $F : X \to \mathbb{R}$ [98]. A simple example is the *max* function on a vector or function.

Often we wish to find extrema of a functional. It is easy to find the extrema for simple functionals. However it is more difficult to find extrema for integral functionals, consider for example the functionals

$$F\{y\} = \int_a^b y(x)dx$$
$$F\{y\} = \int_a^b f(x)y(x)dx$$
$$F\{y\} = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}dx.$$

we will tackle this problem in the next section.

A.3 Euler-Lagrange Equations

Consider the fixed end point problem, where one tries to find a path $y(x)$, $x \in [x_0, x_1]$ that minimises a functional F between two end points (x_0, y_0) and (x_1, y_1). Define the functional $F : X \to \mathbb{R}$, $F\{y\} = \int_{x_0}^{x_1} f(x, y, \frac{dy}{dx})$, where X is the set of C^2 (at least two derivatives exist) continuous functions $f : [x_0, x_1] \times [y_0, y_1] \times [\dot{y}_0, \dot{y}_1] \to \mathbb{R}$ on variables $x, y, \frac{dy}{dx}$ [98]. The implicit function theorem states [100]: suppose f is a C^k map, $1 \leq k \leq \infty$ from a neighbourhood of (x_0, y_0) in $\mathbb{R}^n \times \mathbb{R}^m$ into $\mathbb{R}^m$ such that the derivative of the map $y \to f(x_0, y)$ is invertible at y_0. Write $c = f(x_0, y_0)$, then there are open neighbourhoods U of x_0 and V of y_0 and a C^k map $g : U \to V$ such that for every $(x, y) \in U \times V$, we have $f(x, y) = c$ if and only if $y = g(x)$.

As such, in a neighbourhood of (x_0, y_0), y can be parameterised in terms of x provided f is suitably well behaved. So the space of possible curves is $S = \{y \in C^2[x_0, x_1] \mid y(x_0) = y_0, y(x_1) = y_1\}$ [98]. Suppose we choose a path in S and perturb it by a function $\eta \in \mathcal{H}$ where $\mathcal{H} = \{\eta \in C^2[x_0, x_1] \mid \eta(x_0) = 0,\ \eta(x_1) = 0\}$. For some $\epsilon > 0$ we find the new path to be $\hat{y}(x) = y(x) + \epsilon\eta(x)$. Applying Taylor's theorem,

$$f(x, \hat{y}, \dot{\hat{y}}) = f(x, y, \dot{y}) + \epsilon\left[\eta\frac{\partial f}{\partial y} + \dot{\eta}\frac{\partial f}{\partial \dot{y}}\right] + \mathcal{O}(\epsilon^2). \tag{A.16}$$

So the functional value changes as,

$$F\{\hat{y}\} - F\{y\} = \epsilon \int_{x_0}^{x_1} \left[\eta \frac{\partial f}{\partial y} + \dot{\eta}\frac{\partial f}{\partial \dot{y}}\right] dx + \mathcal{O}(\epsilon^2) \tag{A.17}$$

We can define a quantity called the first variation by taking the limit as $\epsilon \to 0$,

$$\delta F(\eta, y) = \lim_{\epsilon \to 0} \frac{F\{\hat{y}\} - F\{y\}}{\epsilon} = \int_{x_0}^{x_1} \left[\eta \frac{\partial f}{\partial y} + \dot{\eta}\frac{\partial f}{\partial \dot{y}}\right] dx \tag{A.18}$$

At local extrema, $\delta F(\eta, y) = 0$, $\forall \eta \in \mathcal{H}$. Integrating by parts on the last term in Eq. (A.18),

$$\delta F(\eta, y) = \left[\eta \frac{\partial f}{\partial \dot{y}}\right]_{x_0}^{x_1} + \int_{x_0}^{x_1} \eta \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right)\right] dx \tag{A.19}$$

$$= \int_{x_0}^{x_1} \eta \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right)\right] dx \tag{A.20}$$

The last line is achieved by applying the boundary conditions on η, $\eta(x_0) = \eta(x_1) = 0$. As η can take any form we wish, and at extrema $\delta F = 0$, then at extrema,

$$\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right)\right] = 0 \tag{A.21}$$

which are called the Euler-Lagrange equations [98].

Notice it is in the same form derived for Lagrange's equations Eq. (A.14), which started from d'Alembert's principle. Using the result of this section we can now define a quantity called the action S for systems studied in section A.1. The action of a system is a functional given by, $S = \int_{t_0}^{t_1} L dt$, where L is the Lagrangian of the system and t is time. Minimisation of this functional allows us to determine the path a body will take between two points t_0 and t_1 in parameter space.

The minimisation occurs for a fixed end point problem when the Euler-Lagrange equations are satisfied. However, the Euler-Lagrange equations can even allow us to predict one or both of the end points. In the above procedure we set $\left[\eta \frac{\partial f}{\partial \dot{y}}\right]_{x_0}^{x_1} = 0$ because of the requirements on the end points. However, when the end points are not fixed but we have an appropriate function f which satisfies $\frac{\partial f}{\partial \dot{y}} = 0$, then the Euler-Lagrange equation can be applied. In the case that one end point is specified, the solution is uniquely determined, otherwise there may be more than one solution.

A.4 Special Relativity Generalisation

In proving d'Alembert's principle we assumed Newtonian mechanics, $F = m\ddot{x}$ or $p = m\dot{x}$, to be the correct theory describing motion. Though at low relative speeds

Newtonian mechanics is very accurate, if we wish to correctly describe Nature at high relative speeds we will need to apply special relativity, and this modifies the Lagrangian L and hence the action S.

Working with the relativistic definition of momentum $p = \gamma mv$ where $\gamma = \frac{1}{\sqrt{1-v^2/c^2}}$. The special relativistic pure force law is,

$$\begin{aligned}
\frac{dp}{dt} &= \frac{d\gamma}{dt}mv + \gamma m\frac{dv}{dt} && \text{(A.22)}\\
&= \frac{d\gamma}{dv}\frac{dv}{dt}mv + \gamma m\frac{dv}{dt} && \text{(A.23)}\\
&= \frac{\gamma^3 v}{c^2}mv\frac{dv}{dt} + \gamma m\frac{dv}{dt} && \text{(A.24)}\\
&= \left(\frac{\gamma^2 v^2 + c^2}{c^2}\right)\gamma m\frac{dv}{dt} && \text{(A.25)}\\
&= \frac{v^2 + c^2}{c(1 + v^2/c^2)}\gamma m\frac{dv}{dt} && \text{(A.26)}\\
&= \gamma m\frac{dv}{dt} && \text{(A.27)}
\end{aligned}$$

Continuing from section A.1 but replacing from Eq. (A.4) till Eq. (A.10) with the following will produce a Lagrangian which incorporates the relativistic law Eq. (A.27),

$$\begin{aligned}
\sum_{\mathbf{i}} \dot{\mathbf{p}}_\mathbf{i} \cdot \delta\mathbf{r}_\mathbf{i} &= \sum_i \gamma m_i \ddot{\mathbf{r}}_\mathbf{i} \cdot \delta\mathbf{r}_\mathbf{i} && \text{(A.28)}\\
&= \sum_{i,j} \gamma m_i \ddot{r}_i \cdot \frac{\partial r_i}{\partial q_j}\delta q_j && \text{(A.29)}\\
&= \sum_{i,j} \frac{d}{dt}\left(\gamma m_i v_i \cdot \frac{\partial r_i}{\partial q_j}\right)\delta q_j - \gamma m_i v_i \cdot \frac{d}{dt}\left(\frac{\partial r_i}{\partial q_j}\right)\delta q_j && \text{(A.30)}\\
&= \sum_{i,j} \frac{d}{dt}\left(\gamma m_i v_i \cdot \frac{\partial v_i}{\partial \dot{q}_j}\right)\delta q_j - \gamma m_i v_i \cdot \frac{\partial v_i}{\partial q_j}\delta q_j && \text{(A.31)}
\end{aligned}$$

Where again we have used the fact that $v_i = \frac{dr_i}{dt} = \sum_j \frac{\partial r_i}{\partial q_j}\frac{\partial q_j}{\partial t} + \frac{\partial r_i}{\partial t}$, and hence $\frac{\partial v_i}{\partial \dot{q}_k} = \frac{\partial r_i}{\partial q_k}$. We want to turn this into an expression for relativistic kinetic energy. To that end notice that,

$$\begin{aligned}
\int \gamma m_i v_i \frac{\partial v_i}{\partial \dot{q}_j} d\dot{q}_j &= \int \gamma m_i v_i dv_i && \text{(A.32)}\\
&= mc^2\sqrt{1 - v_i^2/c^2} && \text{(A.33)}
\end{aligned}$$

and so applying this idea to both terms,

$$\sum_{\mathbf{i}} \dot{\mathbf{p}}_\mathbf{i} \cdot \delta \mathbf{r}_\mathbf{i} = \sum_{i,j} \frac{d}{dt}\left(\gamma m_i v_i \cdot \frac{\partial v_i}{\partial \dot{q}_j}\right)\delta q_j - \gamma m_i v_i \cdot \frac{\partial v_i}{\partial q_j}\delta q_j \tag{A.34}$$

$$= \sum_{i,j} \frac{d}{dt}\left[\frac{\partial}{\partial \dot{q}_j}\left(mc^2\sqrt{1 - v_i^2/c^2}\right)\right]\delta q_j - \frac{\partial}{\partial q_j}\left(mc^2\sqrt{1 - v_i^2/c^2}\right)\delta q_j \tag{A.35}$$

$$= \sum_{i,j} \frac{d}{dt}\left(\frac{\partial T_i}{\partial \dot{q}_j}\right)\delta q_j - \frac{\partial T_i}{\partial q_j}\delta q_j \tag{A.36}$$

By following the procedure in section A.1, we obtain the same Euler-Lagrange equation, except with a different definition of kinetic energy. The remainder of the proof is given in section A.1; by substituting Eq. (A.3) and Eq. (A.36) into Eq. (A.1) (d'Alembert's principle) produces Eq. (A.10), but with $T = mc^2\sqrt{1 - v_i^2/c^2}$. The proof remains the same from that point onward.

We have only derived this in a particular frame, but if the Lagrangian is a Lorentz scalar (meaning it is invariant under Lorentz transformations) then the Lagrangian is valid in any frame. We will find in the proceeding sections that this is the case for simple particle systems.

A.5 Quantum Mechanics and the Path Integral

How does a particle get from a point $q(t_1)$ to point $q(t_2)$? If we consider a classical particle, the path that the particle follows is the one that minimises the action. In the quantum case, we do not know with certainty where the particle is, and we do not know with certainty its momentum. We only know the probability that a particle will be in a certain place or have a certain momentum, if we know its wave function. In fact, until we probe at the particle, it does not even exist at a single point. So a unique path does not exist. You might consider that a particle fluctuates around a central minimised path, or alternatively that all paths are possible, though not equally likely.

To explain this further, consider the thought experiment given by Feynman in [101]. Referring to Fig. A.1, at point A we have a source of electrons, at point B we have a screen with two holes in it and at point C we have a screen that detects electrons and tells us when and where they landed. The source fires single electrons. By running the experiment for a time, we can build a probability distribution of where the electrons land.

In the case where one path is blocked and the other is open we achieve Fig. A.2a if the top hole is blocked, or Fig. A.2b if the bottom hole is blocked. These are

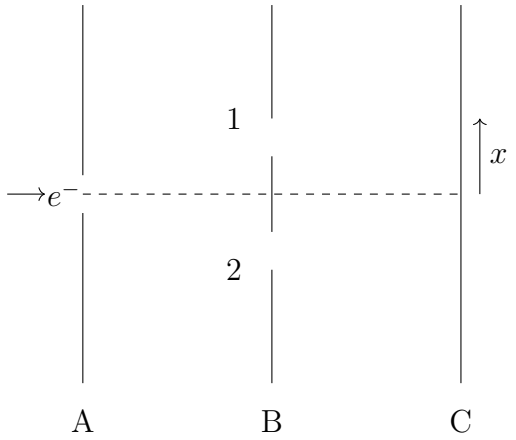


Figure A.1: Double slit thought experiment using an electron [101].

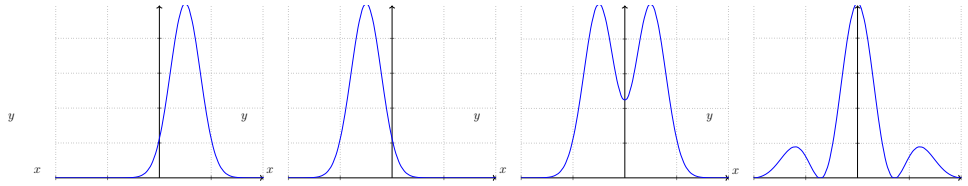


Figure A.2: Plots (a),(b) and (c) show the sort of distribution we would expect if electrons experienced diffraction or acted as rays, but the slits did not interfere with each other [101]. Plot (d) is a pictorial representation of what the real distribution looks like.

both probability distributions which agree with our classical intuition. However if both holes are kept open we achieve the interference pattern show in Fig. A.2d. Notice this is not just the sum of the probabilities of the first two graphs as we would expect classically. Additionally, suppose I were to put a light source on the right hand side of the screen B, firing along screen B. If these photons allow enough resolution to determine if an electron came out of one particular hole or the other, then (whether we actually measure this or not) the distribution at C will be what we would expect classically for a double slit experiment. That is, simply the sum of the probabilities of the two single hole experiments. In this way, if we do not try to resolve which hole the electron went through, it goes through both holes.

In quantum systems it is not generally true that $P = P_1 + P_2$, for probabilities P, because these two choices are not necessarily independent. The above behaviour is reminiscent of a double slit experiment with light, which uses the concept of a wave to make sense of the experimental results. In a similar way, the above

situation can be described by complex valued probability amplitudes ϕ, where $P_1 = |\phi_1|^2$, so that $P = |\phi_1 + \phi_2|^2$. More generally, consider a countably infinite number of slits and partitions, the probability amplitude is $K(b,a) = \sum_{i=0}^{N} \phi_i$, where ϕ_i is the probability amplitude for the i_{th} distinct path to get from a to b. For an uncountable infinite set of paths, (that is a continuously deformable path) we have

$$K(b,a) = \sum_{\text{all paths}} \phi[x(t)] \tag{A.37}$$

where ϕ is a functional of the set of paths $x(t)$. In such a scenario the screens and holes are no longer there, however the particle travels along each conceivable path [25].

Feynman postulated that the probability amplitude should be $\phi[x(t)] = Ae^{iS[x(t)]/\hbar}$, for some constant A where S is the classical action. In the limit that δS is much larger than $\hbar$, we recover the classical path. Furthermore, no particular path is preferred, as $|e^{iS/\hbar}| = 1$. Rather the superposition of amplitudes decides the probability that a particle be at a certain place; this would suggest that all paths are taken. These motivations do not prove that this is the correct form of the amplitude, however the theory gives accurate predictions in experiments, and so the choice has been validated.

The path integral is normally written in the following notation,

$$K(b,a) = \int_a^b \mathcal{D}x(t)\ e^{iS[x(t)]/\hbar} \tag{A.38}$$

This amplitude can be split into parts by recognising that $S(b,a) = S(b,c)+S(c,a)$ for some intermediate point c, provided we integrate over this position,

$$K(b,a) = \int_{x_c}\int_a^b \mathcal{D}x(t)dx_c\ e^{iS(b,c)/\hbar}e^{iS(c,a)/\hbar} \tag{A.39}$$

$$= \int_{x_c} dx_c\ K(b,c)K(c,a) \tag{A.40}$$

We can continue this process of subdividing the path integral into smaller time intervals. Suppose we subdivide it into time intervals of length ϵ to achieve [101];

$$K(b,a) = \int_{x_{N-1}} ... \int_{x_1} dx_1dx_2...dx_{N-1}\ K(b,x_{N-1})K(x_{N-1},x_{N-2})...K(x_1,a). \tag{A.41}$$

In the limit that $\epsilon \to 0$ the path integral becomes only the straight paths and the action is a linear function of the Lagrangian,

$$K(x_{i+1},x_i) = e^{i\epsilon L_i(x_i,x_{i+1},\epsilon)/\hbar}. \tag{A.42}$$

So the path integral can be written as,

$$K(b,a) = \int_a^b \mathcal{D}x(t)\ e^{iS(b,a)/\hbar} \tag{A.43}$$

$$= \lim_{\epsilon\to 0} \int_{x_{N-1}} ... \int_{x_1} dx_1 dx_2 ... dx_{N-1} \prod_{i=0}^{N-1} e^{i\epsilon L_i(x_i, x_{i+1}, \epsilon)/\hbar}. \tag{A.44}$$

A.6 Dirac's Formulation

To further motivate the use of the action as the phase factor in the above amplitudes, I will show that the path integral formalism can be derived directly from non-relativistic quantum theory starting from $\left\langle q_f \left| e^{-i\int Hdt} \right| q_i \right\rangle$. Furthermore, I will show that the path integral is not just applicable to determining particle positions, but can be broadened to generic system states. We will follow the derivation given in [25].

The notation $\left\langle q_f \left| e^{-i\int Hdt} \right| q_i \right\rangle$ denotes the probability of the particle position ending at q_f given an initial position q_i after time t has elapsed. Here $e^{i\int \hat{H}dt}$ is the time evolution operator, $\hat{H}$ is the Hamiltonian operator and t is time. To begin we break the integral into small time segments, say N parts,

$$\left\langle q_f \left| e^{-i\int Hdt} \right| q_i \right\rangle = \left(\prod_{j=1}^{N-1} \int dq_j \right) \left\langle q_f \left| e^{-iH\delta t} \right| q_{N-1} \right\rangle \left\langle q_{N-1} \left| e^{-iH\delta t} \right| q_{N-2} \right\rangle ... \left\langle q_1 \left| e^{-iH\delta t} \right| q_i \right\rangle \tag{A.45}$$

Consider one of these segments and let $\hat{H} = \hat{p}^2/2m + V(\hat{q})$, where $\hat{p}$ is the momentum operator and $\hat{q}$ is the position operator,

$$\left\langle q_{j+1} \left| e^{-iH\delta t} \right| q_j \right\rangle = \int \frac{dp}{2\pi} \left\langle q_{j+1} \left| e^{-iH\delta t} \right| p \right\rangle \left\langle p \,|q_j\right\rangle \tag{A.46}$$

$$= \int \frac{dp}{2\pi} \left\langle q_{j+1} \left| e^{-iH\delta t} \right| p \right\rangle \left\langle p \,|q_j\right\rangle \tag{A.47}$$

$$= \int \frac{dp}{2\pi} e^{-i(p^2/2m + V(q_{j+1}))\delta t} e^{ip(q_{j+1}-q_j)} \tag{A.48}$$

$$= \left(\frac{-im}{2\pi\delta t}\right)^{\frac{1}{2}} e^{im(q_{j+1}-q_j)^2/(2\delta t) - iV(q_{j+1})} \tag{A.49}$$

$$= \left(\frac{-im}{2\pi\delta t}\right)^{\frac{1}{2}} e^{im\delta t[(q_{j+1}-q_j)/\delta t]^2/2 - i\delta t V(q_{j+1})}. \tag{A.50}$$

Substituting back into the main expression,

$$\left\langle q_f \left| e^{-i\int H dt} \right| q_i \right\rangle = \left(\prod_{j=1}^{N-1} \int dq_j \right) \left(\frac{-im}{2\pi\delta t} \right)^{\frac{N}{2}} e^{\sum_{j=0}^{N-1} im\delta t[(q_{j+1}-q_j)/\delta t]^2/2 - i\delta t V(q_{j+1})} \tag{A.51}$$

Going to the continuum limit as $\delta t \to 0$, $\left[(q_{j+1}-q_j)/\delta t\right]^2 \to \dot{q}^2$, $\sum_{j=0}^{N-1} \to \int_0^T dt$ and $\lim_{N\to\infty} \left(\frac{-im}{2\pi\delta t}\right)^{\frac{N}{2}} \left(\prod_{j=1}^{N-1} \int dq_j\right) \to \int Dq(t)$. So the quantity reduces to,

$$\left\langle q_f \left| e^{-i\int H dt} \right| q_i \right\rangle = \int Dq(t) e^{i\int_0^T dt L} \tag{A.52}$$

where $L = \frac{1}{2}mv^2 - V(q)$, which is the Lagrangian for the non-relativistic system. Generalising to some initial system state I and final state F,

$$\left\langle F \left| e^{i\int H dt} \right| I \right\rangle = \int\int dq_1 dq_2 \left\langle F \right| q_2 \rangle \left\langle q_2 \left| e^{i\int H dt} \right| q_1 \right\rangle \langle q_1 \left| I \right\rangle \tag{A.53}$$

$$= \int\int dq_1 dq_2 \Psi_F^\dagger(q_2) \left\langle q_2 \left| e^{i\int H dt} \right| q_1 \right\rangle \Psi_I(q_1). \tag{A.54}$$

If the initial and final states are the ground state $\left\langle 0 \left| e^{i\int H dt} \right| 0 \right\rangle$, it is conventional to name the quantity Z [25]. This is called the partition function.

A.7 Identical Particles and Spin Statistics

The idea that two alternatives (paths in our previous case) interfere or are in some way identical (we could not tell which slit the particle went through) is not isolated to paths. The same is true for fundamental particles and is reflected in how they interact with one another.

Consider a thought experiment given by Feynman [101]. Referring to Fig. A.3, in the centre of momentum frame, at points A and B, a particle enters the experimental set up. We have detectors at 90° to the beam axis at C and D.

Consider first that the particles are fundamentally different, say a is an α particle and b is some other nucleus. Then the probability of 90° scattering is $|\phi_{AB}(1,2)|^2 + |\phi_{AB}(2,1)|^2 = 2p$, where the two probabilities are equal so we have simplified the notation to simply p. In this case, after the experiment we could look at the particles and determine which particle went where, and hence by not influencing directly the experiment conducted we have determined which particle was which. As such the classical prediction for the probability is upheld.

In the case where both particles were α particles, because they are indistinguishable in every way much like in the slits in the double slit experiment, the

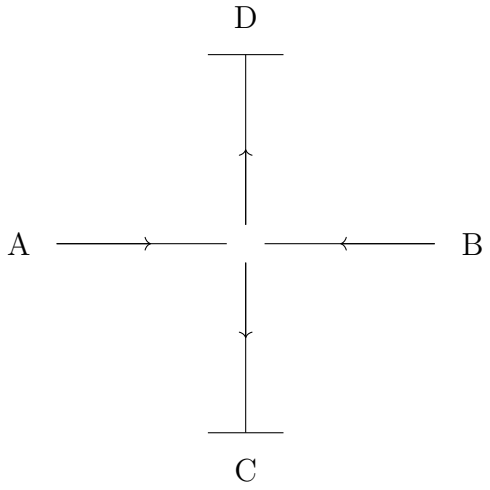


Figure A.3: Two particle scattering with detectors at 90° to the beam axis [101].

probability of 90° scattering is $|\phi_{AB}(1,2) + \phi_{AB}(2,1)|^2 = 4p$. If it was possible to determine which particle was which after the experiment, like in the first case presented, then we would certainly see the same probability as in case 1. By virtue of the two particles producing the above statistics, we know that they are indistinguishable (or interfering) alternatives.

Suppose both particles are electrons with identical spin. In interchanging the electrons there is a phase shift of 180° and so if the electrons are identical, the probability of 90° scattering is $|\phi_{AB}(1,2) + \phi_{AB}(2,1)|^2 = 0$, $\phi_{AB}(1,2) = -\phi_{AB}(2,1)$. Particles which conform to this behaviour are called fermions, and obey Fermi statistics. Particles for which interchange does not alter the phase are called bosons and are said to obey Bose statistics [101], as the above α particles do.

APPENDIX B

Field theory

The concept that the most fundamental particles in our universe are in every way identical leads to odd conclusions. It suggests that an electron at one space-time point is indistinguishable from every other electron (apart from its location and momentum). This idea is reconciled by the concept of a field. Perhaps there is a field on which the electron is an excitation, and this field permeates the entirety of space-time.

In this chapter we develop the rules by which such a concept must abide. In section B.1 I introduce the concept of plane waves and show that the wave vector N is a Lorentz invariant. I then explain that particles satisfy the de Broglie's equation, hence act like waves. In section B.2 I explain how these matter waves satisfy relativistic energy conservation when free of potentials, and write down the Lagrangian and path integral which governs their motion. I also include the more general case of a non-free matter waves. In section B.3 I solve the path integral equation for the free case. In section B.4 I use the free field solution to produce a perturbative expansion to obtain approximate solutions to the non-free path integral equation.

B.1 Plane Waves

Consider a series of plane disturbances a wavelength λ apart, progressing in a unit direction $\vec{n} = (l, m, n)$ at speed u relative to an inertial frame S. The equation of one such plane in S, where the origin in S is at point t_0, x_0, y_0, z_0, is given by

$$l(x - x_0) + m(y - y_0) + n(z - z_0) = u(t - t_0) \tag{B.1}$$

where the plane coincides with the origin at $t = t_0$. Any of the disturbance planes can then be described by this same equation with an additional $L\lambda$, $L \in \mathbb{Z}$ added to the right-hand side of Eq. (B.1). Hence, define $N^\mu \Delta s_\mu = L$, where $N = \frac{1}{\lambda}(u/c, \vec{n}) = \nu\left(\frac{1}{c}, \frac{\vec{n}}{u}\right)$. Though inertial observers may not agree on where the planes reside or even their spacing, they can agree on whether a plane is disturbed or not; that a set of planes does indeed form a set. They can also tell to which plane another observer is referring. In other words, inertial observes can tell if planes are in phase. This means that L is Lorentz invariant, and is often called the phase of a wave. As Δs is a Lorentz vector, then we have just shown that N is as well. We call this vector the wave vector.

At the turn of the 20th century, seeking a way to fix his theory on the black body spectrum [18], Planck suggested that radiation (a wave) was emitted in quanta, with energy $E = h\nu$. Shortly after, Einstein suggested that radiation was emitted, travelled and absorbed as quanta, as this explained the photoelectric effect. If we look at $p = (E/c, \vec{p})$ and $N = \nu\left(\frac{1}{c}, \frac{\vec{n}}{u}\right)$, we see that the spatial vectors point in the same direction (the direction of travel), and if $E = h\nu$ then perhaps $p = hN$. De'Broglie later suggested that this relation can be applied to particles, and this was confirmed when electron diffraction was observed. It is now called de Broglie's equation.

B.2 The Klein-Gordon Equation

For any matter wave, the wave function $\phi(t, \vec{x})$ can be decomposed into plane waves via Fourier decomposition,

$$\phi(t, \vec{x}) = \int d^4k \; a(k) e^{-ik^\mu x_\mu} \tag{B.2}$$

where we define $p = \hbar k$. By applying the derivative ∂^μ to this decomposition we can see that the momentum operator must be $p^\mu = i\hbar\partial^\mu$. If the matter wave is free of potentials and complies with special relativity, the wave function must satisfy the dynamic equation $p^2 = m_0^2 c^2$, where m_0 is the rest mass. This implies $-\hbar^2 \partial^\mu \partial_\mu \phi = m_0^2 c^2 \phi$, which is called the Klein-Gordon equation.

In the classical limit, where the Euler-Lagrange equations specify the motion of the system, we will require that the dynamics of any free field satisfy the Klein-Gordon equation, as this is the equation that fits with the classical world around us. One of the simplest Lagrangians which satisfy this is $L = \frac{1}{2}\hbar^2(\partial\phi)^2 - \frac{1}{2}m_0^2 c^2 \phi^2$, and the simplest case of this is where ϕ is a single component field; that is, one with a single value at each point in space-time. This is known as the free scalar field Lagrangian.

The Lagrangian can be generalised to include potentials $V(\phi)$ (which includes derivatives of ϕ), $L = \frac{1}{2}\hbar^2(\partial\phi)^2 - \frac{1}{2}m_0^2c^2\phi^2 - V(\phi)$ which will not satisfy the Klein-Gordon equation. Quite generally a single component field would satisfy the partition function,

$$Z = \int D\phi e^{\frac{i}{\hbar}\int d^4x \frac{1}{2}\hbar^2(\partial\phi)^2 - V(\phi)}. \tag{B.3}$$

Terms in the potential that are of the form $J(x)\phi(x)$ act like sources and sinks – we can see this by applying the Euler-Lagrange equations. The function $J(x)$ tells us where the sources and sinks are [25]. Likewise, terms that are of the form $M\phi(x)^2$ act like mass terms. Higher order $K_n\phi(x)^n$ mix the local free fields, and so act like interactions. Typically, the latter two coefficients do not depend on position.

B.3 Solving the Free Theory

Exact solutions to Eq. (B.3) have not been found, except in the special case where $L = \frac{1}{2}\hbar^2(\partial\phi)^2 - \frac{1}{2}m^2c^2\phi^2 + J\phi$. From this point on we will be working in Gaussian units $c = \hbar = 1$ and the rest mass m_0 will be written as just m. The partition function Z is then (integrating by parts),

$$Z = \int D\phi e^{i\int d^4x - \frac{1}{2}\phi(\partial^2 + m^2)\phi + J\phi}. \tag{B.4}$$

If we change to the discrete version of the path integral with lattice spacing a, the integrals become sums which reduces all the multiplications to dot products. For example, $\int d^4x J(x)\phi(x) \to a^4\sum_i J_i\phi_i$ [25]. Partial derivatives are discretised, for example in one direction $\partial\phi(ka) \to (1/a)(\phi_{k+1} - \phi_k) = M_{kj}\phi_j$ and so $\int d^4x\phi i(\partial^2 + m^2)\phi \to a^4\phi_k i(M_{ki}M_{ij} + m^2I_{kj})\phi_j = \phi^t \cdot A \cdot \phi$, for some four-dimensional double derivative M^2. The exact discretised form of A does not really concern us at this point because we will revert to the continuum very shortly.

We can diagonalise the matrix A by changing variable ϕ using a linear transformation; by changing variable again we can complete the square. This allows us to determine the integral using the identity $\int_{-\infty}^{\infty} dx e^{\frac{1}{2}iax^2 + iJx} = \left(\frac{2\pi i}{a}\right)^{1/2} e^{-i\frac{J^2}{2a}}$. Details can be found in [25], the solution is given below,

$$Z = \int ... \int d\phi_1 ... d\phi_N e^{-\frac{1}{2}\phi\cdot A\cdot\phi + iJ\cdot\phi} \tag{B.5}$$

$$= \left(\frac{(-2\pi)^N}{\det A}\right)^{\frac{1}{2}} e^{(1/2)J\cdot A^{-1}\cdot J}. \tag{B.6}$$

Shifting back to the continuum,

$$Z = \mathcal{C} e^{1/2 \int d^4x \int d^4y J(x) D(x-y) J(y)}, \tag{B.7}$$

$$= Z(J=0) e^{W(J)} \tag{B.8}$$

where $\mathcal{C}$ is an unknown constant that is fixed by the sourceless vacuum, $Z(J=0)$ and,

$$W(J) = \frac{1}{2} \int \int d^4x d^4y J(x) D(x-y) J(y). \tag{B.9}$$

Since $D(x-y)$ is the inverse of A, we need to find the solution to $-i(\partial^2 + m^2) D(x-y) = \delta^4(x-y)$. The function D is called the propagator. Using the fact that $\delta^4(x-y) = \int \frac{d^4k}{(2\pi)^4} e^{ik(x-y)}$, the solution to D is,

$$D(x-y) = \int \frac{d^4k}{(2\pi)^4} \frac{i e^{ik(x-y)}}{k^2 - m^2 + i\epsilon}, \tag{B.10}$$

which can be checked by applying the operator $i(\partial^2 + m^2)$.

B.4 Perturbative Field Theory

Though we can not find an exact solution to

$$Z = \int D\phi e^{i \int d^4x \frac{1}{2}(\partial\phi)^2 - V(\phi)} \tag{B.11}$$

we can find solutions to equations of the form,

$$Z = \int D\phi e^{i \int d^4x - \frac{1}{2}\phi(\partial^2 + m^2)\phi + J\phi} \phi^n. \tag{B.12}$$

We can approximate the full solution by truncated expansions of the exponential. Consider for instance,

$$Z = \int D\phi e^{i \int d^4x - \frac{1}{2}\phi(\partial^2 + m^2)\phi + J\phi + \lambda\phi^4}. \tag{B.13}$$

By expanding the exponential $e^{\int d^4w \lambda \phi^4(w)} = 1 + \lambda \int d^4w \phi^4(w) + \lambda^2 \int d^4w_1 \int d^4w_2 \phi^4(w_1)\phi^4(w_2)/2 +$ ... we can get closer and closer to the exact solution. We can rewrite this in terms of derivatives of J so that we can use our solution from before,

$$Z = e^{i\lambda \int d^4w \left(-i \frac{\partial}{\partial J(w)}\right)^4} \int D\phi e^{i \int d^4x - \frac{1}{2}\phi(\partial^2 + m^2)\phi + J\phi} \tag{B.14}$$

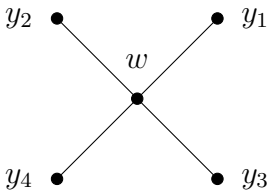
$$= e^{i\lambda \int d^4w \left(-i \frac{\partial}{\partial J(w)}\right)^4} Z(J=0) e^{1/2 \int d^4x \int d^4y J(x) D(x-y) J(y)}. \tag{B.15}$$

The first order connected correction is,

$$\sim i\lambda \int d^4w(-i)^4 \left(\int d^4y D(w-y)J(y) \right)^4 . \tag{B.16}$$

Note for compactness I've used the same variable for the four different integrals over y. For sources that are delta functions at points y_i,

$$\sim i\lambda \int d^4w D(w-y_i)^4 = \tag{B.17}$$



Evaluating the correction,

$$i\lambda \int d^4w D(w-y_i)^4 = i\lambda \int d^4w \left(\int \frac{d^4k_i}{(2\pi)^4} \frac{ie^{ik_i(w-y_i)}}{k_i^2 - m^2 + i\epsilon} \right)^4 \tag{B.18}$$

$$= i\lambda \left(\int \frac{d^4k_i}{(2\pi)^4} \frac{ie^{-ik_i(y_i)}}{k_i^2 - m^2 + i\epsilon} \right)^4 \int d^4w e^{iw(k_1+k_2+k_3+k_4)} \tag{B.19}$$

$$= i\lambda \left(\int \frac{d^4k_i}{(2\pi)^4} \frac{ie^{-ik_i(y_i)}}{k_i^2 - m^2 + i\epsilon} \right)^4 (2\pi)^4 \delta^4(k_1+k_2+k_3+k_4) \tag{B.20}$$

APPENDIX C

Fermion fields

In the previous chapter we introduced fundamental constraints that fields must abide by to be congruent with our understanding of the Universe. We introduced the concept of matter waves and showed that they must satisfy the energy conservation constrains of the Klein-Gordon (KG) equation. The particles in the previous chapter however, have no additional structure to produce the exclusion and phase shift properties shown in experiments with fermions. Though historically not derived to solve this issue, the Dirac equation did solve this issue by happen stance. Principally additional solutions to the KG equation were sought and Dirac's idea was to find solutions that also satisfied the linear energy conservation equation $\hat{p}^\mu a_\mu \psi = mc\psi$ for constant Lorentz vector a_μ.

In section C.1 we derive the constrains on the coefficients a_μ which are necessary so that the Dirac equation can be satisfied. In section C.2 I show that a_μ can be a true Lorentz vector and enforce the covariance of the Dirac equation to determine the transformations of the field vectors ψ under a Lorentz transformation. In section C.3 we derive one possible way to write the Dirac Lagrangian and generalise it to the non-free case. In section C.4 I derive the general form of the ψ fields. In section C.5 I provide a solution to removing negative energy solutions in the ψ field general solution by introducing anti-commuting numbers as coefficients. The anti-commuting coefficients give the fields the fermionic properties we see in experiments. In section C.6 we determine the free solution to the fermion partition function by introducing the concept of Grassmann integration.

C.1 The Dirac Equation

If one could find a free field ψ which satisfied $i\hbar\partial_\mu a^\mu\psi = mc\psi$ Lorentz covariantly for some coefficients a^μ, not dependent on x, then it could be made to satisfy the Klein-Gordon equation. Under this constraint,

$$(i\hbar\partial^\mu a_\mu)(i\hbar\partial^\nu a_\nu\psi) = -\hbar^2\partial^\mu a_\mu a_\nu\partial^\nu\psi = -\hbar^2\partial^2\psi = m^2c^2\psi \tag{C.1}$$

In doing this, the cross terms, for example $\partial_0 a^0\partial_1 a^1\psi$, disappear. This means

$$\{a_\mu, a_\nu\} = 2g_{\mu\nu}I, \tag{C.2}$$

where I is the identity. The components of the wave function derivative at a point in space-time seem to have an anti-commuting structure, and this occurs at all space-time locations. Note that, since only the derivative and ψ depend on x, Lorentz covariance would suggest if x transforms like Λ and ψ transforms like $S(\Lambda)$ that $S^{-1}a^\mu S = \Lambda^\mu{}_\alpha a^\alpha$. More on this later.

Scalar numbers cannot represent an anti-commuting algebra, but an anti-commuting algebra can be represented by matrices. Matrices do not form a group under the anti-commutation bracket $\{\}$, because this would include the zero matrix, which has no inverse. They do form a group under matrix multiplication, which includes all the products of a_μ and $\pm I$ where each a_μ appears only once.

We see from the relation in Eq. (C.2) that the first matrix is self-inverse, but the other matrices produce $-I$ when multiplied by themselves. Suppose however that $\{a_\mu, a_\nu\} = 2\delta_{\mu\nu}I$. Given the anti-commuting nature of the group, we see that $(a_0a_1)^2 = -a_1a_0a_0a_1 = -I$, and so we can easily construct the more complicated $g_{\mu\nu}$ structure if we first find a solution to $\{a_\mu, a_\nu\} = 2\delta_{\mu\nu}I$.

The anti-commuting condition states $a_\mu a_\nu = -a_\nu a_\mu$ $\mu \neq \nu$,

$$\det(a_\mu a_\nu) = \det(-a_\nu a_\mu) = (-1)^N \det(a_\nu)\det(a_\mu) \tag{C.3}$$

where N is the dimension of the square matrix a_ν. We can show the last step is true because the determinant is the product of the eigenvalues of the matrix. This shows that N must be an even number, and we already know the dimension must be greater than 1. Ideally we want the smallest matrix to describe the group, so let's start with $N = 2$. From the relation $\{a_\mu, a_\nu\} = 2\delta_{\mu\nu}I$ all of the matrices are self inverse,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2+bc & b(a+d) \\ c(a+d) & bc+d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{C.4}$$

Hence $c = d = 0$ or $a + d = 0$. If $c = d = 0$ then $a^2 = 1$ and $d^2 = 1$ hence the possibilities are I and $\sigma_1 = \pm\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. If $a + d = 0$ then the possibilities

are $\pm\begin{pmatrix}\sqrt{1-bc} & b\\ c & -\sqrt{1-bc}\end{pmatrix}$. The on diagonal components are proportional to the matrix σ_1 in the first case, but as we want these new matrices to anti-commute with σ_1, the components proportional to σ_1 need to be zero. Enforcing this condition, we produce the matrices $\sigma_2 = \pm\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \sigma_3 = \pm\begin{pmatrix}0 & i\\ -i & 0\end{pmatrix}$. If we fix the signs out of the front of the matrices, we find that they anti-commute. Unfortunately, there are only three anti-commuting matrices with dimension $N = 2$, but we need four to describe our group.

The next smallest matrix is $N = 4$. Using what we learned from $N = 2$ we can quickly construct four matrices that anti-commute and are self inverse,

$$\pm\begin{pmatrix}\sigma_i & 0\\ 0 & -\sigma_i\end{pmatrix}, \pm\begin{pmatrix}0 & I\\ I & 0\end{pmatrix},\ i \in \{1,2,3\} \tag{C.5}$$

where we have used block notation. There is more than one way to write four anti-commuting matrices using block diagonal form, the one above is called the Weyl representation. Another choice is,

$$\pm\begin{pmatrix}I & 0\\ 0 & -I\end{pmatrix}, \pm\begin{pmatrix}0 & \sigma_i\\ \sigma_i & 0\end{pmatrix},\ i \in \{1,2,3\} \tag{C.6}$$

called the Dirac representation.

Using these (after fixing the sign) and the identity we can generate a matrix representation for the entire group. These matrices are called Γ matrices. If we set $a_0 = \begin{pmatrix}0 & I\\ I & 0\end{pmatrix}$ and $a_i = \begin{pmatrix}0 & -\sigma_i\\ \sigma_i & 0\end{pmatrix}$ (Weyl representation) we produce the relation $\{a_\mu, a_\nu\} = 2g_{\mu\nu}I$. Alternatively using the Dirac representation, $a_0 = \begin{pmatrix}I & 0\\ 0 & -I\end{pmatrix}$ and $a_i = \begin{pmatrix}0 & \sigma_i\\ -\sigma_i & 0\end{pmatrix}$ produces the same group.

For brevity we often write the Dirac equation as $i\hbar\not{\partial}\psi = mc\psi$ where $\not{\partial} = a^\mu\partial_\mu$. To write down the simplest Lagrangian we need to know how to produce a Lorentz invariant out of the fields we have been given. To this end, and for further understanding of ψ we will explore the structure of these new fields ψ, that satisfy the Dirac equation.

C.2 Covariance of the Dirac Equation

We assumed from the outset that we could satisfy both a constant coefficient a^μ and Lorentz invariance of the Dirac equation. We now need to show that this is

not contradicted. In the process we will determine how the ψ vector transforms under a Lorentz transformation. Using the Lorentz invariant condition, we want to determine how the field ψ transforms under Lorentz transformations.

$$i\hbar g_{\mu\nu}a^\mu\partial^\nu\psi = mc\psi \tag{C.7}$$
$$i\hbar g_{\mu\nu}a^\mu\Lambda^\nu{}_\alpha\partial^\alpha S(\Lambda)\psi = mcS(\Lambda)\psi \tag{C.8}$$
$$i\hbar S^{-1}(\Lambda)a^\mu S(\Lambda)g_{\mu\nu}\Lambda^\nu{}_\alpha\partial^\alpha\psi = mc\psi \tag{C.9}$$
$$i\hbar S^{-1}(\Lambda)a^\mu S(\Lambda)g_{\mu\nu}\Lambda^\nu{}_\alpha\partial^\alpha\psi = i\hbar g_{\mu\nu}a^\mu\partial^\nu\psi \tag{C.10}$$

Which would imply that,

$$S^{-1}(\Lambda)a^\mu S(\Lambda)g_{\mu\nu}\Lambda^\nu{}_\alpha = a^\beta g_{\beta\alpha} \tag{C.11}$$
$$S^{-1}(\Lambda)a^\mu S(\Lambda) = \Lambda^\mu{}_\beta a^\beta \tag{C.12}$$

By comparing the expansions of the transformations about the identity it can be shown that if $\Lambda = \exp(-\frac{i}{2}\omega_{\mu\nu}M^{\mu\nu})$, where M are the generators of the Lorentz group, then $S(\Lambda) = \exp\left(-\frac{i}{4}\omega_{\mu\nu}\sigma^{\mu\nu}\right)$ where $\sigma^{\mu\nu} = \frac{i}{2}\left[\gamma^\mu, \gamma^\nu\right]$ is the generator for the spinor transformation. Note that the coefficients ω are shared. This proof is long and so I do not include it here. As a solution exists, we have shown that there is no contradiction. It can also be shown that for boosts,

$$S(\Lambda) = \frac{\gamma^0\not{p} + m}{\sqrt{2m(p^0 + m)}}. \tag{C.13}$$

This proof is also long and so we will not present it here. Using this transformation we can boost our solutions of ψ from the rest frame.

C.3 Dirac Lagrangian

In order to create the Lagrangian we need to create a Lorentz scalar from the field ψ. We will need to use $\psi^\dagger$ in some way, because to produce a scalar we need an element from the dual space of ψ. To find the equation that governs $\psi^\dagger$ consider,

$$i\hbar\not{\partial}\psi = mc\psi \tag{C.14}$$
$$-i\hbar\partial^\mu\psi^\dagger a_\mu^\dagger = mc\psi^\dagger \tag{C.15}$$

Using the Dirac representation for the Γ matrices as previously, $a_0 = a_0^\dagger$ and $a_i^\dagger = \begin{pmatrix} 0 & -\sigma_i^\dagger \\ \sigma_i^\dagger & 0 \end{pmatrix}$. However, $\sigma_i^\dagger = \sigma_i$ and so $a_i^\dagger = -a_i = a_0 a_i a_0$; more generally

$a_\mu^\dagger = a_0 a_\mu a_0$. So we can rewrite the $\psi^\dagger$ Dirac equation as,

$$-i\hbar\partial^\mu\psi^\dagger a_0 a_\mu a_0 = mc\psi^\dagger \tag{C.16}$$

$$-i\hbar\partial^\mu\psi^\dagger a_0 a_\mu = mc\psi^\dagger a_0 \tag{C.17}$$

$$-i\hbar\partial^\mu\bar{\psi} a_\mu = mc\bar{\psi} \tag{C.18}$$

The $\psi^\dagger$ field transforms like $S^\dagger$, however looking at the form for $S = e^{i\omega_{\mu\nu}\sigma^{\mu\nu}}$, S satisfies $S^\dagger a^0 = a^0 S^{-1}$. Hence, we can create a Lorentz scalar by $\bar{\psi}\psi$. By making the Lagrangian, $L = \bar{\psi}(i\partial\!\!\!/ - mc)\psi$ we produce a Lorentz invariant. The Lagrangian can be verified using the Euler-Lagrange equations, which produce the Dirac equation for both fields correctly. This is the simplest Lagrangian that produces such a result for a free field. The Lagrangian can be generalised to the non-free case by adding a potential term, $L = \bar{\psi}(i\partial\!\!\!/ - mc)\psi + V(\psi, \bar{\psi})$.

C.4 Four Component Spinors

Let's look a little closer at the equation we have produced and try to determine what ψ looks like. The field ψ has four components, but these are not Lorentz components, the vector is a module of the group of Γ matrices. It is not obvious from inspection how the four components of the field are related;

$$i\hbar\partial^0 a_0\psi = -i\hbar\vec{\nabla}\cdot\vec{a}\psi + mc\psi. \tag{C.19}$$

To get a better idea we will take the limit $\vec{p} = 0$, where in this limit $\vec{p} = i\hbar\vec{\nabla}$. The Dirac equation reduces to,

$$i\hbar a_0\frac{\partial\psi}{\partial ct} = mc\psi \tag{C.20}$$

If we choose the set of Γ matrices in the Dirac representation, then it is clear the two upper components decouple from the lower two components, so we write $\psi = \begin{pmatrix}\eta\\ \chi\end{pmatrix}$. Hence, we get two uncoupled differential equations,

$$i\hbar\frac{\partial\eta}{\partial t} = mc^2\eta, \tag{C.21}$$

$$-i\hbar\frac{\partial\chi}{\partial t} = mc^2\chi. \tag{C.22}$$

Solving these equations we produce, $\eta = e^{-imc^2t/\hbar}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$, $\chi = e^{imc^2t/\hbar}\begin{pmatrix}\kappa\\ \nu\end{pmatrix}$ for constants $\alpha, \beta, \kappa, \nu$. This is pretty clearly not Lorentz invariant, as we don't have

a product of Lorentz vectors producing the scalar in the exponential; in fact we only have part of a known scalar $k \cdot x$. Additionally we are looking for plane wave solutions for the free Dirac equation. This can be remedied without changing the result, by replacing the solution with $\psi = e^{-ikx}\begin{pmatrix}\eta\\0\end{pmatrix}$ and $\psi = e^{ikx}\begin{pmatrix}0\\\chi\end{pmatrix}$, which are plane waves.

The solutions seem to fit into two distinct classes. Solutions with e^{ikx} seem to be negative energy solutions (which is a little absurd), while e^{-ikx} have positive energy solutions. We want to see if we can determine the form of the solution away from the rest frame. Plane wave solutions (free particle solutions) that are not restricted to the rest frame will satisfy the Dirac equation. Hence, for positive energy solutions, $(\not{p} - m)u(p,\eta) = 0$ and for negative energy solutions, $(-\not{p} - m)v(p,\chi) = 0$. However this means that u and v are only functions of momentum, and there is only one relation involving p which the particles satisfy apart from the Dirac equation; the Klein-Gordon (KG) equation. These solutions are proportional to relations which result in the KG equation upon applying the Dirac equation. These are $u(p,\eta) = C(\not{p} + m)\begin{pmatrix}\eta\\0\end{pmatrix}e^{-ikx}$ and $v(p,\chi) = C'(\not{p} - m)\begin{pmatrix}0\\\chi\end{pmatrix}e^{ikx}$. Of course, in the $\vec{p} = 0$ limit the solutions reduce to the rest frame solutions. As such we can fix the constants,

$$u(p,\eta) = \frac{(\not{p} + m_0)}{E/c + m_0 c}\begin{pmatrix}\eta\\0\end{pmatrix}e^{-ikx} \tag{C.23}$$

$$v(p,\chi) = \frac{(\not{p} - m_0 c)}{E/c - m_0 c}\begin{pmatrix}0\\\chi\end{pmatrix}e^{ikx} \tag{C.24}$$

Where I've explicitly written m_0 to make it clear this is the rest mass.

The general form of ψ can be written as the sum of plane wave solutions,

$$\psi(x) = \int \frac{d^4k}{(2\pi)^4}\left[\alpha(p,\eta)u(p,\eta) + \beta(p,\chi)v(p,\chi)\right] \tag{C.25}$$

$$\bar{\psi}(x) = \int \frac{d^4k}{(2\pi)^4}\left[\alpha^\dagger(p,\eta)\bar{u}(p,\eta) + \beta^\dagger(p,\chi)\bar{v}(p,\chi)\right] \tag{C.26}$$

for coefficients α and β. Since ψ has four components and the vectors η and ξ are related to the spin vectors of quantum mechanics, these vectors are called four component spinors.

C.5 Vacuum Energy

What may have set off alarm bells in the above derivation is negative energy solutions. How can one have negative energy, can this make sense and apply to

the real world?

To this end let us see if our vacuum energy makes sense. Starting from the Dirac equation, $i\hbar\partial^0 a_0\psi = -i\hbar\vec{\nabla}\cdot\vec{a}\psi + mc\psi$ pre-multiply by $\bar{\psi}$ to produce a scalar which is proportional to the vacuum energy. If we set the normalisation $\psi^\dagger\psi = 1$;

$$H = \int \frac{d^4k}{(2\pi)^4}[\alpha^\dagger\bar{u} + \beta^\dagger\bar{v}][\alpha(-\not{p} + \gamma^0 p_0 + m)u + \beta(\not{p} - \gamma^0 p_0 + m)v]. \tag{C.27}$$

By using the identity $(\not{p} - m)u = 0$, then taking the Hermitian conjugate and commuting the γ^0 $\bar{u}\not{p} = m\bar{u}$, similarly we find that $\bar{v}\not{p} = -m\bar{v}$. We can also use the identity $\bar{u}\gamma^\mu u = \frac{p^\mu}{m}\bar{u}u$, $\bar{v}\gamma^\mu v = \frac{p^\mu}{m}\bar{v}v$. We'll also take the normalisation $\bar{u}u = 1$ and $\bar{v}v = -1$, and using the rest frame we show $\bar{v}u = \bar{u}v = 0$ which is a Lorentz invariant equation. So we find,

$$H = \int \frac{d^4k}{(2\pi)^4}\alpha^\dagger\alpha\bar{u}(-\not{p} + \gamma^0 p_0 + m)u + \beta^\dagger\beta\bar{v}(\not{p} - \gamma^0 p_0 + m)v \tag{C.28}$$

$$= \int \frac{d^4k}{(2\pi)^4}\frac{E^2(p)}{m}(\alpha^\dagger(p,\eta)\alpha(p,\eta) - \beta^\dagger(p,\chi)\beta(p,\chi)). \tag{C.29}$$

If $\alpha(p)$ or $\beta(p)$ are non-zero, then we have effectively excited a mode at a given momentum, with positive or negative energy respectively. The coefficients $\alpha^\dagger$ and $\beta^\dagger$ destroy modes with positive and negative energy. However, the above Hamiltonian does not make much sense; it does not have a minimum. It says that I can create as many positive modes as I want, as long as I also create the same number of negative modes. But the u and v modes work in different spaces, so the negative modes v do not destroy u, they destroy $\bar{v}$. It would be better if we only created positive modes and only destroyed positive modes in the two solutions, as this makes physical sense. Such a thing could occur if we could change the order of the negative energy solution, so $\bar{v}$ came before v. This can be achieved, in a way, by making $\{\beta^\dagger, \beta\} = 0$, meaning they are anti-commuting numbers. This abstracts the order of the processes from the physical calculation of the matrix elements, which we wish to remain the same. Of course, we also have to apply anti-commutation relations for $\{\alpha^\dagger, \alpha\} = 0$, because if we were to change the order of $\bar{u}$ and u, the process would be put in terms of negative energy. The Hamiltonian becomes,

$$H = \int \frac{d^4k}{(2\pi)^4}\frac{E^2(p)}{m}(\alpha^\dagger(p,\eta)\alpha(p,\eta) + \beta(p,\chi)\beta^\dagger(p,\chi)) \tag{C.30}$$

By interpreting things this way, we must create positive energy before we can destroy it and we no longer have to worry about negative energy. The anti-commutation of the field coefficients gives the fields the fermionic property, as we see in experiments. Our Hamiltonian now has a minimum energy as well.

C.6 The Path Integral for Dirac Fields

We would like to write down the path integral for Dirac fields like so (setting $c = \hbar = 1$),

$$Z = \int D\psi D\bar{\psi} e^{i\int d^4x \bar{\psi}(i\partial\!\!\!/ - m)\psi - V(\psi,\bar{\psi})} \tag{C.31}$$

However we have not properly defined what an integral of ψ would mean, given its anti-commuting coefficients. For such anti-commuting quantities we introduce the Grassmann algebra. For Grassmann numbers ξ, η, we have the conditions $\eta^2 = 0$ and $\xi\eta = -\eta\xi$. The Taylor series of any function truncates exactly to $f(\eta) = a + b\eta$ for regular numbers a, b. Grassmann integrals obey the rule $\int d\eta\ f(\eta) = \int d\eta f(\eta + \xi)$, hence $\int d\eta b\xi = 0$ which only holds if $\int d\eta b = 0$. The product of two Grassmann numbers is a normal number, so in general $\int d\eta\eta = 1$, where the choice 1 fixes the normalisation of η [25]. One can consider $\int d\eta$ to represent the size of the set in the direction η. Two very helpful identities we obtain are $\int d\eta \int d\bar{\eta} e^{\bar{\eta}a\eta} = \int d\eta \int d\bar{\eta}(1 + \bar{\eta}a\eta) = a$ and $\int d\eta \int d\bar{\eta} e^{\bar{\eta}A\eta} = \det A$.

We can now confidently write the partition function with $\psi, \bar{\psi}$ understood to act like Grassmann numbers, and the integrals over them to be Grassmann integrals. For spinor sources $\eta, \bar{\eta}$,

$$Z(\eta, \bar{\eta}) = \int D\psi D\bar{\psi} e^{i\int d^4x \bar{\psi}(i\partial\!\!\!/ - m)\psi + \bar{\eta}\psi + \bar{\psi}\eta} \tag{C.32}$$

If we discretise and set $K = -(\partial\!\!\!/ + im)$, and complete the square to get $\bar{\psi}K\psi + i\bar{\eta}\psi + i\bar{\psi}\eta = (\bar{\psi} + i\bar{\eta}K^{-1})K(\psi + iK^{-1}\eta) + \bar{\eta}K^{-1}\eta$, then following the same procedure as for the scalar field we find,

$$Z(\eta, \bar{\eta}) = C' \det(K) e^{\int d^4x \int d^4y \bar{\eta}(x)(-\partial\!\!\!/ - im)^{-1}\eta(y)} \tag{C.33}$$

Then the solution to $(-\partial\!\!\!/ - im)S(x - y) = \delta^4(x - y)$, is simply,

$$S(x - y) = \int \frac{d^4p}{(2\pi)^4} \frac{ie^{ip\cdot(x-y)}}{p\!\!\!/ - m + i\epsilon}. \tag{C.34}$$

APPENDIX D

Yang-Mills Gauge theory

In chapter B and chapter C I outlined what rules quantum fields, which make up fundamental particles, obey when they are not acted upon by potentials. Using perturbation theory and the free theory solution, we produced approximate (perturbative) solutions for non-zero interacting potentials V. As we will use Quantum Electrodyanmics (QED) and Quantum Chromodynamics (QCD) to study hadron bound states, in this chapter I will discuss the interaction terms, or potentials, in the Lagrangians of these field theories which allow particles with electric charge (QED) and colour charge (QCD) to interact. I approach this topic in the broad sense of Yang-Mills theories because QCD and QED have many similarities. Essentially they are the same type of force, except QCD has an $SU(3)$ gauge group while QED has a $U(1)$ gauge group. I give more detail on these particulars in chapter 2. In this chapter my goal is to describe what a Yang-Mills type gauge theory is and derive the partition function for this class of interacting field theories.

In section D.1 I introduce the concepts of a gauge field and determine the kinetic term for the gauge field which satisfies the gauge symmetry. In section D.2 I state the free action for the Yang-Mills gauge theory, rearrange the terms in the Lagrangian and attempt to find an integral solution to the partition function. However, complications arise when attempting to invert to produce the propagator. In section D.3 the inversion complication is resolved by gauge fixing the theory. We determine the full partition function and propagators for the gauge field and ghost propagator.

D.1 Yang-Mills Non-Abelian Gauge Theory

The gauge principle states that under some transformation, $\phi \to \phi'$ the action remains invariant. In such cases, there is not a unique solution for ϕ but rather a class of solutions. Furthermore, if we have one of the solutions we can generate the rest using the transformation.

Suppose for some complex N component scalar field, $L = \partial\phi^\dagger\partial\phi - m^2\phi^\dagger\phi + V(\phi^\dagger, \phi)$, the Lagrangian is invariant under some $SU(N)$ transformation U. A Lagrangian with such a symmetry would always have ϕ and $\phi^\dagger$ occurring in pairs in each potential term. This is called a global symmetry, because the field component at each space-time point is transformed by the same matrix U.

Yang and Mills proposed a deeper symmetry [20]. Suppose the same N component field was invariant under an $SU(N)$ transformation, but possibly a different element of $SU(N)$ at each point in space-time $U(x)$. This is called a local symmetry, as the field can transform independently at each local space-time point x. For potential terms with no derivative this follows easily, with some constraints. For derivative terms we get $\partial_\mu\phi \to \partial_\mu(U(x)\phi) = U[\partial_\mu\phi + (U^\dagger\partial_\mu U)\phi$. Hence, the term $(U^\dagger\partial_\mu U)\phi$ must cancel with some other term in the Lagrangian. To develop this term, we introduce the concept of a covariant derivative D_μ. It is defined so that $D_\mu\phi(x) \to U(x)D_\mu\phi(x)$ and we can write its form generally as $D_\mu\phi(x) = \partial_\mu\phi(x) - iA_\mu(x)\phi(x)$, where $A_\mu(x)$ is called a gauge field. By using the ϕ transform we can construct how A_μ transforms: $A_\mu \to UA_\mu U^\dagger - i(\partial_\mu U)U^\dagger = UA_\mu U^\dagger + iU\partial_\mu U^\dagger$.

Before continuing, let us learn a little about the field A_μ. Firstly, the field A_μ leaves the action invariant under the transformation $U = e^{i\theta\cdot T}$, with T^a the generators of $SU(N)$. We consider the infinitesimal transformation to be sufficient in describing this transformation, as consecutive applications of this produces any element in the group. Hence, we are letting $U \approx 1 + i\theta \cdot T$ and,

$$A_\mu \to A_\mu + i\theta^a[T^a, A_\mu] + \partial_\mu\theta^a T^a. \tag{D.1}$$

There are several interesting facts that we should know about the fields A_μ. They are $N \times N$ matrices. $A_\mu - A_\mu^\dagger = 0$, can be shown to be gauge invariant, hence we can take A_μ to be Hermitian. The trace of A_μ is gauge invariant, hence we can set A_μ to be traceless. As T^a are also traceless, they span the space hence, $A_\mu = a_\mu^a T^a$. Note, the scalar fields a_μ^a make up a Lorentz vector. T^a are also orthogonal under the trace operation, Hermitian and self-inverse. The generators T^a satisfy $[T^a, T^b] = if^{abc}T^c$, for structure constants f^{abc}. Hence we can write,

$$A_\mu^b \to A_\mu^b + i\theta^a a_\mu^d[T^a, T^d] + \partial_\mu\theta^a T^a \tag{D.2}$$

$$= A_\mu^b - f^{adc}\theta^a a_\mu^d T^c + \partial_\mu\theta^a T^a. \tag{D.3}$$

Right multiplying by T^b and taking the trace of both sides produces,

$$a^b_\mu \to a^b_\mu - f^{adb}\theta^a a^d_\mu + \partial_\mu \theta^b. \tag{D.4}$$

The gauge field A_μ is also a field, and hence in free space should also satisfy the KG equation. However thus far our Lagrangian does not have a kinetic energy term for the field A_μ which would involve $(\partial A)^2$ in some way. We would also desire this new term to be gauge invariant. This is most easily determined using differential forms. Let $B_\mu = iA_\mu$, then $B = B_\mu dx^\mu$ is a 1-form. The gauge transformation in this notation is,

$$B \to UBU^\dagger + UdU^\dagger. \tag{D.5}$$

Note d acts on anything to the right of it in a given term. The quantity we want is something like the 4-form $(dB)^2$. As such let us look at the 2-form dB transformation,

$$dB \to UdBU^\dagger + dUBU^\dagger - UBdU^\dagger + dUdU^\dagger. \tag{D.6}$$

It does not transform how we would like, seemingly we only want $UdBU^\dagger$ so that when we square we get something that is at least covariant with the group transformation. The only other 2-form we can add is B^2, which transforms like,

$$B^2 \to UB^2U^\dagger + UBdU^\dagger + UdU^\dagger UBU^\dagger + UdU^\dagger UdU^\dagger. \tag{D.7}$$

Noting that $UU^\dagger = 1$ and $UdU^\dagger = -dUU^\dagger$, hence

$$B^2 \to UB^2U^\dagger + UBdU^\dagger - dUBU^\dagger - dUdU^\dagger. \tag{D.8}$$

Now the terms in the B^2 transformation cancel many of the terms in the dB transformation, $dB + B^2 \to U(dB + B^2)U^\dagger$. The square of this transforms in the same way. Hence, define the Yang-Mills field strength tensor as $F = dB + B^2$. In more explicit terms

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu, A_\nu]. \tag{D.9}$$

We can also put this in terms of the scalar component fields a^b_μ by right multiplying $F_{\mu\nu}$ by T^b and taking the trace, $F^b_{\mu\nu} = \partial_\mu a^b_\nu - \partial_\nu a^b_\mu + f^{adb}a^a_\mu a^d_\nu$. Finally, the kinematic term we add to the Lagrangian is $-\frac{1}{2g^2}tr\left(F_{\mu\nu}F^{\mu\nu}\right)$, for coupling constant g. The full Yang-Mills Lagrangian is,

$$L = D\phi^\dagger D\phi - m^2\phi^\dagger\phi - \frac{1}{2g^2}tr\left(F_{\mu\nu}F^{\mu\nu}\right) + V(\phi^\dagger, \phi, \partial_\mu\phi^\dagger, \partial_\mu\phi, A_\mu, \partial A). \tag{D.10}$$

D.2 Quantising the Yang-Mills Lagrangian

Let us work with just the gauge field part of the Lagrangian, $L = -\frac{1}{2g^2} tr\,(F_{\mu\nu}F^{\mu\nu})$. In terms of component fields,

$$S(a^a) = -\frac{1}{4}\int d^4x(\partial_\mu a^a_\nu - \partial_\nu a^a_\mu)(\partial^\mu a^\nu_a - \partial^\nu a^\mu_a) \tag{D.11}$$

$$+ 2g(\partial_\mu a^a_\nu - \partial_\nu a^a_\mu) f^{abc} a^\mu_b a^\nu_c + g^2 f^{abc} f^{ade} a^b_\mu a^c_\nu a^\mu_d a^\nu_e. \tag{D.12}$$

Note the component index in latin letters does not have an up and down index convention, it has only been placed opposite to the Lorenz index for clarity and convenience. Also, I have normalised $[T^a, T^b] = \frac{1}{2}\delta_{ab}$, and redefined $A \to gA$. The kinematic term on the first line can be rearranged using integration by parts to give (dropping the component index for convenience),

$$S(a^a) = -\frac{1}{4}\int d^4x g^{\mu\alpha} g^{\nu\beta}(\partial_\mu a_\nu \partial_\alpha a_\beta - \partial_\mu a_\nu \partial_\beta a_\alpha - \partial_\nu a_\mu \partial_\alpha a_\beta + \partial_\nu a_\mu \partial_\beta a_\alpha) \tag{D.13}$$

$$= \frac{1}{4}\int d^4x g^{\mu\alpha} g^{\nu\beta}(a_\nu \partial_\mu \partial_\alpha a_\beta - a_\nu \partial_\mu \partial_\beta a_\alpha - a_\mu \partial_\nu \partial_\alpha a_\beta + a_\mu \partial_\nu \partial_\beta a_\alpha) \tag{D.14}$$

$$= \frac{1}{4}\int d^4x(a_\nu \partial_\mu \partial^\mu a^\nu - a_\nu \partial_\mu \partial^\nu a^\mu - a_\mu \partial_\nu \partial^\mu a^\nu + a_\mu \partial_\nu \partial^\nu a^\mu) \tag{D.15}$$

$$= \frac{1}{2}\int d^4x a_\nu (g^{\mu\nu}\partial^2 - \partial^\mu \partial^\nu) a_\mu \tag{D.16}$$

Taking the remaining terms as parts of the potential, the partition function is then given by,

$$Z = \int DA e^{i\int d^4x \frac{1}{2} tr\left[A_\nu(\partial^2 g^{\mu\nu} - \partial^\mu\partial^\nu)A_\nu\right] + V(A)} \tag{D.17}$$

At this point we would normally discretise the path integral and solve the free theory, which involves inverting the matrix between the two A fields, $Q^{\mu\nu} = (\partial^2 g^{\mu\nu} - \partial^\mu \partial^\nu)$. However, applying Q to the derivative of a scalar field Λ, $Q^{\mu\nu}\partial_\nu \Lambda(x) = 0$. Hence, the null space of the matrix Q is non-trivial and the matrix has no inverse.

D.3 Gauge Fixing Yang-Mills

The nature of the gauge fixing problem is surprisingly mundane [25]. The matrix between the two fields in the partition function, the part we would normally invert to produce the propagator, has no inverse. This means that the matrix is not

linearly independent, or rather its null space is not trivial. In other words, if we were able to parameterise using less dimensions the problem would not have arisen at all.

As a concrete example, consider the integral $\int_{-\infty}^{\infty} dAe^{-A\cdot K\cdot A}$ for $A = (a, b)$ and $K = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. The integral becomes $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dadbe^{-a^2}$, which is not finite and does not make much sense. The problem is fixed by including the delta function $\delta(b - \xi)$. This inclusion does not depend on ξ and gives us a finite answer to the initial problem we wanted to know, but had posed incorrectly.

In field theory we wish to do this same thing. Suppose we had the integral $I = \int DAe^{iS(A)}$, where the action S is invariant under some transformation, $A \to A_g$. These transformations form a group. Our goal is to transform the coordinates so that we factor out this freedom (or null space), $I = \left(\int Dg\right)\int DA'e^{iS(A')}$, which will give us a unique solution to the inverse. If the group is compact it will also give us a finite answer, if it is not then we have separated the infinite part from the rest of the integral. This factorisation of the group integral will also occur at the vacuum level, $Z(0)$, and hence in any measurement it will cancel out, rendering the total answer finite. A good example of this procedure is $I = \int dxdye^{iS(x,y)}$, where $S(x, y)$ is only a function of $x^2 + y^2$. By changing variables $I = (\int d\theta)\int drre^{iS(r)}$ [25].

Now we are in a position to gauge fix the Yang-Mills Lagrangian. The problem is clear; there are some additive vector field directions for the fields a^b which are unconstrained but are integrated over. Hence, we have infinite repeats of the same dynamics; the same integral is repeated with different gauges. To preserve Lorentz invariance, it is a good idea to work with Lorentz scalars. So that we do not add any additional fields, our choices are x^μ or ∂^μ. We know that a free particle has constant momentum, hence we use ∂^μ,

$$\partial a^b \to \partial a^b - f^{adb}\partial^\mu(\theta^a a^d_\mu) + \partial^2\theta^b. \tag{D.18}$$

Now at each space-time point we set $\partial a^b(x) - \sigma^b = 0$, where σ^b is some scalar. In addition we also require $-f^{adb}\partial^\mu(\theta^a a^d_\mu) + \partial^2\theta^b = 0$, otherwise $\partial a^b(x)$ would be able to gauge transform away from the fixed value σ^b. We this for each field a^b.

In the below demonstration I will factorise the part of the partition function that integrates over the gauge freedom. Starting with the partition function for Yang-Mills theory, $Z = \int DAe^{iS(A)}$, then multiply by 1,

$$Z = \frac{\int DAe^{iS(A)}\int D\theta\frac{\delta(-f^{adb}\partial^\mu(\theta^a a^d_\mu)+\partial^2\theta^b)}{\delta(\partial a^b-\sigma^b)}}{\int D\theta\frac{\delta(-f^{adb}\partial^\mu(\theta^a a^d_\mu)+\partial^2\theta^b)}{\delta(\partial a^b-\sigma^b)}} \tag{D.19}$$

Rearranging and evaluating the delta functions,

$$Z = \frac{\int DAe^{iS(A)}\delta(\partial a^b - \sigma^b)\int D\theta \frac{\delta(-f^{adb}\partial^\mu(\theta^a a^d_\mu)+\partial^2\theta^b)}{\delta(\partial a^b-\sigma^b)}}{\int D\theta\delta(-f^{adb}\partial^\mu(\theta^a a^d_\mu)+\partial^2\theta^b)} \tag{D.20}$$

$$= \frac{\int DA' e^{iS(A')}\left(\int D\theta\right)}{\int D\theta'} \tag{D.21}$$

This implies that our gauge is one of an infinite number of gauges. If we only allow for paths with $\partial a^b - \sigma^b = 0$ and $-f^{adb}\partial^\mu(\theta^a a^d_\mu) + \partial^2\theta^b = 0$, simply remove $\int D\theta$ from the top line,

$$Z = \frac{\int DAe^{iS(A)}\delta(\partial a^b - \sigma^b)}{\int D\theta\delta(-f^{adb}\partial^\mu(\theta^a a^d_\mu)+\partial^2\theta^b)} \tag{D.22}$$

and the partition function is now gauge fixed.

Looking at the term on the bottom of Eq. (D.22) we can write,$-f^{adb}\partial^\mu(\theta^a(x)a^d_\mu)+\partial^2\theta^b(x) = \int d^4y\left[-f^{adb}\partial^\mu a^d_\mu + \partial^2\delta^{ba}\right]\delta^4(x-y)\theta^a(y)$. Hence, we let $K^{ab}(x,y) = \left[-f^{adb}\partial^\mu a^d_\mu + \partial^2\delta^{ba}\right]\delta^4(x-y)$, so the denominator can be written like $\int d\theta K\theta = \frac{1}{\det K}$ for matrix K. We can use the Grassmann field identity from section C.6 to rewrite it in terms of an exponential,

$$\begin{aligned}
\det(K) &= \int DcDc^\dagger e^{iS_{ghost}(c^\dagger,c)} \\
S_{ghost}(c^\dagger,c) &= \int d^4x \int d^4y\ c^\dagger_a(x)K^{ab}(x,y)c_b(y) \\
&= \int d^4x\ c^\dagger_a(x)\partial^2 c_a(x) - c^\dagger_a(x)\partial^\mu(f^{abc}a^c_\mu(x)c_b(x))
\end{aligned} \tag{D.23}$$

Faddeev and Popov derived a rigorous procedure which allows one to gauge fix non-abelian $SU(N)$ gauge fields [20]. To ensure that a gauge fixing condition with a functional delta produces the correct normalisation, we need to ensure that the measure is adjusted; not unlike changing variables in normal integration using a Jacobian. The form of this correction is obtained by generalising the identity in discrete n-dimensional vectors [20],

$$1 = \left(\prod_i \int da_i\right)\delta^{(n)}(\vec{G}(\vec{a}))\ \det\left(\frac{\partial g_i}{\partial a_j}\right), \tag{D.24}$$

In analogy to this identity (though I will not prove it) the continuously infinite (functional) generalisation is,

$$1 = \int D\alpha(x)\delta(G(A))\ \det\left(\frac{\delta G(A)}{\delta\theta}\right). \tag{D.25}$$

The factorised partition function Z becomes,

$$Z = \int DAe^{iS[A]} = \int D\theta \int DAe^{iS[A]}\delta(G(A))\ \det\left(\frac{\delta G(A)}{\delta\theta}\right). \tag{D.26}$$

If we set the gauge fixing condition $G(a^a) = \partial^\mu a^a_\mu(x) - \omega^a(x) = 0$, then

$$\frac{\delta G(A)}{\delta\theta} = \frac{1}{g}\partial^\mu D_\mu, \tag{D.27}$$

and so using the Grassmann integral identity from section C.6, the Faddeev-Popov determinant is,

$$\det\left(\frac{1}{g}\partial^\mu D_\mu\right) = \int DcD\bar{c}\exp\left[i\int d^4x\bar{c}(-\partial^\mu D_\mu)c\right], \tag{D.28}$$

$$= \int DcD\bar{c}\exp\left[i\int d^4x L_{ghost}\right], \tag{D.29}$$

where $L_{ghost} = \bar{c}^a(-\partial^2\delta^{ac} - g\partial^\mu f^{abc}a^b_\mu)c^c$.

To tie up some loose ends, the gauge transform is a symmetry of the action, and so we should not force it on our integral. Instead of a δ we can use $e^{\frac{i}{2\xi}\int d^4x\ tr\left[(\partial A)^2\right]}$, for some real constant ξ – notice ξ can not be zero. We can think of it as the path integral definition of a delta function, provided we are integrating over A. This new condition does not strictly enforce the constraint but suppresses solutions that violate it. The other advantage is that we fix all the fields at the same time when we use $(\partial A)^2$; taking the trace separates out all the component fields. Another way of arriving at the same conclusions; because we divide by $Z(0)$ for any measurement, we can change Z by a multiplicative factor. By multiplying Z by $\int D\sigma e^{\frac{i}{2\xi}\int dx\ tr(\sigma^2)}$, we allow an infinite number of gauges centred around $\partial a^b = 0$, Gaussian weighted.

Our gauge fixed partition function becomes,

$$Z = \int DADc^\dagger Dce^{iS(A)+iS_{ghost}(c^\dagger,c)-\frac{i}{2\xi}\int d^4x\ tr\left[(\partial A)^2\right]} \tag{D.30}$$

By looking at the final partition function, we can see the additional terms that we have added are Lagrange multipliers. Once we determine the constraints $\partial a^b(x) - \sigma^b = 0$ and $-f^{adb}\partial^\mu(\theta^a a^d_\mu) + \partial^2\theta^b = 0$, we can then use Lagrange multipliers to enforce the constraints. The Faddeev-Popov procedure ensures we get a suitable Lagrange multiplier, but following the procedure is not required; other Lagrange multipliers are possible.

Finally, we can write the gauge boson propagator as,

$$G^{ab}_{\mu\nu} = \int \frac{d^4k}{(2\pi)^4}\frac{-ie^{ixk}}{k^2}\left[g_{\mu\nu} - (1-\xi)\frac{k_\mu k_\nu}{k^2}\right]\delta_{ab} \tag{D.31}$$

and the ghost propagator as,

$$K_{ab} = \int \frac{d^4k}{(2\pi)^4} \frac{ie^{ixk}}{k^2} \delta_{ab} \tag{D.32}$$