Least-Squares Finite Element Methods for Quantum Electrodynamics

by

Christian W. Ketelsen

B.S., Washington State University, 2003

M.S., Washington State University, 2005

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Applied Mathematics

2010

This thesis entitled:
Least-Squares Finite Element Methods for Quantum Electrodynamics
written by Christian W. Ketelsen
has been approved for the Department of Applied Mathematics

Prof. Thomas A. Manteuffel

Prof. Stephen F. McCormick

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Ketelsen, Christian W. (Ph.D., Applied Mathematics)

Least-Squares Finite Element Methods for Quantum Electrodynamics

Thesis directed by Prof. Thomas A. Manteuffel

The numerical solution of the Dirac equation is the main computational bottleneck in the simulation of quantum electrodynamics (QED) and quantum chromodynamics (QCD). The Dirac equation is a first-order system of partial differential equations coupled with a random background gauge field. Traditional finite-difference discretizations of this system are sparse and highly structured, but contain random complex entries introduced by the background field. For even mildly disordered gauge fields the near kernel components of the system are highly oscillatory, rendering standard multilevel iterative methods ineffective. As such, the solution of such systems accounts for the *vast* majority of computation in the simulation of the theory.

In this thesis, two discretizations of a simplified model problem are introduced, based on least-squares finite elements. The first discretization is obtained by direct discretization of the governing equation using least-squares finite elements. The second is obtained by applying the same discretization methodology to a transformed version of the original system. It is demonstrated that the resulting linear systems satisfy several desirable physical properties of the continuum theory and agree spectrally with the continuum the operator. To date, these are the first discretizations to accomplish these goals without extending the theory to a costly extra dimension.

Finally, it is shown that the resulting linear systems are amenable to effective preconditioning by algebraic multigrid methods. Specifically, classical algebraic multigrid (AMG) and adaptive smoothed aggregation (αSA) multigrid are employed. The result is a solution process that is efficient and scalable as both the lattice size and the disorder of the background field is increased, and the simulated fermion mass is decreased.

Dedication

To Mom and Dad.

Acknowledgements

This work was influenced by a great number of people. First and foremost, I want to thank my advisors, Tom Manteuffel and Steve McCormick. From them I learned how to be a better researcher, teacher, and professional. They taught me that sometimes you have to play hard before you can work hard, and that often times the best mathematics is done on a bicycle, a mountain top, or a ski lift. This work would not have been possible without the help of John Ruge and Marian Brezina. Their patience with my programming ineptitude was the stuff of legends. In addition, I must thank the entire Grandview Gang: James Adler, Eunjung Lee, Josh Nolting, Minho Park, Geoff Sanders, and Lei Tang. Without them, the pun-ishment of Tuesday marathon meetings would surely have been unbearable. I also need to thank every physicist that I ever bugged with endless conceptual questions. James Brannick, Rich Brower, Mike Clark, Tom DeGrand, Anna Hasenfratz, Claudio Rebbi, and Pavlos Vranas each provided lifesaving insight at some time in the course of this work. I must also thank Markus Berndt and Dave Moulton of Los Alamos National Laboratory for putting up with me for three summers. Many of the ideas in this thesis were born in the sweltering heat of a cubicle at 7300 feet. Finally, I want to thank my family and the rest of my Boulder friends for their endless support during the last four years. They kept me mostly sane.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

Introduction

The numerical solution of the Dirac equation is the main computational bottleneck in the numerical simulation of both quantum electrodynamics (QED) and quantum chromodynamics (QCD), both of which are part of the Standard Model of particle physics [30]. In general, the Dirac equation describes the interaction of spin-$\frac{1}{2}$ particles, or *fermions*, and the particles that carry force between them, or *bosons*. QED describes the electroweak interactions between *electrons* and their force carrying *photons*. QCD describes the strong interaction between *quarks* and their force carrying *gluons*. The dimension and complexity of the formulation of the Dirac equation depends on the specific theory that it describes [32]. Compared to QED, QCD is an extremely complex theory. As such, it is common practice to develop new methods first for QED, and then latter extend them to QCD. In this thesis, we focus on a simplified model of QED known as the *Schwinger model*.

The primary purpose of any numerical simulation of QED (or QCD) is to verify the validity of the theory by comparing numerical predictions to experiments. Values of physical observables, like particle mass and momenta, are computed via Monte Carlo methods and compared to like quantities measured in particle accelerator experiments [29]. The *vast* majority of computation time in such simulations is spent inverting the discrete Dirac operator. As such, it is of utmost necessity to develop efficient numerical solvers for the solution of such systems. In traditional discretizations of the

Dirac equation the resulting matrix operator is large, sparse, and highly structured, but has random coefficients and is extremely ill-conditioned. The two main parameters of interest are the temperature (β) of the background gauge (boson) field and the fermion mass (m). For small values of β (< 10), the entries in the Dirac matrix become highly disordered. Moreover, as the fermion mass approaches its true physical value, performance of the community standard Krylov solvers degrades - a phenomenon known as *critical slowing down* [14]. As a result, the development of sophisticated preconditioners for the solution process has been a priority in the physics community for some time. The use of multilevel methods as preconditioners for solution of the Dirac equation were first used in the 1990's [22]. Though greatly improving solver performance, they did not successfully eliminate critical slowing down. Recently, adaptive multilevel preconditioners such as *adaptive smoothed aggregation multigrid* (αSA) have proved capable of eliminating critical slowing down [13], [14].

In addition to developing fast solvers for traditional discretizations of the Dirac operator, it is also necessary to consider alternate discretizations of the governing equations that, while capturing the physical properties of the continuum system, also lend themselves to efficient solution by iterative methods. The vast majority of popular discretizations utilize finite difference techniques. Due to the first-order nature of the operator, the naive application of finite difference techniques results in a problem that the physics community refers to as *species doubling* [29]. In the applied mathematics community, this phenomenon is known as red-black instability. As such, modifications of simple finite difference discretizations are necessary to avoid this issue. In the popular Dirac-Wilson discretization the problem of species doubling is remedied by adding artificial diffusion to the main diagonal of the operator [52]. In [45], a nonlocal approximation to the continuum normal equations is formulated using finite differences. Recently, new methods have been have been developed, based on finite-differences, that retain many important physical properties. However, these methods require formulating the prob-

lem in five dimensions, greatly increasing the computational cost needed to simulate the theory [35], [38], [48]. The primary purpose of this thesis is to introduce consistent discretization of the simplified Schwinger model using least-squares finite elements.

Finite element discretizations have been largely overlooked in lattice gauge theory. Attempts were made in the 1990's to employ finite element methods but were quickly abandoned, primarily because typical Galerkin-like formulations fail to avoid the problem of species doubling. In [31], the continuum equations are expanded in an infinite set of Bloch wave functions and an approximation is obtained by restriction to the lowest mode wave functions, which are very similar finite element basis functions. The primary focus of this dissertation is the development of an alternate discretization of the Dirac equation using least-squares finite elements. This formulation leads to a discrete system that is consistent with the physical properties of the continuum governing equations, avoids the problem of species doubling, and is amenable to solution via adaptive multilevel methods. It is the first formulation to date that accomplishes all of these tasks without extending the model to a costly fifth dimension.

The layout of the thesis is as follows. In Chapter 2, we introduce the general continuum Dirac operator. We discuss the formulations of the model for QED, QCD, and the simplified Schwinger model. We also discuss an alternative formulation of the Schwinger model that will prove useful in our analysis of discretization methods. Finally, we discuss various physical properties that the continuum operator satisfies, including *gauge covariance* and *chiral symmetry*. In Chapter 3, we discuss various discrete approximations to the continuum model. We introduce the method of *covariant finite differences* and traditional discretizations that result from their use, including the *Naive* discretization and *Wilson's* discretization. We formulate discrete analogues of gauge covariance and chiral symmetry, as well as discuss the concept of species doubling. Finally, we discuss the numerical advantages and disadvantages of building simulations around these discretizations. Chapters 2 and 3 are intended to serve as a stand-alone

introduction to discretizations of the Dirac operator for the applied mathematician with little to no background in lattice gauge theory.

Chapter 4 introduces the general least-squares finite element discretization that is the primary method used in this thesis. We discuss the method applied to a general first-order system of PDEs. In Chapter 5, we apply the least-squares finite element methodology to the Schwinger model of QED. A gauge covariant solution process is obtained directly from the standard formulation of the governing equations by applying a method known as *gauge fixing*. It is then demonstrated that the resulting linear system satisfies chiral symmetry and does not suffer from species doubling. The spectrum of the resulting discrete operator is then compared to that of the continuum operator. Finally, it is demonstrated that the least-squares functional satisfies an H^1-ellipticity property, and implications of this are discussed. In the remainder of the chapter, the application of an algebraic multigrid as a preconditioner for the solution process is investigated. It is demonstrated that the resulting linear system of equations can be solved efficiently using such a multilevel method. Finally, we compare the performance of two algebraic multigrid methods applied to the discrete least-squares system and a traditional discretization based on covariant finite differences. It is shown that the former can be solved roughly twice as fast as the latter.

In Chapter 6, the least-squares methodology is applied to an alternate formulation of the governing equations that employs a transformation based on a Helmholtz decomposition of the gauge field. This effectively removes the gauge field from the differential operators and yields a linear system similar to a diffusion equation with variable coefficients. It is shown that the resulting linear system satisfies gauge covariance and chiral symmetry, and does not suffer from species doubling. Next, the H^1-ellipticity of the least-squares functional is established. Finally, the use of two algebraic multigrid methods are investigated as preconditioners in the solution process. These methods are shown to be *very* effective at solving the resulting system of equations.

Finally, in Chapter 7, we make concluding remarks and discuss future directions of the project.

Chapter 2

The Continuum Model

In this chapter, detailed background of the Dirac operator is presented. First, the continuum Dirac operator is introduced together with several properties that it must satisfy, including gauge covariance and chiral symmetry. Next, the discrete Dirac operator and the discrete analogues of the previously mentioned physical properties are discussed. To motivate this, two traditional discretizations of the Dirac operator are considered: the so-called *naive* discretization and *Wilson's* discretization. The properties of gauge covariance and chiral symmetry are discussed in terms of the two resulting discrete operators and, in the process, the curious phenomenon of *species doubling* is introduced.

2.1 Continuum Dirac Operator

The Dirac equation is the relativistic analogue of the Schrödinger equation [32]. It describes the interaction between spin-$\frac{1}{2}$ particles, called *fermions*, and the particles that carry force between them, aptly termed *force carriers*. Depending on the specific gauge theory, the operator can take on several forms, the most general of which is given by

$$\mathcal{D}\psi \quad = \quad \sum_{\mu=1}^{d} \gamma_\mu \otimes \left(\partial_\mu I - i\mathcal{A}_\mu\right)\psi + m\psi. \tag{2.1}$$

Here, d is the problem dimension, γ_μ is a set of matrix coefficients, ∂_μ is the usual partial derivative in the x_μ direction, m is the particle mass, and $\mathcal{A}_\mu(x)$ is a matrix operator describing the gauge field. In quantum electrodynamics (QED) and quantum chromodynamics (QCD) we wish to describe particles of different spins and colors. The dimensions of matrices γ_μ and $\mathcal{A}_\mu$ depend on the number of spins in the theory, n_s, and the number of colors in the theory, n_c. Specifically, each γ_μ is $n_s \times n_s$ and $\mathcal{A}_\mu$ is $n_c \times n_c$. Operator $\mathcal{D}$ acts on $\psi : \mathbb{R}^d \mapsto \mathbb{C}^{n_s} \otimes \mathbb{C}^{n_c}$, a tensor field (multicomponent wavefunction) describing the particle. It is common to refer to the inverse Dirac operator, $\mathcal{D}^{-1}$, as the *fermion propagator*. We introduce the shorthand notation $\nabla_\mu = \partial_\mu I - i\mathcal{A}_\mu$ for the μ^{th} *covariant derivative*. Occasionally, we wish to explicitly indicate that operator $\mathcal{D}$, and its propagator, depend on gauge field $\mathcal{A}$. Thus, we denote the Dirac operator by $\mathcal{D}(\mathcal{A})$ and its propagator by $\mathcal{D}^{-1}(\mathcal{A})$.

Quantum electrodynamics, with which this thesis is most concerned, is the study of the interaction of electrically charged fermions, electrons, and their force carriers, photons. In the full physical model of QED, particles can have one of four different spins, so $n_s = 4$. The term *spin* here is slightly ambiguous. In addition to a particle having a specific angular momentum, either spin-up or spin-down, it also has a specific energy, either positive or negative. The energies here distinguish between particles and *anti-particles*. The anti-particle of the electron is the *positron*. The two possible spins and energies then lead to four possible types of particles: spin-up electrons, spin-down electrons, spin-up positrons, and spin-down positrons. The concept of color is specific to QCD, so, in the case of QED, $n_c = 1$.

The set of matrix coefficients, γ_μ, do not have a definite form. They must simply form a basis for the set of $n_s \times n_s$ unitary, anti-commuting matrices. A Traditional choice are the so-called Dirac matrices:

$$\gamma_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \gamma_2 = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix},$$

$$\gamma_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad \gamma_4 = \begin{bmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{bmatrix}.$$

The full physical theory has four dimensions (one temporal and three spatial). With the above given choice for γ_μ, the Dirac operator becomes

$$\begin{bmatrix} mI & 0 & \nabla_3 - i\nabla_4 & \nabla_1 - i\nabla_2 \\ 0 & mI & \nabla_1 + i\nabla_2 & -\nabla_3 - i\nabla_4 \\ \nabla_3 + i\nabla_4 & \nabla_1 - i\nabla_2 & mI & 0 \\ \nabla_1 + i\nabla_2 & -\nabla_3 + i\nabla_4 & 0 & mI \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix}, \quad (2.2)$$

where ψ_s and f_s, $s = 1, \ldots, 4$, correspond to the s^{th} spin component of ψ and f, respectively. With $n_c = 1$, $\partial_\mu I$ and $\mathcal{A}_\mu$ become scalar operators. The gauge field, $\mathcal{A}$, has four components, $\mathcal{A}_\mu(x) \in \mathbb{R}$, each representing the photons in one of the four directions.

Quantum chromodynamics describes the interaction of color charged particles. The fermion in this case is a *quark*. The carriers of the color-force are the *gluons*. Color charged particles can be red, blue, or green. Like QED, particles come in four different spins (spin-up quark, spin-down quark, spin-up anti-quark, and spin-down anti-quark). However, each of these spin components can be a specific color as well, leading to twelve distinct types of particles. (Note that color here does not refer to the actual visible

color of the particle: it is simply an arbitrary construct of the theory.) Again, like in QED, the theory is typically represented in four dimensions. The formulation of the Dirac equation in quantum chromodynamics is structurally similar to (2.2), but the dimensions of ∇_μ must be altered to account for the three colors of particles. Because the gauge field must describe the interaction of particles of three different colors, we have $\mathcal{A}_\mu \in \mathrm{su}(3)$, the space of 3×3, traceless, Hermitian matrices. The μ^{th} covariant derivative operator acting on the s^{th} spin component of ψ appears as

$$\nabla_\mu \psi_s = \left(\begin{bmatrix} \partial_{x_\mu} & 0 & 0 \\ 0 & \partial_{x_\mu} & 0 \\ 0 & 0 & \partial_{x_\mu} \end{bmatrix} - i\mathcal{A}_\mu \right) \begin{bmatrix} \psi_{s,r} \\ \psi_{s,g} \\ \psi_{s,b} \end{bmatrix}, \tag{2.3}$$

where r, g, and b indicate the color of the particle. The QCD form of the Dirac equation appears exactly as in (2.2), but instead of ψ_s being a scalar quantity, it is a three-component wavefunction, with each component describing a particle of a certain color.

Describing ψ as a wavefunction necessarily indicates a relationship between ψ and a probability amplitude. That is, suppose that p represents, for instance, the state of a quark being spin-up, having positive energy, and being green. Then

$$\int_V |\psi_p|^2 dV, \tag{2.4}$$

is the probability that the particle in question is spin-up, has positive energy, is green, and can be found in the spatial region V [32].

Finally, it is useful to make some observations about the spectral properties of $\mathcal{D}$. Since we are discretizing $\mathcal{D}$, it is desirable that the spectrum of the resulting discrete operator has a spectrum at least somewhat similar to that of the continuum operator.

Note that, in both the QED and QCD cases, the covariant derivative operators are anti-Hermitian. That is,

$$\nabla_\mu^* = -\nabla_\mu, \tag{2.5}$$

where $\mathcal{L}^*$ represents the formal adjoint of differential operator $\mathcal{L}$. Then, it is easy to see that (2.2) can be written as

$$\mathcal{D} = \begin{bmatrix} mI & \mathcal{B} \\ -\mathcal{B}^* & mI \end{bmatrix}, \tag{2.6}$$

where $\mathcal{B} : \mathbb{R}^d \mapsto \mathbb{C}^{n_s/2} \otimes \mathbb{C}^{n_c}$. Further, $\mathcal{D}$ can be decomposed into a sum of Hermitian and anti-Hermitian matrices, according to

$$\mathcal{D} = \begin{bmatrix} mI & 0 \\ 0 & mI \end{bmatrix} + \begin{bmatrix} 0 & \mathcal{B} \\ -\mathcal{B}^* & 0 \end{bmatrix}. \tag{2.7}$$

It then follows that the eigenvalues of $\mathcal{D}$ lie on a vertical line in the complex plain, intersecting the real axis at m. That is,

$$\Sigma\left(\mathcal{D}\right) = m + is, \tag{2.8}$$

where $m, s \in \mathbb{R}$. Then, the eigenvalues of the fermion propagator, $\mathcal{D}^{-1}$, are given by

$$\Sigma\left(\mathcal{D}^{-1}\right) = \frac{1}{m + is}, \tag{2.9}$$

where, again, $s \in \mathbb{R}$. That is, the eigenvalues of the propagator lie on a circle in the complex plane with radius $\frac{1}{2m}$, and centered on the real axis at $\frac{1}{2m}$.

2.2 The Schwinger Model

The operator $\mathcal{D}$, in both QED and QCD, is complicated, to say the least. In four dimensions, with four spins, and three colors, the size of any discrete representation can get intractably large. As such, it is common to work with a simplified systems rather than the full physical models [14]. The so-called Schwinger Model is a two-spin model of QED in two spatial dimensions. Like the full physical model of QED it can be considered a model of the interaction between electrons and photons. Rather than four different types of particles though, it models only two, which we refer to as left- and right-handed. Here, *handedness*, or *helicity*, is a characterization of a particle's angular momentum relative to its direction of motion. For massive particles, left- and right-handed are analogous to spin-up and spin-down particles, respectively. In two dimensions, with $n_s = 2$, γ_μ in (2.1) are replaced by the *Pauli matrices*:

$$\gamma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \gamma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \tag{2.10}$$

Notice that these matrices are unitary and anti-commuting. Naturally, then, the gauge field, $\mathcal{A}$, has two components as well, namely, $\mathcal{A}_1$ and $\mathcal{A}_2$, associated with photons moving in the x- and y-direction, respectively. The wavefunction, ψ, takes the form $\psi = [\psi_{_R}, \psi_{_L}]^t$, where $\psi_{_L}$ and $\psi_{_R}$ represent the left- and right-handed components of the fermion, respectively. Similarly, $f = [f_{_R}, f_{_L}]^t$, where $f_{_L}$ and $f_{_R}$ are the left- and right-handed components of the source term, respectively. For convenience, we associate the $\mu = 1$ direction with x and the $\mu = 2$ direction with y. Substituting these representations into (2.1), we obtain the governing equations of the Schwinger model:

$$\begin{bmatrix} m & \nabla_x - i\nabla_y \\ \nabla_x + i\nabla_y & m \end{bmatrix} \begin{bmatrix} \psi_{_R} \\ \psi_{_L} \end{bmatrix} = \begin{bmatrix} f_{_R} \\ f_{_L} \end{bmatrix}. \tag{2.11}$$

Naturally, the physical objects that are modeled by equations such as these can act on extremely large spatial domains. It is natural then to restrict our attention to a small physical domain and require that the wavefunction, ψ, and the gauge field, $\mathcal{A}$, be periodic on that domain. Let that domain be $\mathcal{R} = [0,1] \times [0,1]$. Then, let $\mathcal{V}_{\mathbb{R}}$ be some space of real-valued, periodic functions on $\mathcal{R}$, and $\mathcal{V}_{\mathbb{C}}$ be some space of complex-valued, periodic functions on $\mathcal{R}$. The specific characteristics of spaces $\mathcal{V}_{\mathbb{R}}$ and $\mathcal{V}_{\mathbb{C}}$ are discussed in a later chapter. Let $\psi(x,y) = [\psi_R(x,y), \psi_L(x,y)]^t \in \mathcal{V}_{\mathbb{C}}^2$ be the fermion field with right- and left-handed components ψ_R and ψ_L, respectively. Assume that $\mathcal{A}(x,y) = [\mathcal{A}_1(x,y), \mathcal{A}_2(x,y)]^t \in \mathcal{V}_{\mathbb{R}}^2$. With periodic boundary conditions on ψ, the 2D Schwinger model becomes

$$\begin{bmatrix} mI & \nabla_x - i\nabla_y \\ \nabla_x + i\nabla_y & mI \end{bmatrix} \begin{bmatrix} \psi_R \\ \psi_L \end{bmatrix} = \begin{bmatrix} f_R \\ f_L \end{bmatrix} \text{ in } \mathcal{R}, \tag{2.12}$$

$$\begin{aligned} \psi(0,y) &= \psi(1,y) \quad \forall y \in (0,1), \\ \psi(x,0) &= \psi(x,1) \quad \forall x \in (0,1), \end{aligned}$$

Note that the Schwinger model *is* a specific version of the Dirac equation. As such, in the remainder of this thesis, the two terms are used interchangeably. If, at some time, the full physical version of the Dirac operator is discussed, it will be made clear from context that we are speaking of something *other* than the Schwinger model.

2.2.1 Gauge Covariance

In any gauge theory, like QED or QCD, several physical symmetries of the system must be captured by the Dirac operator. This section is concerned with the manner in which the Dirac operator transforms under both local and global modifications of the fermion field. The first such symmetry discussed is a local gauge symmetry. In this case, the fermion propagators, $\mathcal{D}^{-1}$, must transform *covariantly* under local gauge

transformations [29]. A local gauge transformations, denoted by $\Omega(x,y)$, is a member of the gauge group of the theory. In QED, $\Omega(x,y) \in \mathrm{U}(1)$, the set of complex scalars with unit magnitude. In QCD, $\Omega(x,y) \in \mathrm{SU}(3)$, the set of complex, unitary, 3×3 matrices with determinant one. The transformation is *local* because it depends on x and y.

Definition 2.2.1. *Suppose a local gauge transformation, $\Omega(x,y)$, is applied to each component of the fermion field ξ. Then, a modified propagator, $\widetilde{\mathcal{D}}^{-1}$, must exist such that*

$$\mathcal{D}^{-1}\left[\Omega(x,y)\, I_{n_s}\right]\xi \;=\; \left[\Omega(x,y)\, I_{n_s}\right]\widetilde{\mathcal{D}}^{-1}\xi, \tag{2.13}$$

where I_{n_s} is the $n_s \times n_s$ identity operator, and transformation $\Omega(x,y)$ is a member of the gauge group.

To find the appropriate form of $\widetilde{\mathcal{D}}$, we consider the form of the Schwinger model appearing in (2.12). In the Schwinger case, the condition for the gauge covariance of $\mathcal{D}^{-1}$ can be written as

$$\mathcal{D}^{-1}\begin{bmatrix} \Omega(x,y) & 0 \\ 0 & \Omega(x,y) \end{bmatrix}\xi \;=\; \begin{bmatrix} \Omega(x,y) & 0 \\ 0 & \Omega(x,y) \end{bmatrix}\widetilde{\mathcal{D}}^{-1}\xi. \tag{2.14}$$

Let $\Omega(x,y) = e^{i\omega(x,y)}$ for some real, periodic function ω. Suppose further that $\mathcal{D}$ is constructed using the gauge field, $\mathcal{A} = [\mathcal{A}_1, \mathcal{A}_2]^t$. We make the ansatz that the modified Dirac operator, $\widetilde{\mathcal{D}}$, has the same form as the original operator, but is built using a modified gauge field, $\widetilde{\mathcal{A}} = \left[\widetilde{\mathcal{A}}_1, \widetilde{\mathcal{A}}_2\right]^t$. For simplicity, write $\widetilde{\mathcal{D}}$ as

$$\widetilde{\mathcal{D}} \;=\; \gamma_1 \otimes \widetilde{\nabla}_x + \gamma_2 \otimes \widetilde{\nabla}_y + mI,$$

where $\widetilde{\nabla}_x = \partial_x - i\widetilde{\mathcal{A}}_1$ and $\widetilde{\nabla}_y = \partial_y - i\widetilde{\mathcal{A}}_2$. Equating the right-hand side of (2.14) with some fermion field, ζ, yields

$$\Omega(x,y)\left[\gamma_1\otimes\widetilde{\nabla}_x+\gamma_2\otimes\widetilde{\nabla}_y+mI\right]^{-1}\xi \;=\; \zeta.$$

Then,

$$\begin{aligned}
\xi &= \left[\gamma_1\otimes\widetilde{\nabla}_x+\gamma_2\otimes\widetilde{\nabla}_y+mI\right]\Omega^*(x,y)\,\zeta \\
&= \gamma_1\otimes\widetilde{\nabla}_x\left(e^{-i\omega}\zeta\right)+\gamma_2\otimes\widetilde{\nabla}_y\left(e^{-i\omega}\zeta\right)+mI\left(e^{-i\omega}\zeta\right) \\
&= \gamma_1\otimes\left(\partial_x-i\widetilde{\mathcal{A}}_1\right)\left(e^{-i\omega}\zeta\right)+\gamma_2\otimes\left(\partial_y-i\widetilde{\mathcal{A}}_2\right)\left(e^{-i\omega}\zeta\right)+mI\left(e^{-i\omega}\right)\zeta \\
&= \Omega^*(x,y)\left[\gamma_1\otimes\left(\partial_x-i\{\widetilde{\mathcal{A}}_1+\omega_x\}\right)+\gamma_2\otimes\left(\partial_y-i\{\widetilde{\mathcal{A}}_2+\omega_y\}\right)+mI\right]\zeta,
\end{aligned}$$

where $\omega_x=\partial_x\omega$ and $\omega_y=\partial_y\omega$. Thus,

$$\left[\gamma_1\otimes\left(\partial_x-i\{\widetilde{\mathcal{A}}_1+\omega_x\}\right)+\gamma_2\otimes\left(\partial_y-i\{\widetilde{\mathcal{A}}_2+\omega_y\}\right)+mI\right]^{-1}\Omega(x,y)\,\xi \;=\; \zeta.$$

Then, equating ζ with the left-hand side of (2.14), it is clear that gauge covariance of the propagator is obtained precisely when $\widetilde{\mathcal{D}}$ is constructed using the modified gauge field

$$\widetilde{\mathcal{A}} \;=\; \mathcal{A}-\nabla\omega. \tag{2.15}$$

A consequence of the gauge covariance of the propagator is that, in solving the equation $\mathcal{D}(\mathcal{A})\,\psi=f$, we are not restricted to working specifically with $\mathcal{D}(\mathcal{A})$. In fact, for *any* transformation of the form $\Omega(x,y)=e^{i\omega(x,y)}\in\mathrm{U}(1)$,

$$\psi \;=\; \begin{bmatrix}\Omega(x,y) & 0\\ 0 & \Omega(x,y)\end{bmatrix}\mathcal{D}^{-1}(\mathcal{A}-\nabla\omega)\begin{bmatrix}\Omega^*(x,y) & 0\\ 0 & \Omega^*(x,y)\end{bmatrix}f. \tag{2.16}$$

Then, it is possible to solve the original problem for ψ by first applying the *inverse* transform to the source term, f, applying the transformed propagator, $\mathcal{D}^{-1}(\mathcal{A} - \nabla\omega)$, and then applying the transform to the result.

An interesting physical implication for the property of gauge covariance is more clearly explained in the context of QCD. The application of a gauge transformation to a fermion field, ψ, can be viewed as a change in the color reference frame. A trivial example would be if the roles of blue and red particles were interchanged in the model. Since the gauge transformation depends on the location in the domain, it is also possible to, for instance, interchange the role of blue and red particles in the first quadrant of the domain, interchange the role of red and green particles in the second quadrant, and leave the remainder of the domain unaffected. Let $\mathcal{C}$ denote the original color reference frame, and $\widetilde{\mathcal{C}}$ denote the modified reference frame described above. Then, in accordance with (2.16), given the problem $\mathcal{D}(\mathcal{A})\psi = f$ in reference frame $\mathcal{C}$, we can solve for ψ by transforming the source data into reference frame $\widetilde{\mathcal{C}}$, inverting the analogous Dirac operator there and, then, transforming the result back to reference frame $\mathcal{C}$.

This property illuminates a fundamental relationship between gauge fields, $\mathcal{A}$ and $\widetilde{\mathcal{A}}$, which differ only by the gradient of a periodic function. Essentially, any computation that can be done with $\mathcal{A}$ could, in fact, be done with $\widetilde{\mathcal{A}}$ instead. Such pairs of gauge fields are said to be in the same *equivalence class*.

Definition 2.2.2. *Gauge fields $\mathcal{A}$ and $\widetilde{\mathcal{A}}$ are said to be in the same equivalence class if there exists some differentiable periodic function, $\omega(x, y)$, such that*

$$\widetilde{\mathcal{A}} = \mathcal{A} - \nabla\omega.$$

Gauge covariance is perhaps the most crucial property in a theory such as QED. The Monte Carlo simulation at the heart of these simulations is, in fact, an approxi-

mation of an infinite-dimensional Feynman path integral [27]. In such an integral, it is necessary to integrate over all possible gauge fields. With the property of gauge covariance of the propagator intact, the problem is reduced to integrating over all possible gauge field *equivalence classes* instead. This clearly reduces the dimensionality the continuum problem, as well as any discrete approximation of the process.

2.2.2 Chiral Symmetry

The second physical property that should be conserved is *chiral symmetry*. In the broadest sense, chiral symmetry is the global symmetry property that independent transformations of the right- and left-handed fields do not change the physics of the model in the massless case [29]. This property is manifested mathematically by the property that, when $m = 0$, the inner product $\langle \xi, \gamma_1 \mathcal{D} \xi \rangle$ remains invariant under transformations of the form $\xi \mapsto \Lambda \xi$, where $\langle \cdot\,, \cdot \rangle$ is the usual L^2 inner product, and

$$\Lambda \;=\; \begin{bmatrix} e^{i\lambda_{_R}} & 0 \\ 0 & e^{i\lambda_{_L}} \end{bmatrix}, \tag{2.17}$$

for $\lambda_{_R}$, $\lambda_{_L} \in \mathbb{R}$.

Definition 2.2.3. *Given $\lambda_{_R}$, $\lambda_{_L} \in \mathbb{R}$, and transformation Λ defined by*

$$\Lambda \;=\; \begin{bmatrix} e^{i\lambda_{_R}} & 0 \\ 0 & e^{i\lambda_{_L}} \end{bmatrix}, \tag{2.18}$$

operator $\mathcal{D}$ satisfies a chiral symmetry if, for $m = 0$,

$$\langle \Lambda\xi, \gamma_1 \mathcal{D} \Lambda \xi \rangle \;=\; \langle \xi, \gamma_1 \mathcal{D} \xi \rangle \,. \tag{2.19}$$

where $\langle \cdot\,, \cdot \rangle$ is the usual L^2 inner product.

It is important to note the difference between the requirements of chiral symmetry and that of gauge covariance. First, chiral symmetry is a *global* symmetry, indicated by the fact that $\lambda_{_R}$ and $\lambda_{_L}$ in Λ do not have spatial dependence. All right- and left-handed fields are rotated by the same transformation at each point. Second, $\mathcal{D}$ cannot be altered to make (2.19) hold.

To see that the continuum Dirac operator of the Schwinger model satisfies chiral symmetry, set $m = 0$ in (2.11). Notice that the elements of Λ are constants. Then,

$$\begin{aligned}
\gamma_1\mathcal{D}\Lambda &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & \nabla_x - i\nabla_y \\ \nabla_x + i\nabla_y & 0 \end{bmatrix} \begin{bmatrix} e^{i\lambda_R} & 0 \\ 0 & e^{i\lambda_L} \end{bmatrix}, \\
&= \begin{bmatrix} \nabla_x - i\nabla_y & 0 \\ 0 & \nabla_x + i\nabla_y \end{bmatrix} \begin{bmatrix} e^{i\lambda_R} & 0 \\ 0 & e^{i\lambda_L} \end{bmatrix}, \\
&= \begin{bmatrix} e^{i\lambda_R} & 0 \\ 0 & e^{i\lambda_L} \end{bmatrix} \begin{bmatrix} \nabla_x - i\nabla_y & 0 \\ 0 & \nabla_x + i\nabla_y \end{bmatrix}. \\
&= \Lambda\gamma_1\mathcal{D}
\end{aligned}$$

Thus,

$$\begin{aligned}
\langle \Lambda\xi, \gamma_1\mathcal{D}\Lambda\xi \rangle &= \langle \Lambda\xi, \Lambda\gamma_1\mathcal{D}\xi \rangle, \\
&= \langle \Lambda^*\Lambda\xi, \gamma_1\mathcal{D}\xi \rangle, \\
&= \langle \xi, \gamma_1\mathcal{D}\xi \rangle,
\end{aligned}$$

as desired. Note that the chiral symmetry of operator $\mathcal{D}$ relies upon the commutativity of $\gamma_1\mathcal{D}$ and transformation Λ. Notice also that this commutativity is made possible by the fact that, in the massless case, operator $\mathcal{D}$ has zeros on its main diagonal.

2.2.3 An Alternate Formulation

For simplicity, denote the off-diagonal block of (2.12) by $\mathcal{B}(\mathcal{A}) = \nabla_x - i\nabla_y = (\partial_x - i\mathcal{A}_1) - i(\partial_y - i\mathcal{A}_2)$. The matrix form of the Dirac equation then becomes

$$\begin{bmatrix} mI & \mathcal{B} \\ -\mathcal{B}^* & mI \end{bmatrix} \begin{bmatrix} \psi_R \\ \psi_L \end{bmatrix} = \begin{bmatrix} f_R \\ f_L \end{bmatrix}. \tag{2.20}$$

First, note that operator $\mathcal{B}$ transforms covariantly under a transformation of the form e^z, where z is any complex-valued, periodic function. That is to say, if some component wavefunction $\xi \in \mathcal{V}$ is transformed according to $\xi \mapsto e^z\xi$, then it is possible to specify some modified operator $\widetilde{\mathcal{B}}$ such that

$$\mathcal{B}e^z\xi = e^z\widetilde{\mathcal{B}}\xi.$$

To see that $\mathcal{B}$ transforms covariantly under such a transformation, let $r(x,y)$ and $s(x,y)$ be real, periodic functions, and set $z = r + is$. Then

$$\begin{aligned} \mathcal{B}(\mathcal{A})\, e^z\xi &= [(\partial_x - i\mathcal{A}_1) - i(\partial_y - i\mathcal{A}_2)]\, e^z\xi \\ &= e^z\{[\partial_x - i(\mathcal{A}_1 + r_y - s_x)] - i[\partial_y - i(\mathcal{A}_2 - r_x - s_y)]\}\xi \\ &= e^z\mathcal{B}\left(\mathcal{A} - \nabla^\perp r - \nabla s\right)\xi. \end{aligned}$$

Thus, the correct modification of $\mathcal{B}(\mathcal{A})$ corresponding to transformation e^z is $\widetilde{\mathcal{B}} = \mathcal{B}\left(\mathcal{A} - \nabla^\perp r - \nabla s\right)$. Notice that the real part of z appear as a curl-like term in the modified gauge field, and the imaginary part of z appears as a gradient term. Now, suppose that real, periodic functions u and v form a Helmholtz decomposition of the gauge field, according to

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}_1 \\ \mathcal{A}_2 \end{bmatrix} = \nabla^{\perp} u + \nabla v + \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}, \tag{2.21}$$

where k_1 and k_2 are constants. Setting $z = u + iv$ yields

$$\mathcal{B}(\mathcal{A})\, e^{z}\xi = e^{z}\mathcal{B}_k\xi, \tag{2.22}$$

where $\mathcal{B}_k := (\partial_x - ik_1) - i(\partial_y - ik_2)$. In addition, it is easy to verify that the adjoint operator, $\mathcal{B}^*$, transforms covariantly under a similar transformation. Specifically,

$$\mathcal{B}^*(\mathcal{A})\, e^{-\bar{z}}\xi = e^{-\bar{z}}\mathcal{B}_k^*\xi. \tag{2.23}$$

From (2.22) and (2.23), it is easy to see that $\mathcal{B}(\mathcal{A})$ and its adjoint can be written as

$$\mathcal{B}(\mathcal{A}) = e^{z}\mathcal{B}_k e^{-z}, \tag{2.24}$$

$$\mathcal{B}^*(\mathcal{A}) = e^{-\bar{z}}\mathcal{B}_k^* e^{\bar{z}}, \tag{2.25}$$

where $z = u + iv$. This representation gives insight into the nullspace of $\mathcal{B}$. First, consider the transformed operator $\mathcal{B}_k$, with the non-constant portion of the gauge field removed. Let $\phi = e^{i(k_1 x + k_2 y)}$. Then

$$\begin{aligned} \mathcal{B}_k\phi &= [(\partial_x - ik_1) - i(\partial_y - ik_2)]\, e^{i(k_1 x + k_2 y)}, \\ &= [(ik_1 - ik_1) - i(ik_2 - ik_2)]\, e^{i(k_1 x + k_2 y)}, \\ &= 0. \end{aligned}$$

But, recall that operator $\mathcal{B}_k$ acts on complex-valued periodic functions, and that

$$e^{i(k_1x+k_2y)} = [\cos(k_1x) + i\sin(k_1x)][\cos(k_2y) + i\sin(k_2y)].$$

Then, $\phi = e^{i(k_1x+k_2y)}$ is periodic on $\mathcal{R}$ only if $k_1 = 2\pi l_1$ and $k_2 = 2\pi l_2$ for some $l_1, l_2 \in \mathbb{Z}$. Thus, it is easy to see from (2.24) that operator $\mathcal{B}(\mathcal{A})$ is singular, with nullspace vector $\phi = e^{z+i(k_1x+k_2y)}$, only if k_1 and k_2 are integer multiples of 2π. Similarly, from (2.25) it is clear that, under these conditions on $\underline{k}$, $\mathcal{B}^*(\mathcal{A})$ is singular with nullspace vector $\phi = e^{-\bar{z}+i(k_1x+k_2y)}$ [36].

Next, notice that (2.20) can be reformulated as

$$\begin{bmatrix} mI & e^{z}\mathcal{B}_k e^{-z} \\ -e^{-\bar{z}}\mathcal{B}_k^* e^{\bar{z}} & mI \end{bmatrix} \begin{bmatrix} \psi_R \\ \psi_L \end{bmatrix} = \begin{bmatrix} f_R \\ f_L \end{bmatrix}. \tag{2.26}$$

Then, if the constant portions of the gauge field are integer multiples of 2π, $\mathcal{D}$ has two eigenvectors, ϕ_+ and ϕ_-, associated with purely real eigenvalue m, where

$$\phi_+ = \begin{bmatrix} e^{-\bar{z}+i(k_1x+k_2y)} \\ e^{z+i(k_1x+k_2y)} \end{bmatrix}, \tag{2.27}$$

$$\phi_- = \begin{bmatrix} e^{-\bar{z}+i(k_1x+k_2y)} \\ -e^{z+i(k_1x+k_2y)} \end{bmatrix}. \tag{2.28}$$

We refer to a gauge field with these properties as an exceptional configuration.

Definition 2.2.4. *Let gauge field $\mathcal{A}$ have the Helmholtz decomposition given in (2.21). Gauge field $\mathcal{A}$ is an exceptional configuration if constants k_1 and k_2 are integer multiples of 2π.*

Furthermore, if $\mathcal{A}$ is an exceptional configuration, the massless Dirac operator, $\mathcal{D}_0$, is singular, with nullspace vectors in the span of ϕ_+ and ϕ_-. This clearly has implications for any discretization of $\mathcal{D}$, because any matrix approximation, $\mathbb{D}$, that shares this

spectral characteristic will become increasingly ill-conditioned as $m \to 0$. However, if either k_1 or k_2 is *not* an integer multiple of 2π, $\mathcal{B}$ will not be singular and, by extension, $\mathcal{D}$ will not have any purely real eigenvalues.

Chapter 3

The Discrete Model

In numerical simulations of QED, many solutions of the discrete Dirac equation must be computed for varying gauge fields and source vectors. Solutions of systems of this type are needed both for computing observables and for generating gauge fields with the correct probabilistic characteristics [13]. In traditional lattice formulations, the continuum domain, $\mathcal{R}$, is replaced by an $N \times N$ regular, periodic lattice. The continuum wavefunction, ψ, and source, f, are replaced by periodic discrete analogues, $\underline{\psi}$ and $\underline{f}$, with values specified only at the lattice *sites*. The continuum gauge field, $\mathcal{A}$, is represented by the periodic discrete field $\underline{A} = [\underline{A}_1, \underline{A}_2]^t$, with information specified on each of the lattice *links*. The components of the gauge field, $\underline{A}_1$ and $\underline{A}_2$, represent values on the horizontal and vertical lattice links, respectively. A discrete solution process of the 2D Schwinger model then takes the source, $\underline{f}$, specified at the lattice sites and gauge field, $\underline{A}$, specified at the lattice links, and returns the discrete fermion field, $\underline{\psi}$, with values again specified at the lattice sites. The discrete solution can be written as

$$\underline{\psi} \;=\; [\mathbb{D}\,(\underline{A})]^{-1}\,\underline{f},$$

where $\mathbb{D}$ is some discrete approximation of the continuum Dirac operator.

For completeness, let $\mathcal{N}_{\mathbb{C}}$ be the space of discrete complex-valued vectors with values associated with the sites on the lattice. Let $\mathcal{N}_{\mathbb{R}} \subset \mathcal{N}_{\mathbb{C}}$ be the space of discrete

real-valued vectors, with values associated with the lattice sites. Then, the discrete fermion field is given by $\underline{\psi} = [\underline{\psi}_R, \underline{\psi}_L]^t \in \mathcal{N}_{\mathbb{C}}^2$, which specifies complex values of both the right- and left-handed components of the fermion field at each lattice site. Similarly, $\underline{f} = [\underline{f}_R, \underline{f}_L]^t \in \mathcal{N}_{\mathbb{C}}^2$. Let $\mathcal{E}$ be the space of discrete real-valued vectors with values associated with the lattice links. Then $\underline{A} = [\underline{A}_1, \underline{A}_2]^t \in \mathcal{E}$.

3.1 The Naive Discretization

Traditional discretizations of the Dirac equation are based on covariant finite differences. As the name suggests, this method produces a discrete operator by applying a finite difference-like approximation of the covariant derivative, ∇_μ. In this chapter, we consider two such discretizations: the *Naive* discretization and the *Wilson* discretization [29]. The former produces the following discrete operator in the Schwinger case:

$$\mathbb{D}_N = \begin{bmatrix} mI & \nabla_x^h - i\nabla_y^h \\ \nabla_x^h + i\nabla_y^h & mI \end{bmatrix}, \tag{3.1}$$

where ∇_x^h and ∇_y^h, acting on the right-hand component of $\underline{\psi}$, have the following centered difference formulas:

$$\nabla_x^h \underline{\psi}_{Rj,k} = \frac{1}{2h}\left(e^{i\underline{\theta}_{j+1/2,k}}\underline{\psi}_{Rj+1,k} - e^{-i\underline{\theta}_{j-1/2,k}}\underline{\psi}_{Rj-1,k}\right), \tag{3.2}$$

$$\nabla_y^h \underline{\psi}_{Rj,k} = \frac{1}{2h}\left(e^{i\underline{\theta}_{j,k+1/2}}\underline{\psi}_{Rj,k+1} - e^{-i\underline{\theta}_{j,k-1/2}}\underline{\psi}_{Rj,k-1}\right). \tag{3.3}$$

The values of $\underline{\theta}$, called *phase factors*, are located at the midpoint of lattice links and relate to the continuum gauge field according to

$$\underline{\theta}_{j+1/2,k} = \int_{x_j}^{x_{j+1}} \mathcal{A}_1(x, y_k)\, dx, \tag{3.4}$$

$$\underline{\theta}_{j,k+1/2} = \int_{y_k}^{y_{k+1}} \mathcal{A}_2(x_j, y)\, dy. \tag{3.5}$$

Note that because $\mathcal{A}$ is a real-valued function, the phase factors are real-valued as well. As a result, the coefficients in the covariant finite differences are members of the gauge group U(1). Denote the collection of phase factors associated with both the horizontal and vertical lattice links by $\underline{\theta}$. Notice then that $\underline{\theta} \in \mathcal{E}$. It can be shown that the covariant finite differences converge to the associated continuum covariant derivatives as the lattice spacing goes to zero. Matrix $\mathbb{D}_N$ can be written in the simplified form

$$\mathbb{D}_N = \begin{bmatrix} mI & \mathbb{B}_N \\ -\mathbb{B}_N^* & mI \end{bmatrix}, \tag{3.6}$$

where $\mathbb{B}_N^*$ is the conjugate transpose of matrix $\mathbb{B}_N$, with stencil

$$\mathbb{B}_N = \frac{1}{2h} \begin{bmatrix} & -ie^{i\underline{\theta}_{j,k+1/2}} & \\ -e^{-i\underline{\theta}_{j-1/2,k}} & 0 & e^{i\underline{\theta}_{j+1/2,k}} \\ & ie^{-i\underline{\theta}_{j,k-1/2}} & \end{bmatrix}. \tag{3.7}$$

Notice that (3.6), like its continuum analogue (2.6), can be decomposed into Hermitian and anti-Hermitian parts:

$$\mathcal{D} = \begin{bmatrix} mI & 0 \\ 0 & mI \end{bmatrix} + \begin{bmatrix} 0 & \mathbb{B}_N \\ -\mathbb{B}_N^* & 0 \end{bmatrix}. \tag{3.8}$$

Thus, as in the continuum case, the eigenvalues of $\mathbb{D}_N$ lie on a vertical line in the complex plain intersecting the real axis at m. That is,

$$\Sigma\left(\mathbb{D}_N\right) = m + is_j, \quad j = 1, \ldots, 2n^2, \tag{3.9}$$

where $s_j \in \mathbb{R}$. The fact that the eigenvalues of the discrete operator lie on the same vertical line in the complex plane as the eigenvalues of the continuum operator is en-

couraging. Unfortunately, we will see later that the actual values of the discrete and continuum eigenvalues do not agree well at all.

3.1.1 Gauge Covariance

As in the continuum case, matrix operator $\mathbb{D}_N$ must satisfy discrete analogues of gauge covariance and chiral symmetry. Recall that, in the continuum, a gauge transformation, Ω, is defined by

$$\Omega(x,y) \;=\; e^{i\omega(x,y)},$$

where $\omega(x,y)$ is a periodic, real-valued function. Let $\underline{\omega}$ be the vector whose entries are the values of $\omega(x,y)$ evaluated at each lattice site. That is,

$$\underline{\omega}_{j,k} \;=\; \omega(x_j, y_k).$$

Then, define a discrete gauge transformation, $\underline{\Omega}_{\underline{\omega}}$, by the $n^2 \times n^2$ diagonal matrix, whose entries are given by

$$\left[\,\underline{\Omega}_{\underline{\omega}}\,\right]_{l,l} \;=\; e^{i\underline{\omega}_{j,k}}, \tag{3.10}$$

and the map between l and (j,k) is defined by the usual lexicographic ordering of the unknowns. Note that $\underline{\Omega}_{\underline{\omega}}$ is a unitary matrix and that each of its entries is itself a member of the gauge group U(1). Also, note that the adjoint, $\underline{\Omega}^*_{\underline{\omega}}$, is simply the diagonal matrix whose entries are the complex conjugates of the diagonal entries of $\underline{\Omega}_{\underline{\omega}}$. Finally, define unitary matrices $\mathbb{T}_{\underline{\omega}}$ and $\mathbb{T}^*_{\underline{\omega}}$ according to

$$\mathbb{T}_{\underline{\omega}} = \begin{bmatrix} \underline{\Omega}_{\underline{\omega}} & 0 \\ 0 & \underline{\Omega}_{\underline{\omega}} \end{bmatrix} \qquad \mathbb{T}^*_{\underline{\omega}} = \begin{bmatrix} \underline{\Omega}^*_{\underline{\omega}} & 0 \\ 0 & \underline{\Omega}^*_{\underline{\omega}} \end{bmatrix}.$$

Recalling Definition 2.2.1, we make the following definition of gauge covariance of the discrete solution process:

Definition 3.1.1. *Suppose a discrete local gauge transformation, $\underline{\Omega}_{\underline{\omega}}$, is applied to each component of the discrete fermion field, $\underline{\xi}$. A discrete solution process is gauge covariant if a modified discrete propagator, $\widetilde{\mathbb{D}}^{-1}$, exists such that*

$$\mathbb{D}^{-1}\mathbb{T}_{\underline{\omega}}\,\underline{\xi} \;=\; \mathbb{T}_{\underline{\omega}}\,\widetilde{\mathbb{D}}^{-1}\underline{\xi}. \tag{3.11}$$

A little algebra shows that (3.11) is equivalent to

$$\mathbb{T}_{\underline{\omega}}\,\widetilde{\mathbb{D}}\underline{\zeta} \;=\; \mathbb{D}\mathbb{T}_{\underline{\omega}}\,\underline{\zeta}, \tag{3.12}$$

for $\underline{\zeta} = \widetilde{\mathbb{D}}^{-1}\underline{\xi}$. In general, to determine the form of $\mathbb{D}$, it is only necessary to consider (3.12) for a single arbitrary row of the equation. We wish to show that the Naive Dirac operator, $\mathbb{D}_N$, satisfies this definition of gauge covariance. Without loss of generality, consider only the j,k^{th} component of $\underline{\zeta}_R$. Again, we make the ansatz that $\mathbb{D}_N$ and $\widetilde{\mathbb{D}}_N$ only differ in their gauge data, $\underline{\theta}$ and $\widetilde{\underline{\theta}}$, respectively. For simplicity of notation, we require that the stencils for $\underline{\zeta}_{Rj,k}$ satisfy

$$\begin{aligned}
&\frac{1}{2h}\begin{bmatrix} & -ie^{i\left(\widetilde{\underline{\theta}}_{j,k+1/2}+\underline{\omega}_{j,k+1}\right)} & \\ -e^{-i\left(\widetilde{\underline{\theta}}_{j-1/2,k}-\underline{\omega}_{j-1,k}\right)} & 0 & e^{i\left(\widetilde{\underline{\theta}}_{j+1/2,k}+\underline{\omega}_{j+1,k}\right)} \\ & ie^{-i\left(\widetilde{\underline{\theta}}_{j,k+1/2}-\underline{\omega}_{j,k-1}\right)} & \end{bmatrix} \\
&= \frac{1}{2h}\begin{bmatrix} & -ie^{i\left(\underline{\theta}_{j,k+1/2}+\underline{\omega}_{j,k}\right)} & \\ -e^{-i\left(\underline{\theta}_{j-1/2,k}-\underline{\omega}_{j,k}\right)} & 0 & e^{i\left(\underline{\theta}_{j+1/2,k}+\underline{\omega}_{j,k}\right)} \\ & ie^{-i\left(\underline{\theta}_{j,k+1/2}-\underline{\omega}_{j,k}\right)} & \end{bmatrix}.
\end{aligned} \tag{3.13}$$

Clearly then, (3.12) is satisfied when

$$\begin{aligned}
\widetilde{\underline{\theta}}_{j+1/2,k} &= \underline{\theta}_{j+1/2,k} - \left(\underline{\omega}_{j+1,k} - \underline{\omega}_{j,k}\right), \\
\widetilde{\underline{\theta}}_{j-1/2,k} &= \underline{\theta}_{j-1/2,k} - \left(\underline{\omega}_{j,k} - \underline{\omega}_{j-1,k}\right), \\
\widetilde{\underline{\theta}}_{j,k+1/2} &= \underline{\theta}_{j,k+1/2} - \left(\underline{\omega}_{j,k+1} - \underline{\omega}_{j,k}\right), \\
\widetilde{\underline{\theta}}_{j,k-1/2} &= \underline{\theta}_{j,k-1/2} - \left(\underline{\omega}_{j,k} - \underline{\omega}_{j,k-1}\right),
\end{aligned}$$

which, in turn, implies that $\mathbb{D}_N$ satisfies Definition 3.1.1. Recall that, in the continuum case, the correct modification of $\mathcal{D}$ required subtracting $\nabla\omega$ from the given gauge field. This is precisely the modification that was necessary in the discrete case. In this instance, a constant multiple of a *discrete* gradient of $\underline{\omega}$ is subtracted from each of the original phase factors.

3.1.2 Chiral Symmetry

Similarly, we require a discrete analogue of Definition 2.2.3 for chiral symmetry. In the discrete case, chiral symmetry requires that, in the massless case, the physics remained unchanged under independent transformation of the left- and right-handed components of the discrete fermion field.

Definition 3.1.2. *Suppose $\lambda_{_R}$, $\lambda_{_L} \in \mathbb{R}$, and a transformation $\underline{\Lambda}$ is defined by*

$$\underline{\Lambda} = \begin{bmatrix} e^{i\lambda_{_R}} I & 0 \\ 0 & e^{i\lambda_{_L}} I \end{bmatrix}, \tag{3.14}$$

where I is the $n^2 \times n^2$ identity matrix. Operator $\mathbb{D}$ satisfies a chiral symmetry if, for $m = 0$,

$$\left\langle \underline{\Lambda}\underline{\xi}, \underline{\Gamma}_1 \mathbb{D} \underline{\Lambda}\underline{\xi} \right\rangle = \left\langle \underline{\xi}, \underline{\Gamma}_1 \mathbb{D} \underline{\xi} \right\rangle, \tag{3.15}$$

where

$$\underline{\Gamma}_1 \;=\; \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}, \tag{3.16}$$

and $\langle \cdot\,, \cdot \rangle$ *is the usual discrete* l^2 *inner product.*

In the following lemma, we verify that the Naive discrete Dirac operator, $\mathbb{D}_N$, satisfies chiral symmetry.

Lemma 3.1.3. *The Naive Dirac operator,* $\mathbb{D}_N$*, satisfies chiral symmetry.*

Proof. Let $\lambda_{_R}$, $\lambda_{_L} \in \mathbb{R}$ and define $\mathbb{D}_N$, $\underline{\Lambda}$, and $\underline{\Gamma}_1$ according to (3.6), (3.14), and (3.16), respectively. Setting $m = 0$, $\mathbb{D}_N$ becomes

$$\mathbb{D}_N \;=\; \begin{bmatrix} 0 & \mathbb{B}_N \\ -\mathbb{B}_N^* & 0 \end{bmatrix}. \tag{3.17}$$

Then,

$$\left\langle \underline{\Lambda}\underline{\xi}, \underline{\Gamma}_1 \mathbb{D}_N \underline{\Lambda}\underline{\xi} \right\rangle \;=\; \left\langle \underline{\Lambda}^* \mathbb{D}_N^* \underline{\Gamma}_1 \underline{\Lambda}\underline{\xi}, \underline{\xi} \right\rangle.$$

Simplifying the long matrix product gives

$$\begin{aligned}
\underline{\Lambda}^* \mathbb{D}_N^* \underline{\Gamma}_1 \underline{\Lambda} &= \begin{bmatrix} e^{-i\lambda_{_R}} I & 0 \\ 0 & e^{-i\lambda_{_L}} I \end{bmatrix} \begin{bmatrix} \mathbb{B}_N & 0 \\ 0 & -\mathbb{B}_N^* \end{bmatrix} \begin{bmatrix} e^{i\lambda_{_R}} I & 0 \\ 0 & e^{i\lambda_{_L}} I \end{bmatrix} \\
&= \begin{bmatrix} e^{-i\lambda_{_R}} I & 0 \\ 0 & e^{-i\lambda_{_L}} I \end{bmatrix} \begin{bmatrix} e^{i\lambda_{_R}} I & 0 \\ 0 & e^{i\lambda_{_L}} I \end{bmatrix} \begin{bmatrix} \mathbb{B}_N & 0 \\ 0 & -\mathbb{B}_N^* \end{bmatrix} \\
&= \begin{bmatrix} \mathbb{B}_N & 0 \\ 0 & -\mathbb{B}_N^* \end{bmatrix} \\
&= \mathbb{D}_N^* \underline{\Gamma}_1.
\end{aligned}$$

Notice that commutation performed in the second line above is only valid because λ_R and λ_L are constant. Furthermore, if a mass term were present, the above commutation would not be valid. Finally,

$$\begin{aligned} \left\langle \underline{\Lambda}\underline{\xi}, \underline{\Gamma}_1 \mathbb{D}_N \underline{\Lambda}\underline{\xi} \right\rangle &= \left\langle \mathbb{D}_N^* \underline{\Gamma}_1 \underline{\xi}, \underline{\xi} \right\rangle \\ &= \left\langle \underline{\xi}, \underline{\Gamma}_1 \mathbb{D}_N \underline{\xi} \right\rangle, \end{aligned}$$

as desired.

□

3.1.3 Species Doubling

Thus far, the Naive Dirac operator, $\mathbb{D}_N$, has shown a great deal of promise. The discretization is simple, it agrees spectrally with the continuum operator, and it satisfies discrete versions of gauge covariance and chiral symmetry. There is, however, a major problem with the discretization, which we illustrate here using a 1D version of the Schwinger model. In 1D, the Naive Dirac operator is given by

$$\mathbb{D}_N = \gamma_1 \otimes \nabla_x^h + mI, \tag{3.18}$$

where ∇_x^h has the stencil previously given in (3.3). In the gauge-free case (that is, $\mathcal{A} = 0$), operator ∇_x^h, acting on the right-handed component of $\underline{\psi}$, becomes

$$\nabla_x^h \underline{\psi}_{Rj,k} = \frac{1}{2h} \left(\underline{\psi}_{Rj+1,k} - \underline{\psi}_{Rj-1,k} \right), \tag{3.19}$$

where h is the lattice spacing. In the 1D free case,

$$\mathbb{D}_N = \begin{bmatrix} mI & \mathbb{B}_x \\ \mathbb{B}_x & mI \end{bmatrix}, \tag{3.20}$$

where $\mathbb{B}_x$ is the periodic Toeplitz matrix with stencil $\frac{1}{h}[-1/2 \;\; 0 \;\; 1/2]$. For simplicity, assume that the 1D lattice has N cells and, thus, $N+1$ periodic lattice sites, and that N is even. Then, the eigenvalues of $\mathbb{B}_x$ are given by

$$\nu_k = \frac{i}{h}\sin\left(\frac{2\pi k}{N}\right), \tag{3.21}$$

for $k = -(N/2-1), \ldots, N/2$. Note that ν_k and ν_{-k}, for $k = 1, \ldots, N/2$, are complex conjugate pairs. It is clear from the form of $\mathbb{D}_N$ that the eigenvalues of the discrete propagator, $\mathbb{D}_N^{-1}$, are given by

$$\kappa_k = \frac{h}{mh + i\sin(2\pi k/N)}, \tag{3.22}$$

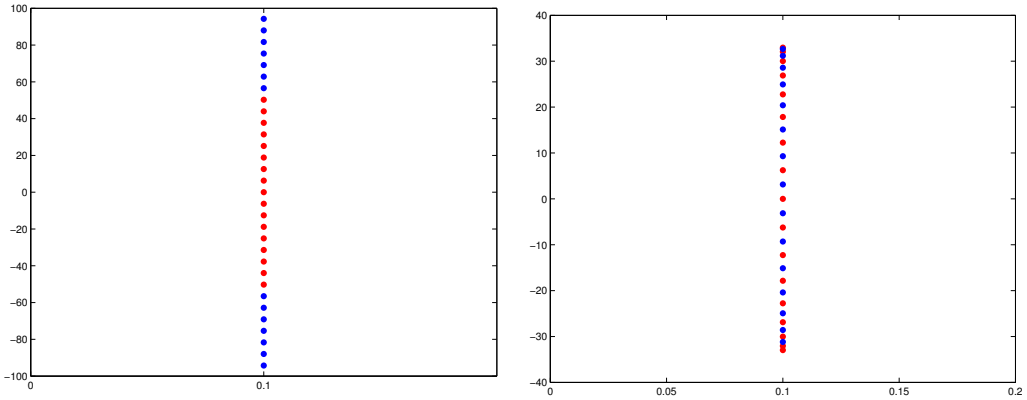
with corresponding eigenvectors

$$\underline{v}_k = \begin{cases} [1, 1, \ldots, 1, 1]^t & k = 0, \\ [\ldots, \cos(2\pi k\ell/N) \pm \sin(2\pi k\ell/N), \ldots]^t & k = \pm 1, \ldots, N/2-1, \\ [1, -1, \ldots, 1, -1]^t & k = N/2, \end{cases} \tag{3.23}$$

where $l = 1, \ldots, n$. Notice the symmetry of κ_k. For every low frequency eigenvector, a corresponding high frequency eigenvector shares the same eigenvalue. The physics community is especially concerned with the correspondence between the eigenvalues of the $k = 0$ and $k = N/2$ modes. In the Naive discretization, the eigenvalues of the propagator, $\mathbb{D}_N^{-1}$, associated with these two modes both approach ∞ as $m \to 0$. Loosely speaking, this represents two particles of different momenta with the same energy, which is impossible. Hence, this phenomenon is referred to as *species doubling* [29].

The presence of species doubling in the Naive discretization can easily be recognized by simply plotting the spectrum of the discrete operator. Figure 3.1 shows the spectra of $\mathcal{D}$ and $\mathbb{D}_N$, respectively. The two colors represent eigenvalues associated with

low and high frequency modes. Notice that, in the spectrum of the continuum operator, the high frequency eigenvalues continue to grow in absolute value, while, in the Naive operator, the high frequency modes double back over the low frequency modes.



(a) Spectrum of $\mathcal{D}$ (b) Spectrum of $\mathbb{D}_N$

Figure 3.1: The spectrum of the 1D, gauge free continuum Schwinger operator, $\mathcal{D}$, and the Naive discrete operator, $\mathbb{D}_N$, respectively, with $m = 0.1$ and $N = 32$. Red and blue dots indicate low and high frequency modes, respectively.

In the applied mathematics community, doubling is known as a red-black instability. There are a number of successful remedies [50]. However, the issue is not only removal of the spurious high frequency components in the discrete solution, but overall accuracy of the discretization process. The addition of the complex gauge field further complicates the situation. The traditional remedy in the physics community is to add an artificial stabilization term to $\mathbb{D}_N$, which we demonstrate below.

3.2 Wilson's Discretization

The addition of the artificial stabilization term to $\mathbb{D}_N$ is the basis for the Dirac-Wilson operator, given by

$$\mathbb{D}_W \quad = \quad \mathbb{D}_N - I \otimes \frac{h}{2} \Delta_G^h, \tag{3.24}$$

where Δ_G^h is a discretization of the so-called *Gauge Laplacian* operator. The stencil for the Gauge Laplacian is given by

$$\Delta_G^h = \frac{1}{h^2}\begin{bmatrix} & e^{i\underline{\theta}_{j,k+1/2}} & \\ e^{-i\underline{\theta}_{j-1/2,k}} & -4 & e^{i\underline{\theta}_{j+1/2,k}} \\ & e^{-i\underline{\theta}_{j,k-1/2}} & \end{bmatrix}. \tag{3.25}$$

Notice that, in the gauge-free case, Δ_G^h coincides with the usual finite difference discretization of the Laplace operator. In matrix form, the Dirac-Wilson operator is given by

$$\mathbb{D}_W = \begin{bmatrix} -\frac{h}{2}\Delta_G^h + mI & \nabla_x^h - i\nabla_y^h \\ \nabla_x^h + i\nabla_y^h & -\frac{h}{2}\Delta_G^h + mI \end{bmatrix}. \tag{3.26}$$

The addition of the Gauge Laplacian ensures that each unknown is self-connected, even in the massless case. Thus, the problem of species doubling is averted, because red-black instability no longer exists. In addition to the centered differences of the first-order terms, the diffusion-like terms now couple each component of the discrete fermion field to its nearest neighbors. This can be seen by again looking at the spectrum of the discrete operator in the 1D gauge free case. Then,

$$\mathbb{D}_W = \begin{bmatrix} \frac{1}{2}\mathbb{H} + mI & \mathbb{B}_x \\ \mathbb{B}_x & \frac{1}{2}\mathbb{H} + mI \end{bmatrix}, \tag{3.27}$$

where $\mathbb{B}_x$ again has the periodic Toeplitz matrix with stencil $\frac{1}{h}[-1/2 \;\; 0 \;\; 1/2]$, and $\mathbb{H}$ is the periodic Toeplitz matrix constructed via the 3-point, periodic, Laplacian stencil $\frac{1}{h}[-1 \;\; 2 \;\; -1]$. Note that $\mathbb{H}$ and $\mathbb{B}_x$ both have eigenvectors corresponding to the discrete Fourier modes defined in (3.23). Again, assume that the 1D lattice has N cells and that

N is even. Then, the eigenvalues of $\mathbb{B}_N$ are again given by ν_k as defined in (3.21). The eigenvalues of $\mathbb{H}$ are given by

$$\alpha_k = \frac{2}{h}\left[1 - \cos\left(\frac{2\pi k}{N}\right)\right]. \tag{3.28}$$

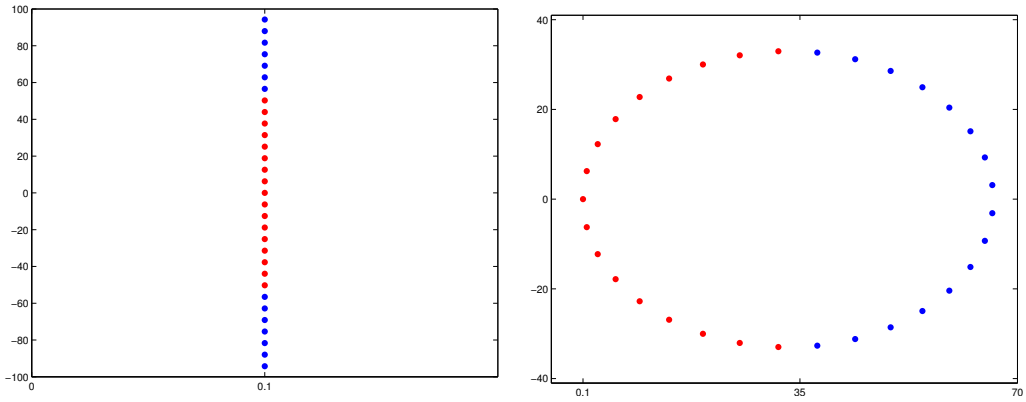
Consequently, the eigenvalues of the Dirac-Wilson propagator, $\mathbb{D}_W^{-1}$, are given by

$$\begin{aligned}\lambda_k &= \left[m + \frac{1}{2}\alpha_k \pm \frac{1}{h}\nu_k\right]^{-1} \\ &= \left[m + \frac{1}{h}\left\{1 - \cos\left(\frac{2\pi k}{N}\right) \pm i \sin\left(\frac{2\pi k}{N}\right)\right\}\right]^{-1},\end{aligned}$$

Note that, in this formulation, the eigenvalue corresponding to the lowest frequency mode, that is, λ_0, does approach ∞ as $m \to 0$, but the eigenvalue corresponding to the highest frequency, that is, $\lambda_{N/2}$, now approaches h. Thus, the Dirac-Wilson operator does not suffer from species doubling.

Figure 3.2 shows the spectrum of the 1D gauge free continuum operator, $\mathcal{D}$, along with the spectrum of the 1D gauge free $\mathbb{D}_W$. Notice that instead of the high frequency modes doubling back on the low frequency modes, they are given a larger real part. Thus, the Dirac-Wilson operator successfully avoids species doubling. This solution comes at a high price however. As seen in Figures 3.2 and 3.3, the spectrum of the discrete operator only corresponds to that of the continuum operator in the very lowest modes. As such, the discrete model can only hope to accurately capture the behavior of the very smoothest modes of the continuum.

The choice to discretize the Gauge Laplacian using covariant finite differences ensures that the operator still satisfies gauge covariance. However, the addition of the Wilson term does have an impact on the chiral symmetry of the model. Recall that the Naive discretization had chiral symmetry because, in the massless case, the discrete operator had zero blocks on the main diagonal. It is obvious from (3.26) that this does



(a) Spectrum of $\mathcal{D}$ (b) Spectrum of $\mathbb{D}_W$

Figure 3.2: The spectrum of the 1D, gauge free, continuum Schwinger operator, $\mathcal{D}$, and Dirac-Wilson operator, $\mathbb{D}_W$, respectively, with $m = .01$ and $N = 32$. Red and blue dots indicate low and high frequency modes, respectively.

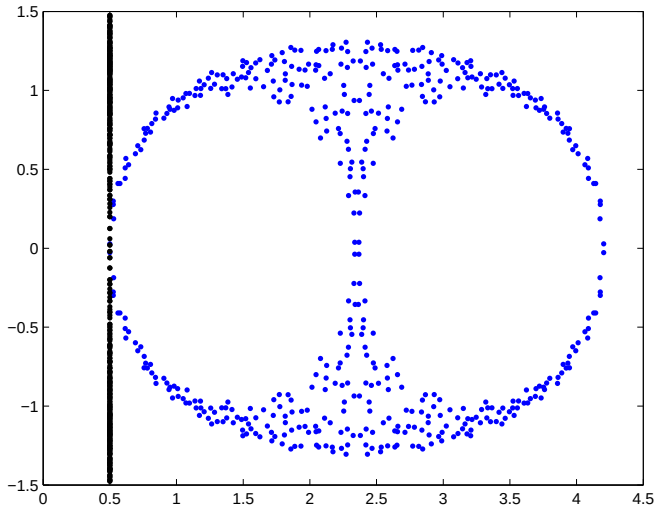


Figure 3.3: The spectrum of the 2D continuum Schwinger operator and the Dirac-Wilson operator, respectively, with nonzero gauge field.

not occur in the Dirac-Wilson case. Proponents of the discretization argue that, in the limit as $h \to 0$, the Wilson term vanishes, ensuring that chiral symmetry is obtained as the lattice approximation approaches the continuum. From Figure 3.3 we see that only a few eigenvalues of $\mathbb{D}_W$ match the continuum. This means that h must be very

small in order for many of the discrete eigenvector-eigenvalue pairs to agree with the continuum.

Chapter 4

The Least-Squares Finite Element Method

The least-squares finite element methodology has been employed in many instances to overcome the shortcomings of the standard Galerkin finite element method applied to nonsymmetric systems of first-order PDEs. Applying the Galerkin methodology to such systems often leads to significant deficiencies. One such deficiency resulting from the Galerkin discretization of first-order operators is red-black instability, or, in the context of lattice gauge theory, species doubling. Many stabilization strategies exist for this difficulty including adding artificial diffusion to the governing equations (as in Wilson's discretization) and employing nonsymmetric discretization techniques, such as upwinding [50]. It has been demonstrated that the former strategy, in the context of covariant finite differences, can do serious damage to the spectrum of the discrete operator. The latter strategy is immediately discounted by the physics community because it leads to lattice actions that give an artificial preference to one spatial direction over another.

Initially applied to first-order system formulations of convection diffusion problems, the least-squares methodology has since been successfully applied to problems in fluid flow (Navier-Stokes) [6], [7], electromagnetism (Maxwell's equations) [41], neutron transport [3], [40], and plasma physics (magnetohydrodynamics) [1]. In general, the method formulates the solution of a system of first-order PDEs in terms of a minimization principle in an infinite dimensional Hilbert space. Then, restricting the resulting

weak form to a finite dimensional space results in a symmetric positive semidefinite system of linear equations.

In this section, we consider a general system of first-order, linear, PDEs of the form $\mathcal{L}\psi = f$. Here, $\mathcal{L} : \mathcal{V} \mapsto \left(L^2\right)^r$, where r is the number of equations in the system and $\mathcal{V} \subset \left(H^1\right)^r$ is a Hilbert space over the complex numbers with appropriate boundary conditions. Let $\mathcal{R}$ be the problem domain and define the following norms and their associated spaces:

$$\begin{aligned} \|\psi\|_0 &= \left(\int_{\mathcal{R}} |\psi(x)|^2 dx\right)^{1/2}, \\ L^2(\mathcal{R}) &= \{\psi : \|\psi\|_0 < \infty\}, \\ \|\psi\|_1 &= \left(\|\psi\|_0^2 + \|\nabla\psi\|_0^2\right)^{1/2}, \\ H^1(\mathcal{R}) &= \{\psi : \|\psi\|_1 < \infty\}. \end{aligned}$$

The linear system, $\mathcal{L}\psi = f$, is then recast as a minimization problem according to

$$\psi = \arg\min_{\varphi \in \mathcal{V}} G(\varphi, f) := \arg\min_{\varphi \in \mathcal{V}} \|\mathcal{L}\varphi - f\|_0^2, \tag{4.1}$$

where ψ is the solution in $\mathcal{V}$. If ψ is the minimizer in (4.1), then clearly ψ satisfies

$$G'(\psi)[v] = 0, \tag{4.2}$$

where $G'(\psi)[v]$ is the Fréchét derivative of G, in the direction of $v \in \mathcal{V}$, evaluated at ψ. Then, (4.2) implies that (4.1) can be cast as the solution of the weak form

$$\text{Find } \psi \in \mathcal{V} \text{ s.t. } \langle \mathcal{L}\psi, \mathcal{L}v \rangle = \langle f, \mathcal{L}v \rangle \quad \forall v \in \mathcal{V}, \tag{4.3}$$

where $\langle \cdot\,, \cdot \rangle$ is the L^2 inner product. It is worth noticing that, for sufficiently smooth ψ and f, (4.3) is equivalent to

$$\text{Find } \psi \in \mathcal{V} \text{ s.t. } \langle \mathcal{L}^*\mathcal{L}\psi, v \rangle = \langle \mathcal{L}^* f, v \rangle \quad \forall v \in \mathcal{V}, \tag{4.4}$$

where $\mathcal{L}^*$ represents the formal adjoint of $\mathcal{L}$. This is the same weak form that would result from applying a Ritz-Galerkin method to $\mathcal{L}^*\mathcal{L}\psi = \mathcal{L}^* f$. Note that this is simply the continuum normal equations of the original problem, $\mathcal{L}\psi = f$. It is common then, to look at the *formal normal*, $\mathcal{L}^*\mathcal{L}$, for insight into the nature of the linear system that results from the least-squares discretization. Note that, in general, $\mathcal{L}^*\mathcal{L}$ is self-adjoint and nonnegative, a quality that the continuum Dirac operator does not have.

A desirable property of the least-squares functional, G, defined in (4.1), is that $G(\psi, 0)$ is *elliptic* with respect to some norm.

Definition 4.0.1. *Quadratic functional $\mathcal{F}$ is said to be elliptic on $\mathcal{V}$ with respect to norm $\|\cdot\|$ if there exists positive constants c and C, independent of ψ, such that*

$$c\|\psi\|^2 < \mathcal{F}(\psi) < C\|\psi\|^2 \quad \forall \psi \in \mathcal{V}. \tag{4.5}$$

The left-hand inequality is known as the coercivity condition. The right-hand inequality is known as the continuity condition.

If functional G is elliptic, then the minimization principle given in (4.1), and by extension, the weak form given in (4.3), has a unique solution on $\mathcal{V}$.

To obtain a discrete linear system that approximates the continuum equations we restrict $\mathcal{V}$ to a suitable finite-dimensional subspace, $\mathcal{V}^h \subset \mathcal{V}$, where $\mathcal{V}^h = \text{span}\{\phi_j\}$ for some set of basis functions $\{\phi_j\}$. Then, the weak form in (4.3) becomes

$$\text{Find } \psi^h \in \mathcal{V}^h \text{ s.t. } \left\langle \mathcal{L}\psi^h, \mathcal{L}v^h \right\rangle = \left\langle f, \mathcal{L}v^h \right\rangle \quad \forall v^h \in \mathcal{V}^h. \tag{4.6}$$

This discrete weak form is equivalent to the system of linear equations

$$\mathbb{A}\underline{\psi} = \underline{f}, \tag{4.7}$$

where

$$[\mathbb{A}]_{jk} = \langle \mathcal{L}\phi_k, \mathcal{L}\phi_j \rangle, \tag{4.8}$$

$$\left[\underline{f}\right]_j = \langle f, \mathcal{L}\phi_j \rangle, \tag{4.9}$$

and the discrete solution, $\underline{\psi}$, is related to the finite element solution, ψ^h, according to

$$\psi^h = \sum_j \underline{\psi}_j \phi_j. \tag{4.10}$$

It is easy to see from (4.8) that the resulting matrix operator, $\mathbb{A}$, is Hermitian positive semidefinite. Finally, if $G(\psi, 0)$ can be shown to be elliptic with respect to the H^1 norm, an optimal multilevel iterative method exists that can efficiently solve the resulting linear system, given in (4.7) [54].

Chapter 5

Least-Squares Finite Elements for the Schwinger Model

This chapter introduces a discretization of the 2D Schwinger model using least-squares finite elements. Since this leads to a discrete solution process that is not automatically gauge covariant, the method is then modified using a process called *gauge fixing*. The resulting process is shown to satisfy discrete gauge covariance and chiral symmetry without suffering from species doubling. Next, H^1-ellipticity of the least-squares functional is demonstrated, the implications of which are discussed for the physical properties of the discrete operator as well as its solution by multilevel iterative methods. Finally, numerical experiments are carried out to test the effectiveness of algebraic multigrid (AMG) and adaptive smoothed aggregation multigrid (αSA) as preconditioners for solving the system of linear equations. The results show that the discrete least-squares operator can be solved effectively by a multilevel method. Furthermore, AMG and αSA, applied to the least-squares system, do not suffer from critical slowing down. Finally, the performance of AMG as a preconditioner for the least-squares operator is compared with the performance of αSA as a preconditioner for the solution of the Dirac-Wilson operator. We find that the least-squares operator can be inverted roughly six times as fast as the Dirac-Wilson operator.

5.1 The Least-Squares Discretization

The goal of this chapter is to discretize the 2D Schwinger model using least-squares finite elements. The governing equations are reiterated here for convenience:

$$\begin{bmatrix} mI & \nabla_x - i\nabla_y \\ \nabla_x + i\nabla_y & mI \end{bmatrix} \begin{bmatrix} \psi_R \\ \psi_L \end{bmatrix} = \begin{bmatrix} f_R \\ f_L \end{bmatrix} \text{ in } \mathcal{R}, \tag{5.1}$$

$$\begin{aligned} \psi(0,y) &= \psi(1,y) \quad \forall y \in (0,1), \\ \psi(x,0) &= \psi(x,1) \quad \forall x \in (0,1), \end{aligned}$$

Let $\mathcal{V}_{\mathbb{R}}$ and $\mathcal{V}_{\mathbb{C}}$ be spaces of continuous, periodic, real- and complex-valued functions, respectively. The solution of (5.1) can be reformulated in terms of a minimization principle:

$$\psi = \arg\min_{\varphi \in \mathcal{V}_{\mathbb{C}}^2} G(\varphi, \mathcal{A}; f) := \arg\min_{\varphi \in \mathcal{V}_{\mathbb{C}}^2} \|\mathcal{D}\varphi - f\|_0^2. \tag{5.2}$$

The decision to cast the minimization problem in terms of the L^2-norm requires that $(\mathcal{D}\varphi - f) \in \left[L^2(\mathcal{R})\right]^2$. It is sufficient to require that the components of the source term, f_R and f_L, belong to $L^2(\mathcal{R})$, the components of the fermion field, ψ_R and ψ_L, belong to $H^1(\mathcal{R})$, and the components of the gauge field, $\mathcal{A}_1$ and $\mathcal{A}_2$, be essentially bounded and belong to $L^\infty(\mathcal{R})$. Then, given the restrictions that ψ and $\mathcal{A}$ be periodic, let $\mathcal{V}_{\mathbb{C}}$ be the space of periodic, complex-valued functions in $H^1(\mathcal{R})$, and let $\mathcal{V}_{\mathbb{R}}$ be the space of periodic, real-valued, essentially bounded functions in $L^\infty(\mathcal{R})$. Furthermore, if $\mathcal{A} \in [L^\infty(\mathcal{R})]^2$, and $\mathcal{A}$ has the Helmholtz decomposition $\mathcal{A} = \nabla^\perp u + \nabla v + \underline{k}$, where u and v are periodic, then u and v must belong to some space $\mathcal{W}_{\mathbb{R}} \subset \mathcal{W}_1^\infty(\mathcal{R})$, where $\mathcal{W}_1^\infty(\mathcal{R})$ is the space of periodic, real-valued functions belonging to $L^\infty(\mathcal{R})$ and whose partial derivatives belong to $L^\infty(\mathcal{R})$ as well. $\mathcal{W}_{\mathbb{R}}$, then, is the subspace of $\mathcal{W}_1^\infty(\mathcal{R})$

consisting of periodic, real-valued functions [17].

Equation (5.2) is equivalent to the weak form

$$\text{Find } \psi \in \mathcal{V}_{\mathbb{C}}^2 \text{ s.t. } \langle \mathcal{D}\psi, \mathcal{D}v \rangle = \langle f, \mathcal{D}v \rangle \quad \forall v \in \mathcal{V}_{\mathbb{C}}^2. \tag{5.3}$$

The formal normal, $\mathcal{D}^*\mathcal{D}$, in this case appears as

$$\begin{aligned} \mathcal{D}^*\mathcal{D} &= \begin{bmatrix} mI & -\nabla_x + i\nabla_y \\ -\nabla_x - i\nabla_y & mI \end{bmatrix} \begin{bmatrix} mI & \nabla_x - i\nabla_y \\ \nabla_x + i\nabla_y & mI \end{bmatrix} \\ &= \begin{bmatrix} m^2 I - \nabla_x^2 - \nabla_y^2 - i\left[\nabla_x, \nabla_y\right] & 0 \\ 0 & m^2 I - \nabla_x^2 - \nabla_y^2 - i\left[\nabla_y, \nabla_x\right] \end{bmatrix}. \end{aligned}$$

Thus, the discrete operator that the least-squares methodology produces has uncoupled Laplacian-like terms on the main diagonal. Note that the term $\nabla_x^2 + \nabla_y^2$ is the continuum version of the gauge Laplacian discussed previously. Though these are not simple constant-coefficient operators (because they include the random background fields), their Hermitian positive semidefinite scalar character should lend themselves to efficient solution by multilevel methods.

The finite element solution is obtained by restricting the minimization problem in (5.2), and, thus, the weak form in (5.3), to a finite-dimensional space, $\mathcal{V}_{\mathbb{C}}^h \subset \mathcal{V}_{\mathbb{C}}$. Then, the finite-element approximation to the solution, ψ^h, must satisfy the weak form

$$\text{Find } \psi^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2 \text{ s.t. } \left\langle \mathcal{D}\psi^h, \mathcal{D}v^h \right\rangle = \left\langle f, \mathcal{D}v^h \right\rangle \quad \forall v^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2. \tag{5.4}$$

Note that $\mathcal{D}$ in (5.4) is the usual Dirac operator in the continuum. However, $\mathcal{D}$ depends on the continuum gauge field, $\mathcal{A}$. Since realistic continuum gauge data is not available we must represent provided discrete gauge data by functions in the continuum. To do this, we interpolate discrete gauge data to the continuum by replacing the continuum

gauge field, $\mathcal{A}$, with finite element gauge data, $A^h \in \mathcal{W}^h_{\mathbb{R}} \subset \mathcal{V}_{\mathbb{R}}$, thus making (5.4) well defined. We assume that A^h is a good approximation to continuum gauge field, $\mathcal{A}$, such that $\|\mathcal{A} - A^h\|_\infty = \mathcal{O}(h)$. Details of this process and the specific form of $\mathcal{W}^h_{\mathbb{R}}$ are provided below.

To compute the inner products in (5.4), a finite element basis is required for fermion field ψ^h as well as the gauge field, A^h. In analogy to the nodal setting, each elementary square on the lattice is represented by a quadrilateral finite element. Recall that $\mathcal{N}_{\mathbb{C}}$ is the space of discrete complex-valued vectors, with values associated with the lattice sites. We equate any discrete vector $\underline{w} = \left[\underline{w}_{_R}, \underline{w}_{_L}\right]^t \in (\mathcal{N}_{\mathbb{C}})^2$ with the piecewise bilinear function $w^h = \left[w^h_{_R}, w^h_{_L}\right]^t \in \left(\mathcal{V}^h_{\mathbb{C}}\right)^2$, where $\mathcal{V}^h_{\mathbb{C}} = \text{span}\{\phi_j\}_{j=1}^{n^2}$ is the space of periodic piecewise bilinear finite element functions over the complex numbers. Here, ϕ_j is the standard nodal basis function associated with lattice site x_j. Then, naturally,

$$\begin{aligned} w^h_{_R} &= \sum_{j=1}^{n^2} \underline{w}_{_R j}\phi_j, \\ w^h_{_L} &= \sum_{j=1}^{n^2} \underline{w}_{_L j}\phi_j. \end{aligned}$$

We wish to represent the discrete gauge field, $\underline{A}$, in the continuum using a finite element function. Recall that $\underline{A} = [\underline{A}_1, \underline{A}_2]^t$ belongs to $\mathcal{E}$, the space of discrete real-valued vectors, with values associated with the lattice links. Gauge data, $\underline{A}_1$ and $\underline{A}_2$, define values on the horizontal and vertical lattice links, respectively. We associate any $\underline{A} \in \mathcal{E}$ with $A^h = \left[A^h_1, A^h_2\right]^t \in \mathcal{W}^h_{\mathbb{R}}$, where A^h is chosen to exactly interpolate the discrete gauge data on the centers of lattice links. To define A^h and $\mathcal{W}^h_{\mathbb{R}}$ precisely we first consider a Helmholtz decomposition of the discrete gauge field:

$$\underline{A} = \begin{bmatrix} \underline{A}_1 \\ \underline{A}_2 \end{bmatrix} = \begin{bmatrix} \mathbb{C}_1 & \mathbb{G}_1 \\ \mathbb{C}_2 & \mathbb{G}_2 \end{bmatrix} \begin{bmatrix} \underline{u} \\ \underline{v} \end{bmatrix} + \begin{bmatrix} \underline{k}_1 \\ \underline{k}_2 \end{bmatrix} \tag{5.5}$$

where $\mathbb{C} = \left[\mathbb{C}_1^t \;\; \mathbb{C}_2^t\right]^t$ and $\mathbb{G} = \left[\mathbb{G}_1^t \;\; \mathbb{G}_2^t\right]^t$ are discrete representations of the curl and gradient operators, respectively. The specific form of $\mathbb{C}$ and $\mathbb{G}$ will be described below. Vectors $\underline{v}$ and $\underline{u}$ are real-valued and are associated with the sites of the standard lattice and the cell-centered lattice, respectively. Then $\underline{u}, \underline{v} \in \mathcal{N}_{\mathbb{R}}$. Note that each row in (5.5) corresponds to gauge data on a link on the standard lattice, with the first block row corresponding to the horizontal links and the second block row corresponding to the vertical links. The rows of the matrix in (5.6), denoted alternatively by $[\mathbb{C}\ \mathbb{G}]$, are defined by the relationship between the individual lattice links, and their contributions from $\underline{u}$ and $\underline{v}$ values on adjoining lattice sites. This relationship is illustrated in Figure 5.1. From this, we see that the rows of matrices $\mathbb{G}$ and $\mathbb{C}$ are defined by the appropriate centered differences that map values of $\underline{v}$ and $\underline{u}$ at sites on the standard and cell-centered lattice, respectively, to values on the links of the standard lattice.

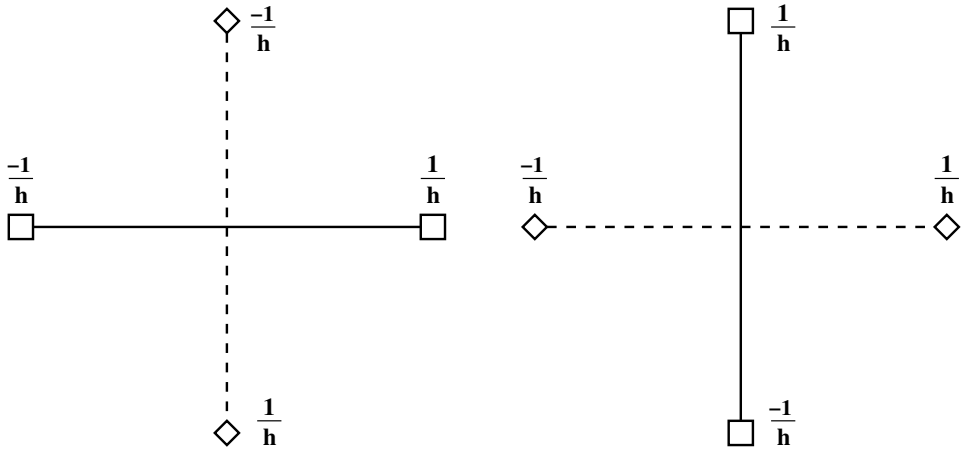


Figure 5.1: Contributions of $\underline{u}$ and $\underline{v}$ to discrete gauge data on a horizontal lattice link (left) and a vertical lattice link (right). $\square$ and $\diamond$ represent sites on the standard and cell-centered lattice, respectively. Solid and dashed lines represent links on the standard and cell-centered lattice, respectively.

The columns of $\mathbb{G}$ and $\mathbb{C}$ are defined by their action on canonical basis functions associated with the standard and cell-centered lattice, respectively. Let $\underline{e}_{jk}$ be the vector, defined on the standard lattice, with value 1 at the $(j,k)^{\text{th}}$ lattice site and zero at all

other sites. Similarly, let $\underline{\epsilon}_{jk}$ be the vector, defined on the cell-centered lattice, with value 1 at the $(j,k)^{\text{th}}$ cell-centered lattice site and zero at all other sites. Here, we use the convention that the $(j,k)^{\text{th}}$ site on the cell-centered lattice is located at the center of the cell with the $(j,k)^{\text{th}}$ standard lattice site in its lower left corner. Figure 5.2 shows the values of $\mathbb{G}\underline{e}_{jk}$ and $\mathbb{C}\underline{\epsilon}_{jk}$ on the appropriate lattice links.

In the following theorem, we establish the existence and uniqueness of the decomposition defined in (5.5).

Theorem 5.1.1. *For discrete gauge field $\underline{A} \in \mathcal{E}$, defined on an $(N+1) \times (N+1)$ periodic lattice, there exist unique vectors, $\underline{v}$ and $\underline{u}$, defined on the standard and cell-centered lattice, respectively, such that*

$$\underline{A} \quad = \quad \mathbb{C}\underline{u} + \mathbb{G}\underline{v} + \begin{bmatrix} \underline{k}_1 \\ \underline{k}_2 \end{bmatrix}, \tag{5.6}$$

and

$$\sum_{jk} \underline{u}_{jk} \quad = \quad \sum_{jk} \underline{v}_{jk} \quad = \quad 0,$$

for some constants, k_1 and k_2.

Proof. Note that the $(N+1) \times (N+1)$ periodic lattice has $2N^2$ distinct lattice links. Thus, $\underline{A} \in \mathbb{R}^{2N^2}$. We begin by showing that, for any discrete gauge field, $\underline{A}$,

$$\underline{A} \in \mathcal{E} \quad = \quad \text{Range}\,(\mathbb{C}) \oplus \text{Range}\,(\mathbb{G}) \oplus \mathcal{K}, \tag{5.7}$$

where $\mathcal{K}$ is the space of vectors that are of a constant value on the horizontal lattice links, and of a (possibly different) constant value on the vertical lattice links. Thus, $\dim(\mathcal{K}) = 2$.

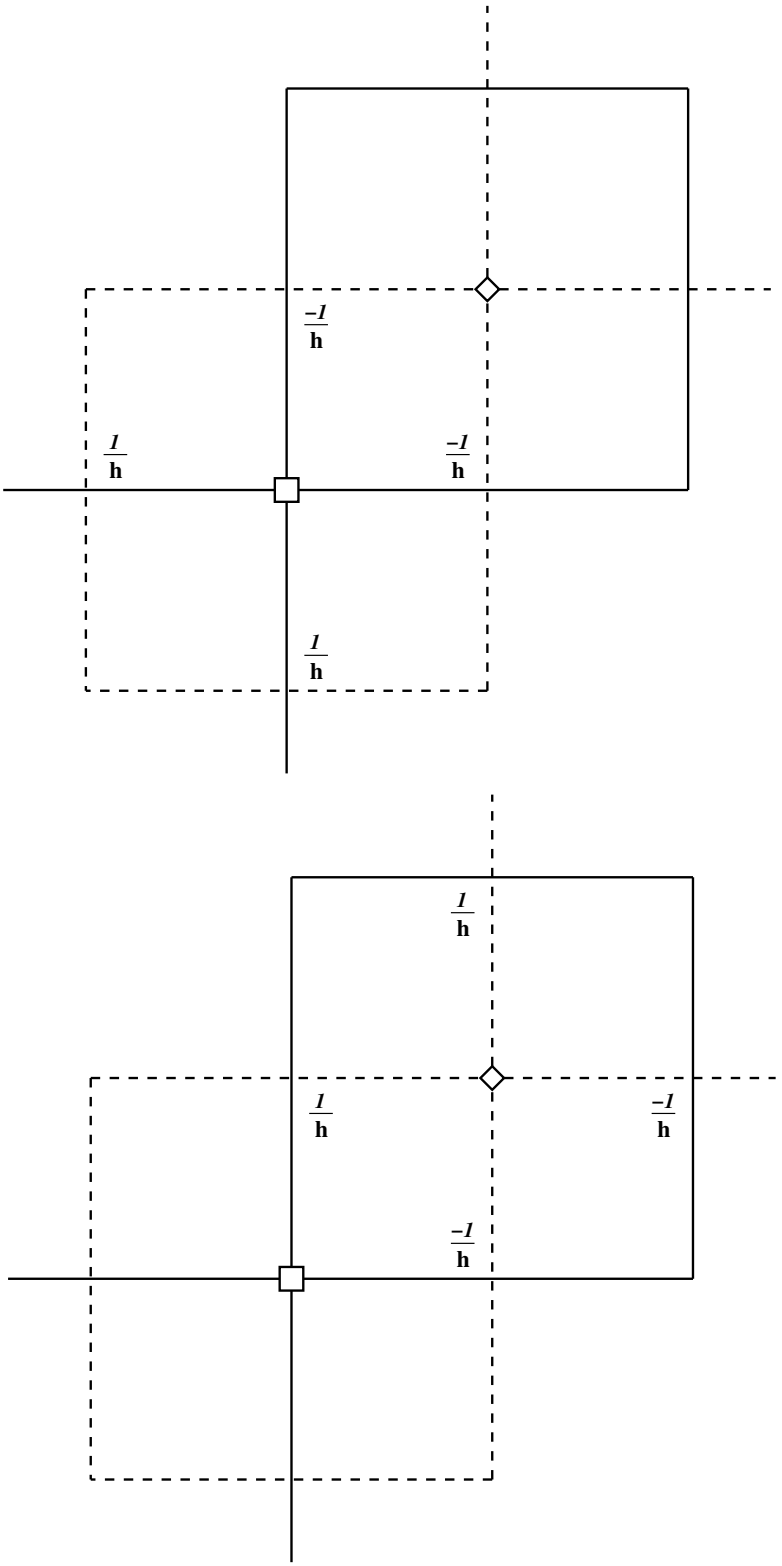


Figure 5.2: The values of $\mathbb{G}\underline{e}_{jk}$ (top) and $\mathbb{C}\underline{\epsilon}_{jk}$ (bottom) associated with the links of the standard and cell-centered lattice, respectively. Here, $\square$ indicates the $(j,k)^{\text{th}}$ site on the standard lattice and $\diamond$ indicates the $(j,k)^{\text{th}}$ site on the cell-centered lattice.

Clearly, $\dim(\text{Range}(\mathbb{G})) = \text{Rank}(\mathbb{G})$ and $\dim(\text{Range}(\mathbb{C})) = \text{Rank}(\mathbb{C})$. Each of the N^2 columns of $\mathbb{G}$ can be associated with the gradient of a bilinear finite element function, ϕ^h_{jk}, on the standard lattice, that has value 1 at site (j,k) and 0 at all other sites. The only null space component of the gradient operator is the constant function, $\phi^h = 1$. Thus, $\text{Rank}(\mathbb{G}) = N^2 - 1$. Similarly, each of the N^2 columns of $\mathbb{C}$ can be associated with the curl of a bilinear finite element function, φ^h_{jk}, on the cell-centered lattice. The only null space component of the curl operator is the constant function, $\varphi^h = 1$. Thus, $\text{Rank}(\mathbb{C}) = N^2 - 1$. Then,

$$\dim(\text{Range}(\mathbb{C})) + \dim(\text{Range}(\mathbb{G})) + \dim(\mathcal{K}) \quad = \quad 2N^2.$$

To establish (5.7) we must show that Range $(\mathbb{C})$, Range $(\mathbb{G})$, and $\mathcal{K}$ are mutually orthogonal. We first show that the columns of $\mathbb{C}$ and $\mathbb{G}$ are orthogonal. To see this, we verify that

$$\langle \mathbb{C}\underline{u}, \mathbb{G}\underline{v} \rangle = 0 \tag{5.8}$$

for all $\underline{v}$ defined on the sites of the standard lattice, and all $\underline{u}$ defined on the sites of the cell-centered lattice. It is sufficient to verify (5.8) for arbitrary basis vectors associated with the standard and cell-centered lattice, $\underline{e}_{jk}$ and $\underline{\epsilon}_{mn}$, respectively. Clearly, if the $(j,k)^{\text{th}}$ site on the standard lattice and the $(m,n)^{\text{th}}$ site on the cell-centered lattice are not part of the same elementary square, then (5.8) trivially holds. Without loss of generality, consider $\underline{e}_{jk}$ and $\underline{\epsilon}_{jk}$, with the convention that the $(j,k)^{\text{th}}$ site on the cell-centered lattice is directly up and to the right of the $(j,k)^{\text{th}}$ site on the standard lattice. Figure 5.2 shows the values of $\mathbb{G}\underline{e}_{jk}$ and $\mathbb{C}\underline{\epsilon}_{jk}$ on the appropriate lattice links. Clearly,

$$\left\langle \mathbb{C}\underline{\epsilon}_{jk}, \mathbb{G}\underline{e}_{jk} \right\rangle \quad = \quad \frac{-1}{h} \cdot \frac{-1}{h} + \frac{-1}{h} \cdot \frac{1}{h} = 0.$$

Thus, (5.8) holds for all $\underline{v}$ on the standard lattice and $\underline{u}$ on the cell-centered lattice, and the columns of $\mathbb{C}$ and $\mathbb{G}$ are orthogonal.

Next, we observe that

$$\langle \mathbb{C}\underline{u}, \underline{k} \rangle = \langle \mathbb{G}\underline{v}, \underline{k} \rangle = 0, \tag{5.9}$$

for all $\underline{v}$ on the standard lattice, $\underline{u}$ on the cell-centered lattice, and arbitrary constant vector $\underline{k} = \left[\underline{k}_1^t \;\; \underline{k}_2^t\right]^t$ defined on the lattice links. Again, it is sufficient to verify (5.9) for arbitrary canonical basis vectors, $\underline{e}_{jk}$ and $\underline{\epsilon}_{jk}$. From the representation of $\mathbb{G}\underline{e}_{jk}$ and $\mathbb{C}\underline{\epsilon}_{jk}$ in Figure 5.2, it is easy to see that multiplying horizontal link values by k_1 and vertical link values by k_2, and then summing the results establishes (5.9). Thus, Range $(\mathbb{C})$ and Range $(\mathbb{G})$ are orthogonal to $\mathcal{K}$. This establishes that any discrete gauge field, $\underline{A}$, has a decomposition of the form (5.6).

Under the current assumptions, the decomposition defined in (5.6) is not unique. Since constant vectors are in the null space of both $\mathbb{C}$ and $\mathbb{G}$, any $\underline{u}$ and $\hat{\underline{u}}$, and any $\underline{v}$ and $\hat{\underline{v}}$, that differ by only a constant, will produce the same $\underline{A}$. To remedy this, we require that the entries of $\underline{u}$ and $\underline{v}$ individually sum to zero. That is, we require that

$$\sum_{jk} \underline{u}_{jk} \quad = \quad \sum_{jk} \underline{v}_{jk} \quad = \quad 0.$$

Under these conditions, the decomposition defined in (5.6) is unique, and the proof is complete.

□

The decomposition of $\mathcal{E}$, given in (5.7), suggests a method of computing $\underline{v}$ and $\underline{u}$, for any gauge field, $\underline{A}$. Specifically, $\underline{u}$ is the orthogonal projection of $\underline{A}$ onto the space of vectors in Range $(\mathbb{C})$ whose entries sum to zero. Likewise, $\underline{v}$ is the orthogonal projection

of $\underline{A}$ onto the space of vectors in Range $(\mathbb{G})$ whose entries sum to zero. Thus, $\underline{u}$ and $\underline{v}$ are the solutions to the following sets of normal equations:

$$\hat{\mathbb{C}}^t\hat{\mathbb{C}}\underline{u} = \hat{\mathbb{C}}^t\underline{A}, \tag{5.10}$$

$$\hat{\mathbb{G}}^t\hat{\mathbb{G}}\underline{v} = \hat{\mathbb{G}}^t\underline{A}, \tag{5.11}$$

where $\hat{\mathbb{C}}$ and $\hat{\mathbb{G}}$ are versions of $\mathbb{C}$ and $\mathbb{G}$ modified to enforce the condition that the entries of $\underline{u}$ and $\underline{v}$ sum to zero. The matrices on the left-hand side of (5.10) and (5.11) are similar to constant coefficient Poisson operators discretized via finite differences. Thus, they can easily be inverted using standard methods. Finally, constants k_1 and k_2 are found by computing

$$\begin{bmatrix} \underline{k}_1 \\ \underline{k}_2 \end{bmatrix} = \underline{A} - \mathbb{C}\underline{u} - \mathbb{G}\underline{v}. \tag{5.12}$$

The development of the Helmholtz decomposition of the discrete gauge field leads us to a convenient representation of the discrete gauge data by a finite element function, A^h. Recall that we related the action of $\mathbb{G}$ and $\mathbb{C}$ on discrete vectors to the application of the gradient and curl operators to bilinear finite element functions defined on the standard and cell-centered lattice, respectively. Then, A^h can be defined in terms of a continuum Helmholtz decomposition involving bilinear finite element functions v^h and u^h defined on the standard and cell centered lattice, respectively, whose entries at lattice sites correspond exactly with the discrete values of $\underline{v}$ and $\underline{u}$. This decomposition is,

$$A^h = \nabla^\perp u^h + \nabla v^h + \underline{k}. \tag{5.13}$$

The definition of v^h as a bilinear finite element function on the standard lattice implies that the gradient portion of the gauge field, ∇v^h, belongs to $\mathcal{W}_v^h$, the Nédélec space

over the real numbers associated with the standard lattice. Similarly, the definition of u^h as a bilinear finite element function on the cell-centered lattice implies that the curl portion of the gauge filed, $\nabla^{\perp} u^h$, belongs to $\mathcal{W}_u^h$, the Raviart-Thomas space over the real numbers associated with a cell-centered lattice. Illustrations of typical basis functions for $\mathcal{W}_u^h$ and $\mathcal{W}_v^h$ can be found in Figures 5.3 and 5.4. Noting that constant vector $\underline{k}$ is represented in both of these spaces, define $\mathcal{W}_{\mathbb{R}}^h = \mathcal{W}_u^h \oplus \mathcal{W}_v^h$. This choice of spaces naturally ensures that the curl and gradient portions of the gauge field are orthogonal.

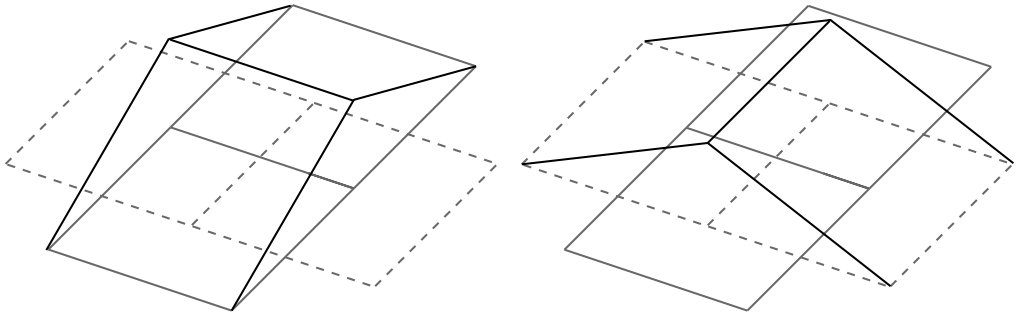


Figure 5.3: Nédélec element (left) and Raviart-Thomas element (right) associated with a horizontal lattice link. The solid grid lines represent the standard lattice, and the dashed represent the cell-centered lattice

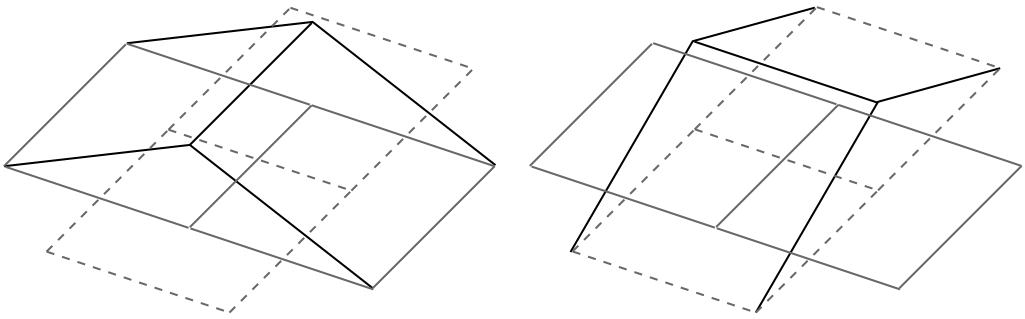


Figure 5.4: Nédélec element (left) and Raviart-Thomas element (right) associated with a vertical lattice link. The solid grid lines represent the standard lattice, and the dashed represent the cell-centered lattice

Our aim is to use the least-squares methodology described above to approximate the solution of (5.1). This process should accept source data, $\underline{f}$, defined on the nodes,

and gauge field data $\underline{A}$, prescribed on the lattice links, and return the discrete wavefunction, $\underline{\psi}$, defined at the nodes. This is accomplished by mapping $\underline{f}$ and $\underline{A}$ into their respective finite element spaces, solving the weak formulation (5.4), and mapping the resulting finite element solution back to $\mathcal{N}_{\mathbb{C}}^2$. This process is summarized in Algorithm 5.1.

ALGORITHM 5.1: Least-Squares Dirac Solve

Input: Gauge field $\underline{A}$, source term $\underline{f}$.

Output: Wavefunction $\underline{\psi}$.

1. Map $\underline{A} \mapsto A^h \in \mathcal{W}_{\mathbb{R}}^h$.
2. Map $\underline{f} \mapsto f^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2$.
3. Find $\psi^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2$ s.t. $\left\langle \mathcal{D}\psi^h, \mathcal{D}v^h \right\rangle = \left\langle f^h, \mathcal{D}v^h \right\rangle \quad \forall v^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2$,
   where $\mathcal{D} = \mathcal{D}\left(A^h\right)$.
4. Map $\psi^h \mapsto \underline{\psi} \in \mathcal{N}_{\mathbb{C}}^2$.

5.1.1 Gauge Covariance

A little reflection on the property of gauge covariance of the fermion propagator in the continuum leads to a test for covariance of the discrete solution process. Let $\omega(x, y)$ be some periodic function and define

$$T_\omega = \begin{bmatrix} e^{i\omega(x,y)} & 0 \\ 0 & e^{i\omega(x,y)} \end{bmatrix}.$$

Note that T_ω is a unitary operator. Consider the following related Dirac equations in the continuum:

$$\mathcal{D}(\mathcal{A})\,\psi \;=\; f, \tag{5.14}$$

$$\mathcal{D}(\mathcal{A}-\nabla\omega)\,\widetilde{\psi} \;=\; T_\omega^*\, f, \tag{5.15}$$

where T_ω^* is the formal adjoint of T_ω. It is easy to see, by the principle of gauge covariance, that ψ and $\widetilde{\psi}$ are related according to

$$\psi \;=\; T_\omega\, \widetilde{\psi}.$$

Transferring these facts to the discrete lattice, let $\underline{\omega}$ be the usual vector of values of $\omega(x,y)$ evaluated at the lattice sites. Let the discrete gauge transformation, $\underline{\Omega}_{\underline{\omega}}$, and the associated matrix, $\mathbb{T}_{\underline{\omega}}$, be as defined in (3.10) and (3.11), respectively. Consider, now, the discrete analogues of (5.14) and (5.15):

$$\mathbb{D}(\underline{A})\,\underline{\psi} \;=\; \underline{f}, \tag{5.16}$$

$$\mathbb{D}(\underline{A}-\mathbb{G}\underline{\omega})\,\underline{\widetilde{\psi}} \;=\; \mathbb{T}_{\underline{\omega}}^*\,\underline{f}. \tag{5.17}$$

Note that (5.16) and (5.17) are simply the discrete Dirac operator constructed using two sets of gauge data, $\underline{A}$ and $\underline{\widetilde{A}}$, and two sets of source data, $\underline{f}$ and $\underline{\widetilde{f}}$, related according to

$$\underline{\widetilde{A}} \;=\; \underline{A}-\mathbb{G}\underline{\omega}, \tag{5.18}$$

$$\underline{\widetilde{f}} \;=\; \mathbb{T}_{\underline{\omega}}^*\,\underline{f}. \tag{5.19}$$

Then, given these two sets of inputs, any discrete solution process should yield solution vectors, $\underline{\psi}$ and $\underline{\widetilde{\psi}}$, such that

$$\underline{\psi} \;=\; \mathbb{T}_{\underline{\omega}}\,\underline{\widetilde{\psi}}.$$

Unfortunately, Algorithm 5.1 does not pass this test. To see this, first consider the test for gauge covariance in the continuum. Suppose the weak form in Algorithm 5.1 is applied to both sets of gauge and source data, but on the entire continuum space $\mathcal{V}_{\mathbb{C}}^2$, rather than $\left(\mathcal{V}_{\mathbb{C}}^h\right)^2$. Then, solutions ψ and $\widetilde{\psi}$ must satisfy, for all $w \in \mathcal{V}_{\mathbb{C}}^2$,

$$\langle \mathcal{D}(\mathcal{A})\,\psi, \mathcal{D}(\mathcal{A})\,w\rangle = \langle f, \mathcal{D}(\mathcal{A})\,w\rangle\,, \tag{5.20}$$

$$\left\langle \mathcal{D}(\mathcal{A}-\nabla\omega)\,\widetilde{\psi}, \mathcal{D}(\mathcal{A}-\nabla\omega)\,w\right\rangle = \left\langle \widetilde{f}, \mathcal{D}(\mathcal{A}-\nabla\omega)\,w\right\rangle, \tag{5.21}$$

respectively. Setting $\widetilde{\psi} = T_\omega^*\,\xi$, $\widetilde{f} = T_\omega^*\,f$, and $w = T_\omega^*\,v$, (5.21) becomes

$$\langle \mathcal{D}(\mathcal{A}-\nabla\omega)\,T_\omega^*\,\xi, \mathcal{D}(\mathcal{A}-\nabla\omega)\,T_\omega^*\,v\rangle = \langle T_\omega^*\,f, \mathcal{D}(\mathcal{A}-\nabla\omega)\,T_\omega^*\,v\rangle\,. \tag{5.22}$$

By gauge covariance, (5.22) is equivalent to

$$\langle T_\omega^*\,\mathcal{D}(\mathcal{A})\,\xi, T_\omega^*\,\mathcal{D}(\mathcal{A})\,v\rangle = \langle T_\omega^*\,f, T_\omega^*\,\mathcal{D}(\mathcal{A})\,v\rangle\,. \tag{5.23}$$

Finally, canceling the gauge transformations, (5.23) becomes

$$\langle \mathcal{D}(\mathcal{A})\,\xi, \mathcal{D}(\mathcal{A})\,v\rangle = \langle f, \mathcal{D}(\mathcal{A})\,v\rangle\,. \tag{5.24}$$

Clearly, (5.20) and (5.24) are equivalent. Thus, $\xi = \psi$ and, as a result, $\psi = T_\omega\widetilde{\psi}$, as desired. This is not surprising since the continuum formulation of the Dirac equation satisfies gauge covariance. Then, restricting the associated weak forms to their respective finite element spaces suggests a test for the gauge covariance of Algorithm 5.1.

Consider discrete gauge fields, $\underline{A}$ and $\underline{\widetilde{A}}$, and source terms, $\underline{f}$ and $\underline{\widetilde{f}}$, related according to (5.18) and (5.19), respectively. Then, for the two data sets, the equalities in the weak forms in Algorithm 5.1 become, for all $w^h \in \left(\mathcal{V}^h_{\mathbb{C}}\right)^2$,

$$\left\langle \mathcal{D}(A^h)\psi^h, \mathcal{D}(A^h)w^h \right\rangle = \left\langle f^h, \mathcal{D}(A^h)w^h \right\rangle, \tag{5.25}$$

$$\left\langle \mathcal{D}(A^h - \nabla\omega^h)\widetilde{\psi}^{\,h}, \mathcal{D}(A^h - \nabla\omega^h)w^h \right\rangle = \left\langle \widetilde{f}^{\,h}, \mathcal{D}(A^h - \nabla\omega^h)w^h \right\rangle, \tag{5.26}$$

where ω^h is the projection of $\underline{\omega}$ into $\left(\mathcal{V}^h_{\mathbb{R}}\right)^2$ and $\widetilde{f}^{\,h}$ is the L^2-projection of $\mathbb{T}^*_{\underline{\omega}}\,\underline{f}$ into $\left(\mathcal{V}^h_{\mathbb{C}}\right)^2$. Proceeding as before, set $\widetilde{\psi}^{\,h} = T^*_{\omega^h}\,\xi^h$ and $w^h = T^*_{\omega^h}\,v^h$. Then, (5.26) becomes

$$\left\langle \mathcal{D}(A^h - \nabla\omega^h)T^*_{\omega^h}\,\xi^h, \mathcal{D}(A^h - \nabla\omega^h)T^*_{\omega^h}\,v^h \right\rangle = \left\langle \widetilde{f}^{\,h}, \mathcal{D}(A^h - \nabla\omega^h)T^*_{\omega^h}\,v^h \right\rangle. \tag{5.27}$$

Again, by gauge covariance of $\mathcal{D}$, (5.27) is equivalent to

$$\left\langle T^*_{\omega^h}\mathcal{D}(A^h)\xi^h, T^*_{\omega^h}\mathcal{D}(A^h)v^h \right\rangle = \left\langle \widetilde{f}^{\,h}, T^*_{\omega^h}\mathcal{D}(A^h)v^h \right\rangle. \tag{5.28}$$

Moving the $T^*_{\omega^h}$ to the other side of the inner product yields

$$\left\langle \mathcal{D}(A^h)\xi^h, \mathcal{D}(A^h)v^h \right\rangle = \left\langle T_{\omega^h}\,\widetilde{f}^{\,h}, \mathcal{D}(A^h)v^h \right\rangle \quad . \tag{5.29}$$

Weak form (5.29) appears very similar to (5.25), except in the inner product on the right-hand side. If it were the case that $T_{\omega^h}\,\widetilde{f}^{\,h} = f^h$, then Algorithm 5.1 would, in fact, be gauge covariant. Unfortunately, since $\widetilde{f}^{\,h}$ is obtained by first multiplying (node-wise) the discrete entries of $e^{i\underline{\omega}}$ and $\underline{f}$, and then projecting the result into $\left(\mathcal{V}^h_{\mathbb{C}}\right)^2$, this equality does not hold. Thus, Algorithm 5.1 is not gauge covariant. Luckily, this problem can be circumvented using a process called *gauge fixing*.

Consider the continuum Dirac equation with gauge field $\mathcal{A}$, given in (5.14). Write the Helmholtz decomposition of the gauge field, $\mathcal{A}$, as

$$\mathcal{A} \;=\; \mathcal{A}_0 + \nabla v,$$

where

$$\mathcal{A}_0 \;=\; \nabla^{\perp} u + \underline{k},$$

for periodic functions u and v and constant vector $\underline{k}$. Note then that $\mathcal{A}_0$ is divergence-free. Equation (5.14) becomes

$$\mathcal{D}\left(\mathcal{A}_0 + \nabla v\right)\psi \;=\; f,$$

to which the solution is

$$\psi \;=\; \left[\mathcal{D}\left(\mathcal{A}_0 + \nabla v\right)\right]^{-1} f.$$

Rewriting the source function as

$$f \;=\; T_v\, g,$$

for some $g \in \mathcal{V}^2_{\mathbb{C}}$, (5.30) becomes

$$\psi \;=\; \left[\mathcal{D}\left(\mathcal{A}_0 + \nabla v\right)\right]^{-1} T_v\, g.$$

But, from gauge covariance of the propagator,

$$\psi \;=\; T_v\, \left[\mathcal{D}\left(\mathcal{A}_0\right)\right]^{-1} g,$$

implying

$$\psi \;=\; T_v \left[\mathcal{D}\left(\mathcal{A}_0\right)\right]^{-1} T_v^* f.$$

Now, consider the continuum equation with modified input data, (5.15). In this case, the Helmholtz decomposition of $\widetilde{\mathcal{A}}$ is

$$\widetilde{\mathcal{A}} \;=\; \mathcal{A}_0 + \nabla\left(v-\omega\right),$$

and the Dirac equation becomes

$$\mathcal{D}\left(\mathcal{A}_0 + \nabla\left(v-\omega\right)\right)\widetilde{\psi} \;=\; \widetilde{f}.$$

Writing the source term as $\widetilde{f} = T_{v-\omega}\,\widetilde{g}$, the solution becomes

$$\widetilde{\psi} \;=\; \left[\mathcal{D}\left(\mathcal{A}_0 + \nabla\left(v-\omega\right)\right)\right]^{-1} T_{v-\omega}\,\widetilde{g}.$$

Again, by gauge covariance, the solution becomes

$$\widetilde{\psi} \;=\; T_{v-\omega}\,\left[\mathcal{D}\left(\mathcal{A}_0\right)\right]^{-1}\widetilde{g},$$

implying

$$\begin{aligned}\widetilde{\psi} \;&=\; T_{v-\omega}\,\left[\mathcal{D}\left(\mathcal{A}_0\right)\right]^{-1} T_{v-\omega}^*\widetilde{f} \\ &=\; T_\omega\left\{T_v\left[\mathcal{D}\left(\mathcal{A}_0\right)\right]^{-1} T_v^* f\right\}.\end{aligned}$$

Thus,

$$\widetilde{\psi} \;=\; T_\omega\, \psi,$$

as desired.

The key to retaining this property in the discrete setting is that the fermion propagator, computed in both cases, is constructed with the same divergence-free gauge field, $\mathcal{A}_0$, and the same source term,

$$\widetilde{f} \;=\; T_v^*\, f.$$

This process of defining the Dirac operator in terms of the same gauge field, $\mathcal{A}_0$, is known as gauge fixing. The decision to choose a divergence-free $\mathcal{A}_0$ is known as fixing the *Coulomb gauge*. A gauge covariant least-squares solution process based on this idea can now be defined.

ALGORITHM 5.2: Gauge Covariant Least-Squares Dirac Solve

Input: Gauge field $\underline{A}$, source term $\underline{f}$.

Output: Wavefunction $\underline{\psi}$.

1. Compute $\underline{A}_0$ and $\underline{v}$ such that $\underline{A} = \underline{A}_0 + \mathbb{G}\underline{v}$
2. Set $\underline{g} = \mathbb{T}^*_{\underline{v}}\, \underline{f}$
3. Map $\underline{A}_0 \mapsto A_0^h \in \mathcal{W}^h_{\mathbb{R}}$
4. Map $\underline{g} \mapsto g^h \in \left(\mathcal{V}^h_{\mathbb{C}}\right)^2$.
5. Find $\zeta^h \in \left(\mathcal{V}^h_{\mathbb{C}}\right)^2$ s.t. $\left\langle \mathcal{D}\zeta^h, \mathcal{D}w^h \right\rangle = \left\langle g^h, \mathcal{D}w^h \right\rangle \quad \forall w^h \in \left(\mathcal{V}^h_{\mathbb{C}}\right)^2$,
   where $\mathcal{D} = \mathcal{D}\left(A_0^h\right)$.
6. Map $\zeta^h \mapsto \underline{\zeta} \in \mathcal{N}^2_{\mathbb{C}}$.
7. Set $\underline{\psi} = \mathbb{T}_{\underline{v}}\, \underline{\zeta}$

Using the nodal basis for $\mathcal{V}_{\mathbb{C}}^h$, the following matrix equation for Step 5 of Algorithm 5.2 can be established:

$$\mathbb{L}\underline{w} \;=\; \mathbb{K}\underline{b}$$

where the entries in vectors $\underline{w}$ and $\underline{b}$ are the coefficients in the expansions of ζ^h and g^h, respectively, and the elements of matrices $\mathbb{L}$ and $\mathbb{K}$ are given by

$$\begin{aligned} [\,\mathbb{L}\,]_{j,k} &= \left\langle \mathcal{D}(A_0^h)\phi_k, \mathcal{D}(A_0^h)\phi_j \right\rangle, \\ [\,\mathbb{K}\,]_{j,k} &= \left\langle \phi_k, \mathcal{D}(A_0^h)\phi_j \right\rangle. \end{aligned}$$

Then, Step 5 in Algorithm 5.2 can be replaced by computing

$$\underline{w} = \mathbb{L}^{-1}\mathbb{K}\underline{b}.$$

and setting

$$\zeta^h = \sum_{j=1}^{n^2} w_j \phi_j.$$

Recalling the relationship between the entries of $\underline{\zeta}$ and $\underline{g}$, and the coefficients in the expansion of ζ^h and g^h, respectively, Steps 3-6 in Algorithm 5.2 can be replaced by

$$\underline{\zeta} = \mathbb{L}^{-1}\mathbb{K}\underline{g}.$$

Specifically, $\mathbb{L}$ and $\mathbb{K}$ have the form

$$\mathbb{L} \;:=\; \begin{bmatrix} m^2\mathbb{M} + \mathbb{L}_{xx} + \mathbb{L}_{yy} + i(\mathbb{L}_{xy} - \mathbb{L}_{yx}) & 0 \\ 0 & m^2\mathbb{M} + \mathbb{L}_{xx} + \mathbb{L}_{yy} - i(\mathbb{L}_{xy} - \mathbb{L}_{yx}) \end{bmatrix}, \quad (5.30)$$

$$\mathbb{K} \;:=\; \begin{bmatrix} m\mathbb{M} & \mathbb{B}_x - i\mathbb{B}_y \\ \mathbb{B}_x + i\mathbb{B}_y & m\mathbb{M} \end{bmatrix}, \tag{5.31}$$

where

$$\begin{aligned} [\,\mathbb{L}_{xx}\,]_{j,k} &=< \nabla_x\phi_k, \nabla_x\phi_j > \qquad [\,\mathbb{M}\,]_{j,k} =< \phi_k, \phi_j > \\ [\,\mathbb{L}_{yy}\,]_{j,k} &=< \nabla_y\phi_k, \nabla_y\phi_j > \qquad [\,\mathbb{B}_x\,]_{j,k} =< \phi_k, \nabla_x\phi_j > \\ [\,\mathbb{L}_{xy}\,]_{j,k} &=< \nabla_x\phi_k, \nabla_y\phi_j > \qquad [\,\mathbb{B}_y\,]_{j,k} =< \phi_k, \nabla_y\phi_j > \\ [\,\mathbb{L}_{yx}\,]_{j,k} &=< \nabla_y\phi_k, \nabla_x\phi_j > . \end{aligned}$$

It is important to notice that the covariant derivative operators, ∇_x and ∇_y, are constructed using the divergence-free gauge field, A_0^h. It is interesting to note the similarity between the form of $\mathbb{L}$ given in (5.30) and the formal normal given in (5.4).

Matrix $\mathbb{L}$ is nonsingular, except in the case that $m = 0$ and the gauge field is an exceptional configuration. To see this, note that, by construction,

$$\|\mathcal{D}\psi^h\|_0^2 \;=\; \left\langle \mathbb{L}\underline{\psi}, \underline{\psi} \right\rangle_{l^2}, \tag{5.32}$$

for any finite element fermion field ψ^h and its lattice counterpart $\underline{\psi}$. From the discussion of the spectrum of the continuum operator, $\mathcal{D}$, in Section 2.1, it is clear from (5.32) that if $\mathcal{D}$ is nonsingular, then $\mathbb{L}$ is positive definite, and thus nonsingular as well. Furthermore, $\mathbb{L}$ is singular only when $m = 0$ and A_0^h is an exceptional configuration.

Note that $\mathbb{K}$ can be written as a sum of Hermitian and skew-Hermitian matrices according to

$$\mathbb{K} \;=\; \begin{bmatrix} m\mathbb{M} & 0 \\ 0 & m\mathbb{M} \end{bmatrix} + \begin{bmatrix} 0 & \mathbb{B}_x - i\mathbb{B}_y \\ \mathbb{B}_x + i\mathbb{B}_y & 0 \end{bmatrix}.$$

By construction, $\mathbb{M}$ is positive definite. Thus, if $m > 0$, then the first term in the decomposition is also positive definite. This implies that $\mathbb{K}$ is positive definite, and thus nonsingular. However, if A_0^h is an exceptional configuration, the skew-Hermitian term is singular, and if $m = 0$, then $\mathbb{K}$ is singular as well.

Using the matrix representations described above, Algorithm 5.2 can be rewritten completely in the discrete setting. This is summarized in Algorithm 5.3.

ALGORITHM 5.3: Discrete Gauge Covariant Least-Squares Dirac Solve

---

Input: Gauge field $\underline{A}$, source term $\underline{f}$.

Output: Wavefunction $\underline{\psi}$.

1. Compute $\underline{A}_0$ and $\underline{v}$ such that $\underline{A} = \underline{A}_0 + \mathbb{G}\underline{v}$, where $\langle \underline{A}_0, \mathbb{G}\underline{w} \rangle = 0 \ \forall \underline{w} \in \mathcal{N}_{\mathbb{C}}$
2. Set $\underline{g} = \mathbb{T}^*_{\underline{v}} \, \underline{f}$
3. Compute $\underline{\zeta} = \mathbb{L}^{-1}\mathbb{K}\underline{g}$
4. Set $\underline{\psi} = \mathbb{T}_{\underline{v}} \, \underline{\zeta}$

Combining Steps 2-4 in Algorithm 5.3 yields an expression for the discrete least-squares propagator, hereafter denoted by $\mathbb{D}_{LS}^{-1}$:

$$\mathbb{D}_{LS}^{-1} \quad = \quad \mathbb{T}_{\underline{v}} \, \mathbb{L}^{-1}\mathbb{K}\mathbb{T}^*_{\underline{v}}. \tag{5.33}$$

From (5.33), the least-squares representation of the Dirac operator is, naturally

$$\mathbb{D}_{LS} \quad = \quad \mathbb{T}_{\underline{v}} \, \mathbb{K}^{-1}\mathbb{L}\mathbb{T}^*_{\underline{v}}. \tag{5.34}$$

Again, the operators given in (5.33) and (5.34) exist except in the case when $m = 0$ and A_0^h is an exceptional configuration.

Finally, note that $\mathbb{L}$ and $\mathbb{K}$ are constructed using $\underline{A}_0$. An encouraging consequence of constructing $\mathbb{L}$ based on $\underline{A}_0$ is that divergence-free gauge fields tend to be smoother than general fields [10]. The decreased disorder in the background field makes $\mathbb{L}$ easier to invert using a multilevel iterative method.

The following theorem establishes the discrete gauge covariance of Algorithm 5.3.

Theorem 5.1.2. *The least-squares solution process defined in Algorithm 5.3 satisfies discrete gauge covariance as described in Definition 3.1.1.*

Proof. Recall that, given the least-squares propagator, $\mathbb{D}_{LS}^{-1}$, and associated discrete gauge transformation matrices, $\underline{\Omega}_{\underline{\omega}}$ and $\mathbb{T}_{\underline{\omega}}$, there must exists a modified propagator, denoted by $\widetilde{\mathbb{D}}_{LS}^{-1}$, such that

$$\mathbb{D}_{LS}^{-1}\,\mathbb{T}_{\underline{\omega}}\,\underline{\xi} \;=\; \mathbb{T}_{\underline{\omega}}\,\widetilde{\mathbb{D}}_{LS}^{-1}\,\underline{\xi}.$$

From (5.33),

$$\mathbb{D}_{LS}^{-1} \;=\; \mathbb{T}_{\underline{v}}\,\mathbb{L}^{-1}\mathbb{K}\mathbb{T}_{\underline{v}}^{*}.$$

Then,

$$\begin{aligned}\mathbb{D}_{LS}^{-1}\,\mathbb{T}_{\underline{\omega}}\,\underline{\xi} &= \mathbb{T}_{\underline{v}}\,\mathbb{L}^{-1}\mathbb{K}\mathbb{T}_{\underline{v}}^{*}\,\mathbb{T}_{\underline{\omega}}\,\underline{\xi}\\ &= \mathbb{T}_{\underline{\omega}}\,\left[\mathbb{T}_{\underline{v}-\underline{\omega}}\,\mathbb{L}^{-1}\mathbb{K}\mathbb{T}_{\underline{v}-\underline{\omega}}^{*}\,\right]\underline{\xi}.\end{aligned}$$

Finally, defining $\widetilde{\mathbb{D}}_{LS}^{-1}$ according to

$$\widetilde{\mathbb{D}}_{LS}^{-1} \;=\; \mathbb{T}_{\underline{v}-\underline{\omega}}\,\mathbb{L}^{-1}\mathbb{K}\mathbb{T}_{\underline{v}-\underline{\omega}}^{*},$$

achieves the desired result.

□

5.1.2 Chiral Symmetry

Recall Definition 3.1.2 regarding the discrete chiral symmetry of a discrete Dirac operator. Given $\lambda_{_R}$ and $\lambda_{_L} \in \mathbb{R}$, and associated transformation $\underline{\Lambda}$, defined in (3.14), the least-squares analogue of the Dirac operator, $\mathbb{D}_{LS}$, must satisfy

$$\left\langle \underline{\Lambda\xi}, \underline{\Gamma}_1 \mathbb{D}_{LS} \underline{\Lambda\xi} \right\rangle \;=\; \left\langle \underline{\xi}, \underline{\Gamma}_1 \mathbb{D}_{LS} \underline{\xi} \right\rangle, \tag{5.35}$$

when $m = 0$. Recall from (5.34) that $\mathbb{D}_{LS}$ is ill-posed in the massless case if the discrete gauge field is an exceptional configuration, because $\mathbb{K}$ is not invertible. Even in this case, the least-squares discretization should satisfy chiral symmetry. The following lemma demonstrates that $\mathbb{D}_{LS}$ satisfies an equivalent formulation of chiral symmetry.

Lemma 5.1.3. *(Chiral symmetry for the discrete least-squares operator). Given any* $\lambda_{_R},\ \lambda_{_L} \in \mathbb{R}$, *and any* $\underline{\psi},\ \underline{f} \in \mathcal{N}^2_{\mathbb{C}}$ *such that*

$$\mathbb{L}\mathbb{T}^*_{\underline{v}}\, \underline{\psi} \;=\; \mathbb{K}\mathbb{T}^*_{\underline{v}}\, \underline{f}$$

for $m = 0$, *then*

$$\begin{aligned} \underline{\hat{\psi}} &= \underline{\Lambda\psi}, \\ \underline{\hat{f}} &= \underline{\Gamma}_1 \underline{\Lambda\Gamma}_1 \underline{f}, \end{aligned}$$

satisfy

$$\mathbb{L}\mathbb{T}^*_{\underline{v}}\, \underline{\hat{\psi}} \;=\; \mathbb{K}\mathbb{T}^*_{\underline{v}}\, \underline{\hat{f}}.$$

Proof. Recalling (5.30) and (5.31), it is easy to see that, in the massless case, $\mathbb{L}$ and $\mathbb{K}$ are of the form

$$\mathbb{L} = \begin{bmatrix} \mathbb{L}_{11} & 0 \\ 0 & \mathbb{L}_{22} \end{bmatrix},$$
$$\mathbb{K} = \begin{bmatrix} 0 & \mathbb{K}_{12} \\ \mathbb{K}_{21} & 0 \end{bmatrix}.$$

Then, $\mathbb{L}\mathbb{T}^*_{\underline{v}}$ and $\mathbb{K}\mathbb{T}^*_{\underline{v}}$ appear as

$$\mathbb{L}\mathbb{T}^*_{\underline{v}} = \begin{bmatrix} \mathbb{L}_{11}\,\underline{\Omega}^*_{\underline{v}} & 0 \\ 0 & \mathbb{L}_{22}\,\underline{\Omega}^*_{\underline{v}} \end{bmatrix}, \tag{5.36}$$
$$\mathbb{K}\mathbb{T}^*_{\underline{v}} = \begin{bmatrix} 0 & \mathbb{K}_{12}\,\underline{\Omega}^*_{\underline{v}} \\ \mathbb{K}_{21}\,\underline{\Omega}^*_{\underline{v}} & 0 \end{bmatrix}. \tag{5.37}$$

The following can be deduced from the block structure of (5.36), and the constant nature of the diagonal blocks of $\underline{\Lambda}$:

$$\mathbb{L}\mathbb{T}^*_{\underline{v}}\,\underline{\Lambda} = \underline{\Lambda}\mathbb{L}\mathbb{T}^*_{\underline{v}}.$$

Then, from the block structure of (5.37) and the constant nature of $\underline{\Lambda}$,

$$\mathbb{K}\mathbb{T}^*_{\underline{v}}\,\underline{\Gamma}_1\,\underline{\Lambda}\,\underline{\Gamma}_1 = \underline{\Lambda}\,\mathbb{K}\mathbb{T}^*_{\underline{v}}.$$

Thus,

$$\begin{aligned} \mathbb{L}\mathbb{T}^*_{\underline{v}}\hat{\psi} &= \mathbb{L}\mathbb{T}^*_{\underline{v}}\underline{\Lambda}\psi \\ &= \underline{\Lambda}\mathbb{L}\mathbb{T}^*_{\underline{v}}\psi, \end{aligned}$$

and

$$\begin{aligned}\mathbb{K}\mathbb{T}^*_{\underline{v}}\,\hat{f} &= \mathbb{K}\mathbb{T}^*_{\underline{v}}\,\underline{\Gamma}_1\,\underline{\Lambda}\,\underline{\Gamma}_1\underline{f} \\ &= \underline{\Lambda}\,\mathbb{K}\mathbb{T}^*_{\underline{v}}\,\underline{f},\end{aligned}$$

which yields the result.

□

5.1.3 Species Doubling

From (5.30), the principle part of the main diagonal blocks of $\mathbb{L}$ is a 9-point Laplacian-like stencil. Since this stencil connects each unknown to its nearest neighbors, matrix $\mathbb{L}$ does not have red-black instability and, thus, the problem of species doubling is averted. This is proved definitively now by examining the spectrum of the effective least-squares operator $\mathbb{D}_{LS}$ in the 1D gauge free case. In this instance,

$$\begin{aligned}\mathbb{L} &= \begin{bmatrix} m^2\mathbb{M}+\mathbb{H} & 0 \\ 0 & m^2\mathbb{M}+\mathbb{H} \end{bmatrix}, \\ \mathbb{K} &= \begin{bmatrix} m\mathbb{M} & \mathbb{B}_x \\ \mathbb{B}_x & m\mathbb{M} \end{bmatrix}, \\ \mathbb{T}_{\underline{v}} &= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.\end{aligned}$$

where $\mathbb{H}$, $\mathbb{B}_x$, and $\mathbb{M}$ are periodic Toeplitz matrices with stencils $\frac{1}{h}\,[-1\;\;2\;\;-1]$, $[-1\;\;0\;\;1]$, and $h\,[1/6\;\;2/3\;\;1/6]$, respectively. The eigenvalues of $\mathbb{M}$ are given by

$$\mu_k = \frac{h}{3}\left[2+\cos\left(\frac{2\pi k}{n}\right)\right]. \tag{5.38}$$

Then, the effective least-squares propagator is given by

$$\mathbb{D}_{LS}^{-1} = \mathbb{L}^{-1}\mathbb{K}, \tag{5.39}$$

and its eigenvalues are given by

$$\tau_k = \frac{m\mu_k \pm i\nu_k}{m^2\mu_k + \alpha_k}. \tag{5.40}$$

Substituting the expressions for μ_k, ν_k, and α_k into (5.40) and simplifying gives

$$\tau_k = \frac{m\ h^2\left[2+\cos\left(2\pi k/N\right)\right] \pm 3i\ h\sin\left(2\pi k/N\right)}{m^2h^2\left[2+\cos\left(2\pi k/N\right)\right] + 6\left[1-\cos\left(2\pi k/N\right)\right]}, \tag{5.41}$$

for $k = -N/2+1, \ldots, N/2$. Recall from Chapter 3 that species doubling is present if the eigenvalue associated with the highest frequency mode, $\tau_{N/2}$, approaches ∞ in the limit as $m \to 0$. As expected, the eigenvalue associated with the lowest frequency mode, τ_0, approaches ∞ as $m \to 0$. Then, setting $k = N/2$ and taking the limit to the massless case, we see that $\tau_{N/2} \to 0$, unlike in the naive propagator, where $\tau_{N/2} \to \infty$. Thus, the least-squares formulation for the 1D Dirac operator does not suffer from species doubling. The generalization of this analysis to the 2D case is straightforward.

It is also easily verified that the effective least-squares operator does not suffer from species doubling by looking at its spectrum graphically. The eigenvalues of the 1D, gauge-free operator are plotted in Figure 5.5. Again, low frequency eigenvalues are given in red, and high frequency eigenvalues are given in blue. Notice that the high frequency modes continue to grow in magnitude as the frequency increases, much like the continuum Dirac operator. And, although the spectrum is not a perfect vertical line in the complex plane, the low modes approximate the continuum much better than those of the Dirac-Wilson operator.

Recall that, in the successful discretization of any PDE, the discrete operator must do a good job of capturing the behavior of the continuum modes associated with

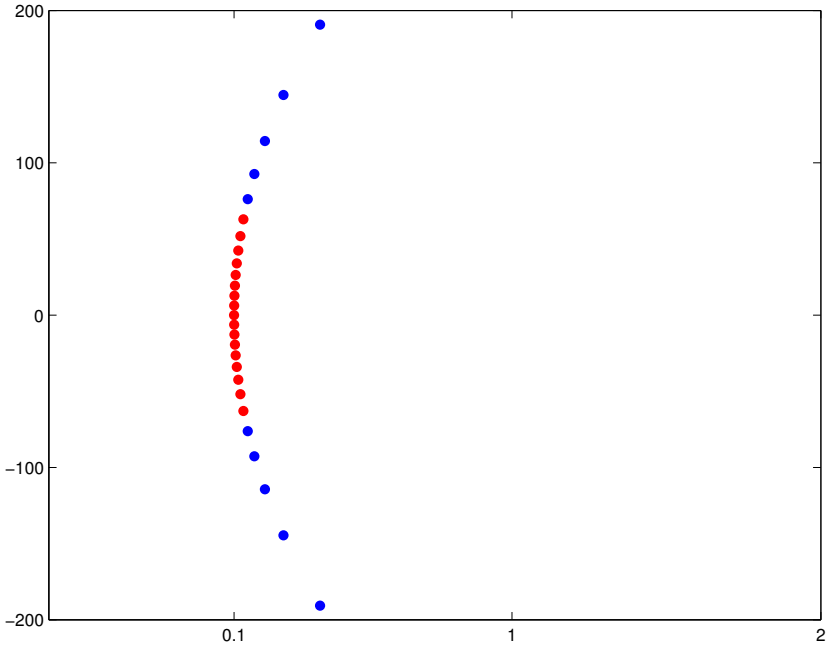


Figure 5.5: The spectrum of the 1D, gauge free least-squares operator, $\mathbb{D}_{LS}$ with $m = 0.1$ and $N = 32$. Red and blue dots indicate low and high frequency modes, respectively.

eigenvalues of smallest modulus. In fact, the least-squares discretization does just this. This point is further illustrated in Figure 5.6, where the lowest modes of the continuum operator, the least-squares operator, and the Dirac-Wilson operator are displayed together. Note that the least-squares operator represents the lowest modes of the continuum operator almost perfectly. On the other hand, only the single lowest mode of the Dirac-Wilson operator comes close to the continuum spectrum.

Naturally, the eigenvectors associated with eigenvalues of small modulus in the operator are associated with eigenvalues of large modulus in the propagator. It is desirable, then, that the high frequency modes of the discrete propagator agree with the high frequency modes of the continuum propagator. The upper end of the spectrum is shown in Figure 5.7 for the continuum, least-squares, and Dirac-Wilson propagators. Recall that the spectrum of the continuum propagator is on a circle in the complex plane, with radius $\frac{1}{2m}$, and centered on the real axis at $\frac{1}{2m}$. Although the eigenvalues of

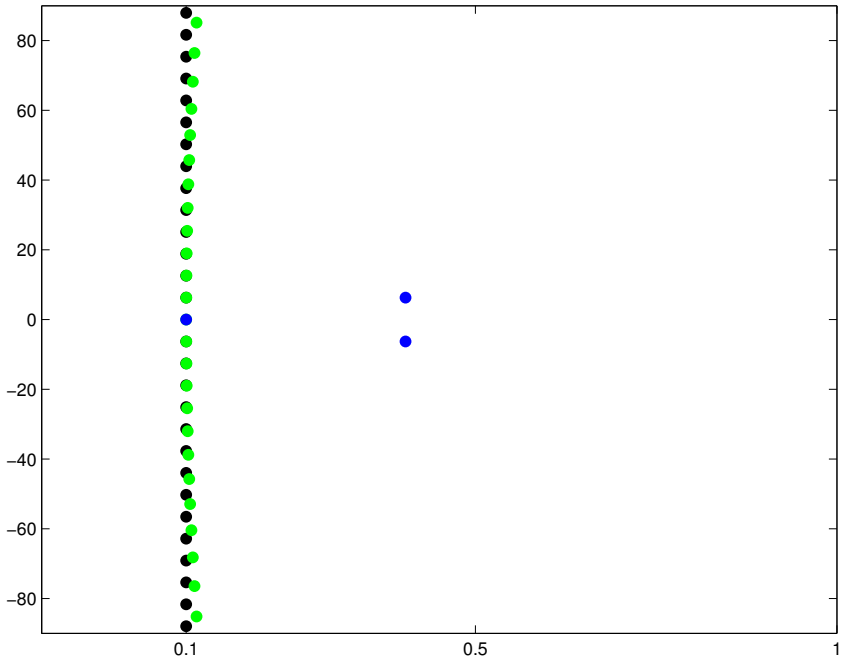


Figure 5.6: Lower end of spectrum of the 1D, gauge free continuum Schwinger operator (Black), $\mathcal{D}$, least-squares operator (Green), $\mathbb{D}_{LS}$, and Wilson operator (Blue), $\mathbb{D}_W$, respectively, with $m = 0.1$ and $N = 32$.

the least-squares propagator is not *exactly* on this circle, they are certainly very close. Note that, again, the eigenvalues of the Dirac-Wilson operator are not even close to those of the continuum.

5.2 H^1-Ellipticity

From (5.2), we see that the least-squares functional is given by

$$G(\psi, \mathcal{A}; f) = \|m\psi_R + \mathcal{B}\psi_L - f_R\|_0^2 + \|m\psi_L - \mathcal{B}^*\psi_R - f_L\|_0^2 \tag{5.42}$$

Several lemmas are required to prove the main result, that the least-squares functional, $G(\psi, \mathcal{A}; 0)$, is elliptic with respect to the H^1 norm, except in the massless case when the gauge field is an exceptional configuration. Under these conditions, $\mathcal{D}$ has a two-

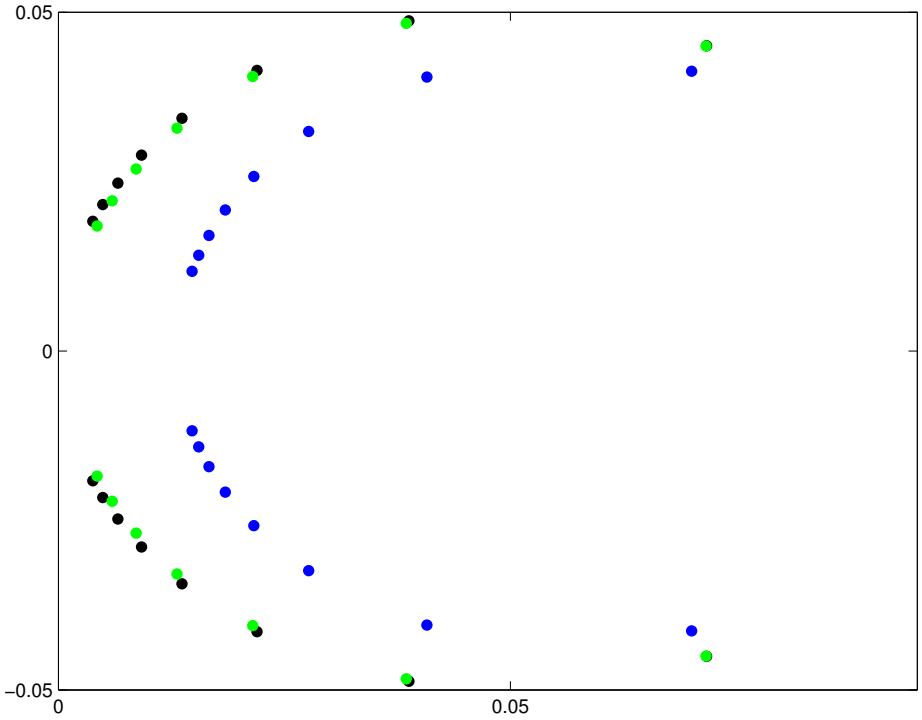


Figure 5.7: Upper end of spectrum of the 1D, gauge free continuum Schwinger propagator (Black), $\mathcal{D}$, least-squares propagator (Green), $\mathbb{D}_{LS}$, and Wilson propagator (Blue), $\mathbb{D}_W$, respectively, with $m = 10$ and $N = 32$.

dimensional null space. In this case, $G(\psi, \mathcal{A}; 0)$ is coercive on the orthogonal complement to that null space.

5.2.1 Main Theorem

Lemma 5.2.1. *Let L be a skew-adjoint operator and $m \in \mathbb{R}$. Then*

$$\|L\psi + m\psi\|_0^2 \quad = \quad \|L\psi\|_0^2 + m^2\|\psi\|_0^2.$$

Proof. Note that

$$\begin{aligned}
\|L\psi + m\psi\|_0^2 &= \langle L\psi + m\psi, L\psi + m\psi\rangle \\
&= \langle L\psi, L\psi\rangle + \langle L\psi, m\psi\rangle + \langle m\psi, L\psi\rangle + \langle m\psi, m\psi\rangle \\
&= \langle L\psi, L\psi\rangle + m\langle L\psi, \psi\rangle + m\langle L^*\psi, \psi\rangle + m^2\langle \psi, \psi\rangle \\
&= \|L\psi\|_0^2 + m\langle (L + L^*)\psi, \psi\rangle + m^2\|\psi\|_0^2.
\end{aligned}$$

Then, noting that $L^* = -L$, because L is skew-adjoint, proves the result.

□

Lemma 5.2.2. *Let $\mathcal{D}_0$ be the massless Dirac operator and let $\mathcal{U}$ be a subspace of $\mathcal{V}_{\mathbb{C}}^2$. If there exists some $c_k > 0$ such that*

$$\|\mathcal{D}_0\psi\|_0^2 \geq c_k\|\psi\|_1^2 \quad \forall\psi \in \mathcal{U},$$

then

$$\|\mathcal{D}\psi\|_0^2 \geq c_k\|\psi\|_1^2 \quad \forall\psi \in \mathcal{U}.$$

Proof. Assume that $\mathcal{D}_0$ is coercive on $\mathcal{U}$. That is, there exists $c_k > 0$ such that

$$\|\mathcal{D}_0\psi\|_0^2 \geq c_k\|\psi\|_1^2 \quad \forall\psi \in \mathcal{U}. \tag{5.43}$$

If $m = 0$, then the proof is complete. Thus, assume $m > 0$. Note that $\mathcal{D}_0$ is skew-adjoint. Then, by Lemma 5.2.1,

$$\|(\mathcal{D}_0 + mI)\psi\|_0^2 = \|\mathcal{D}_0\psi\|_0^2 + m^2\|\psi\|_0^2.$$

Then

$$\begin{aligned}\|\mathcal{D}\psi\|_0^2 &= \|\left(\mathcal{D}_0 + mI\right)\psi\|_0^2 \\ &= \|\mathcal{D}_0\psi\|_0^2 + m^2\|\psi\|_0^2 \\ &\geq c_k\|\psi\|_1^2 + m^2\|\psi\|_0^2 \\ &\geq c_k\|\psi\|_1^2,\end{aligned}$$

as desired.

□

Recall that gauge field $\mathcal{A}$ can be decomposed according to $\mathcal{A} = \nabla^\perp u + \nabla v + \underline{k}$. In the alternate formulation of the Schwinger model, discussed in Section 2.2.3, we use a transformation involving e^z to remove the gauge field from the differential operators in $\mathcal{D}$, where $z = u + iv$. Note that the choice of $u, v \in \mathcal{W}_{\mathbb{R}}$ ensures that $z \in \mathcal{W}_{\mathbb{C}}$, the space of periodic, complex-valued functions in $\mathcal{W}_1^\infty(\mathcal{R})$. This guarantees that e^z is bounded in the ∞-norm.

Lemma 5.2.3. *Let $\mathcal{U}_{_L}$ be a subspace of $\mathcal{V}_{\mathbb{C}}$, and define $\hat{\mathcal{U}}_{_L} = e^z\mathcal{U}_{_L}$ and $\mathcal{B} = e^z\mathcal{B}_k e^{-z}$, where $\mathcal{B} = \nabla_x - i\nabla_y$ and $\mathcal{B}_k = (\partial_x - ik_1) - i(\partial_y - ik_2)$. If there exists $c_k > 0$ such that*

$$\|\mathcal{B}_k\xi\|_0^2 \geq c_k\|\xi\|_1^2 \quad \forall \xi \in \mathcal{U}_{_L},$$

then there exists $c_{_L} > 0$ such that

$$\|\mathcal{B}\xi\|_0^2 \geq c_{_L}\|\xi\|_1^2 \quad \forall \xi \in \hat{\mathcal{U}}_{_L}.$$

Proof. First, note that for any $\xi \in \mathcal{U}_{_L}$, there exists $c > 0$ such that $\|e^z\xi\|_1^2 \leq c\|\xi\|_1^2$. To see this, let ξ be an arbitrary vector in $\mathcal{U}_{_L}$. Then, by repeated use of the triangle and Cauchy's inequalities,

$$\begin{aligned}
\|e^z\xi\|_1^2 &= \|e^z\xi\|_0^2 + \|\nabla\left(e^z\xi\right)\|_0^2 \\
&= \|e^z\xi\|_0^2 + \|\nabla\left(e^z\right)\xi + e^z\nabla\xi\|_0^2 \\
&\leq \|e^z\xi\|_0^2 + 2\left[\|\nabla\left(e^z\right)\xi\|_0^2 + \|e^z\nabla\xi\|_0^2\right] \\
&\leq \|e^z\|_\infty^2\|\xi\|_0^2 + 2\left[\|\nabla\left(e^z\right)\|_\infty^2\|\xi\|_0^2 + \|e^z\|_\infty^2\|\nabla\xi\|_0^2\right] \\
&\leq c_1\left[\|\xi\|_0^2 + \|\nabla\xi\|_0^2\right] \\
&= c_1\|\xi\|_1^2,
\end{aligned}$$

where

$$c_1 = \max\left\{\|e^z\|_\infty^2 + 2\|\nabla e^z\|_\infty^2, 2\|e^z\|_\infty^2\right\}$$

Let $\phi = e^z\xi$ be an arbitrary function in $\hat{\mathcal{U}}_{_L}$. Then,

$$\begin{aligned}
\|\phi\|_1^2 &= \|e^z\xi\|_1^2 \\
&\leq c_1\|\xi\|_1^2 \\
&\leq c_2\|\mathcal{B}_k\xi\|_0^2 \\
&= c_2\|e^{-z}e^z\mathcal{B}_k\xi\|_0^2 \\
&\leq c_2\|e^{-z}\|_\infty^2\|e^z\mathcal{B}_k\xi\|_0^2 \\
&= c_3\|e^z\mathcal{B}_k e^{-z}\left(e^z\xi\right)\|_0^2 \\
&= c_3\|e^z\mathcal{B}_k e^{-z}\phi\|_0^2 \\
&= c_3\|\mathcal{B}\phi\|_0^2,
\end{aligned}$$

where

$$c_3 = c_1\, c_k\, \|e^z\|_\infty^2.$$

Thus, $\mathcal{B}$ is coercive on $\hat{\mathcal{U}}_{L}$, with coercivity constant

$$c_{L} \;=\; c_{k}\left[\left(\|e^{-z}\|_{\infty}^{2}\right)\max\left\{\|e^{z}\|_{\infty}^{2}+2\|\nabla e^{z}\|_{\infty}^{2},2\|e^{z}\|_{\infty}^{2}\right\}\right]^{-1}.$$

□

Lemma 5.2.4. *Let $\mathcal{U}_{R}$ be a subspace of $\mathcal{V}_{\mathbb{C}}$, and define $\hat{\mathcal{U}}_{R}=e^{-\bar{z}}\mathcal{U}_{R}$ and $\mathcal{B}^{*}=e^{-\bar{z}}\mathcal{B}_{k}^{*}e^{\bar{z}}$, where $\mathcal{B}^{*}=-\nabla_{x}-i\nabla_{y}$ and $\mathcal{B}_{k}^{*}=-\left(\partial_{x}-ik_{1}\right)-i\left(\partial_{y}-ik_{2}\right)$. If there exists $c_{k}^{*}>0$ such that*

$$\|\mathcal{B}_{k}^{*}\xi\|_{0}^{2} \;\geq\; c_{k}^{*}\|\xi\|_{1}^{2} \quad \forall\xi\in\mathcal{U}_{R},$$

then there exists $c_{R}>0$ such that

$$\|\mathcal{B}^{*}\xi\|_{0}^{2} \;\geq\; c_{R}\|\xi\|_{1}^{2} \quad \forall\xi\in\hat{\mathcal{U}}_{R}.$$

Proof. The proof is similar to that of Lemma 5.2.3. The resulting coercivity constant is given by

$$c_{R} \;=\; c_{k}^{*}\left[\left(\|e^{\bar{z}}\|_{\infty}^{2}\right)\max\left\{\|e^{-\bar{z}}\|_{\infty}^{2}+2\|\nabla e^{-\bar{z}}\|_{\infty}^{2},2\|e^{-\bar{z}}\|_{\infty}^{2}\right\}\right]^{-1}.$$

□

Lemma 5.2.5. *If $\mathcal{A}$ is not an exceptional configuration then $\mathcal{B}_{k}$ is coercive on $\mathcal{U}_{L}:=\mathcal{V}_{\mathbb{C}}$. That is, there exists $c_{k}>0$ such that*

$$\|\mathcal{B}_{k}\xi\|_{0}^{2} \;\geq\; c_{k}\|\xi\|_{1}^{2},$$

for all $\xi \in \mathcal{U}_{_L}$*, where* $\mathcal{U}_{_L} := \mathcal{V}_{\mathbb{C}}$*. If* $\mathcal{A}$ *is an exceptional configuration then* $\mathcal{B}_k$ *is coercive on* $\mathcal{U}_{_L} := \mathcal{N}(\mathcal{B}_k)^{\perp}$*. That is, there exists* $c_k > 0$ *such that*

$$\|\mathcal{B}_k \xi\|_0^2 \quad \geq \quad c_k \|\xi\|_1^2,$$

for all $\xi \in \mathcal{U}_{_L}$*, where* $\mathcal{U}_{_L} := \mathcal{N}(\mathcal{B}_k)^{\perp}$*.*

Proof. We begin by attempting to show that $\mathcal{B}_k$ is coercive on all of $\mathcal{V}_{\mathbb{C}}$. When the gauge field is an exceptional configuration (that is, k_1 and k_2 are integer multiples of 2π), we will need to restrict $\mathcal{U}_{_L}$ to the orthogonal complement of the nullspace of $\mathcal{B}_k$. For now, though, we take $\mathcal{U}_{_L} = \mathcal{V}_{\mathbb{C}}$ and proceed until a problem becomes apparent. $\mathcal{V}_{\mathbb{C}}$ is the space of periodic, complex-valued functions in $H^1(\mathcal{R})$. As such, we can rewrite

$$\mathcal{U}_{_L} \quad = \quad \text{Span}\left\{ e^{i2\pi(rx+sy)} \text{ s.t. } r, s \in \mathbb{Z} \right\}.$$

Consider the action of $\mathcal{B}_k$ on a single basis element of $\mathcal{U}_{_L}$. That is, let $\phi_{rs} = \alpha_{rs} e^{i2\pi(rx+sy)} \in \mathcal{U}_{_L}$ for some $r, s \in \mathbb{Z}$. Note that α_{rs} is a normalizing constant chosen so that $\|\phi_{rs}\|_0 = 1$. Then

$$\begin{aligned} \mathcal{B}_k \phi_{rs} &= [(\partial_x - ik_1) - i(\partial_y - ik_2)]\,\phi_{rs} \\ &= [(i2\pi r - ik_1) - i(i2\pi s - ik_2)]\,\phi_{rs} \\ &= [i(2\pi r - k_1) + (2\pi s - k_2)]\,\phi_{rs}. \end{aligned}$$

Thus, ϕ_{rs} is an eigenvector of $\mathcal{B}_k$ with associated eigenvalue

$$\lambda_{rs} \quad = \quad [i(2\pi r - k_1) + (2\pi s - k_2)].$$

For any $\phi_{rs} = \alpha_{rs} e^{i2\pi(rx+sy)} \in \mathcal{U}_{_L}$, we have

$$\begin{aligned}\|\mathcal{B}_k\phi_{rs}\|_0^2 &= \|\left[i\left(2\pi r-k_1\right)+\left(2\pi s-k_2\right)\right]\phi_{rs}\|_0^2\\ &= |i\left(2\pi r-k_1\right)+\left(2\pi s-k_2\right)|^2\|\phi_{rs}\|_0^2\\ &= \left(2\pi r-k_1\right)^2+\left(2\pi s-k_2\right)^2.\end{aligned}$$

Furthermore,

$$\begin{aligned}\|\phi_{rs}\|_1^2 &= \|\phi_{rs}\|_0^2+\|\nabla\phi_{rs}\|_0^2\\ &= \|\phi_{rs}\|_0^2+\|\partial_x\phi_{rs}\|_0^2+\|\partial_y\phi_{rs}\|_0^2\\ &= 1+4\pi^2\left(r^2+s^2\right).\end{aligned}$$

Then, clearly

$$\begin{aligned}\frac{\|\mathcal{B}_k\phi\|_0^2}{\|\phi\|_1^2} &\geq \min_{r,s\in\mathbb{Z}}\frac{\left(2\pi r-k_1\right)^2+\left(2\pi s-k_2\right)^2}{1+4\pi^2\left(r^2+s^2\right)} &(5.44)\\ &:= c_k. &(5.45)\end{aligned}$$

That is, the minimum of the ratio in (5.44) occurs when ϕ is some eigenvector of $\mathcal{B}_k$.

It is desirable to have an estimate for the value of c_k. To find a lower bound on c_k, parameters r and s are allowed to take on continuous values and the minimum over $(r,s)\in\mathbb{R}^2$ of the following function is considered:

$$f\left(r,s\right) = \frac{\left(2\pi r-k_1\right)^2+\left(2\pi s-k_2\right)^2}{1+4\pi^2\left(r^2+s^2\right)}.$$

Then, since $\phi\in\mathcal{U}_{_L}$, a lower bound on c_k is obtained by taking the minimizing r and s to be integers near the true minimum of $f(r,s)$. First, notice that $f(r,s)\geq 0$ for all r and s, and that $f(r,s)$ has only one root at

$$(r_+, s_+) \quad = \quad \left(\frac{k_1}{2\pi}, \frac{k_2}{2\pi}\right).$$

Without loss of generality, assume that k_1 and k_2 are both positive. Using a calculus argument, it is easy to see that, in the first quadrant (in general, the quadrant containing (r_+, s_+)), $f(r, s)$ is monotonically increasing along any trajectory moving away from (r_+, s_+). This is illustrated by the cross-section of $f(r, s)$ with $s = 0$ and $k_1 = 3$, given in Figure 5.8.

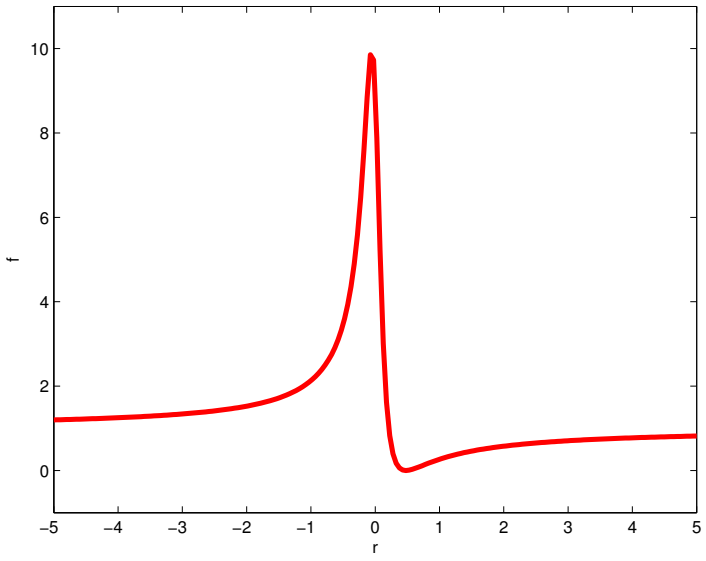


Figure 5.8: A cross-section of $f(r, s)$ with $s = 0$ and $k_1 = 3$.

It is clear then that the integers that minimize $f(r, s)$ are near (r_+, s_+). Care must be taken, though, when $k_1 \approx 2\pi r$ and $k_2 \approx 2\pi s$ for some integers r and s. In this case, $\mathcal{B}_k$ is nearly singular, and c_k is very small. Furthermore, if the gauge field is an exceptional configuration, $\mathcal{B}_k$ is singular with nullspace vector $\phi_0 = e^{i(k_1 x + k_2 y)}$. This case must be handled separately from the case when $\mathcal{B}_k$ is nonsingular.

Case 1: $\mathcal{B}_k$ Nonsingular

Consider the square in the rs-plane given by $r_m \leq r \leq r_M$, $s_m \leq s \leq s_M$, where

$$r_m = \left\lfloor \frac{k_1}{2\pi} \right\rfloor \qquad r_M = \left\lceil \frac{k_1}{2\pi} \right\rceil$$
$$s_m = \left\lfloor \frac{k_2}{2\pi} \right\rfloor \qquad s_M = \left\lceil \frac{k_2}{2\pi} \right\rceil ,$$

and assume that (r_+, s_+) is bounded away from the corners of the square (see Figure 5.9). Here, $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor and ceiling operators, respectively.

For simplicity, write k_1 and k_2 as

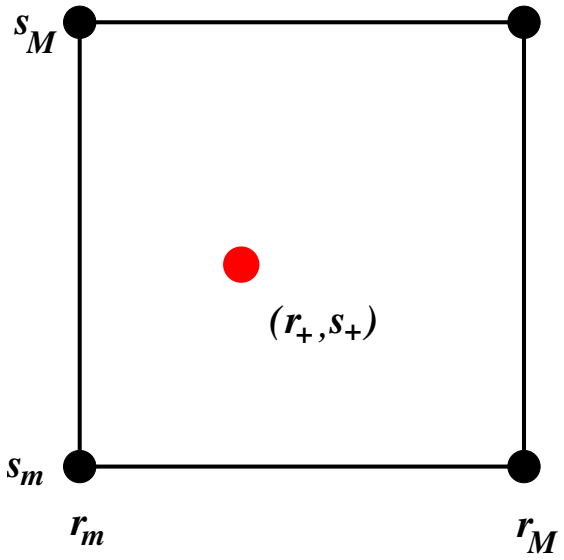


Figure 5.9: Constants k_1 and k_2 bounded away from 2π.

$$k_1 = 2\pi\,(r_m + \delta r) \qquad k_2 = 2\pi\,(s_m + \delta s)\,,$$

where $0 < \delta r < 1$ and $0 < \delta s < 1$. Then, the value of $f(r, s)$ at each of the four corners of the square is given by

$$
\begin{aligned}
f(r_m, s_m) &= \frac{4\pi^2\delta r^2 + 4\pi^2\delta s^2}{1 + 4\pi^2 r_m^2 + 4\pi^2 s_m^2} \\
f(r_M, s_m) &= \frac{4\pi^2 (1-\delta r)^2 + 4\pi^2\delta s^2}{1 + 4\pi^2 (r_m+1)^2 + 4\pi^2 s_m^2} \\
f(r_m, s_M) &= \frac{4\pi^2\delta r^2 + 4\pi^2 (1-\delta s)^2}{1 + 4\pi^2 r_m^2 + 4\pi^2 (s_m+1)^2} \\
f(r_M, s_M) &= \frac{4\pi^2 (1-\delta r)^2 + 4\pi^2 (1-\delta s)^2}{1 + 4\pi^2 (r_m+1)^2 + 4\pi^2 (s_m+1)^2}
\end{aligned}
$$

From these expressions, it is clear the the size of $f(r,s)$ on the corners of the box depend on both the size of δr and δs, as well as the size of r_m and s_m (and, by extension, the size of k_1 and k_2). In the case when k_1 and k_2 are large, the integer pair that minimizes $f(r,s)$ occurs at the point closest to (r_+, s_+). However, in the case when k_1 and k_2 are small, it is possible that the minimizing corner point will *not* be the one closest to (r_+, s_+). In any case, a lower bound for $f(r,s)$ can be established. That is,

$$
c_k := \min_{r,s\in\mathbb{Z}} f(r,s) = \geq \frac{4\pi^2 (\delta r^*)^2 + 4\pi^2 (\delta s^*)^2}{1 + 4\pi^2 \left[(r_m+1)^2 + (s_m+1)^2\right]}, \tag{5.46}
$$

where

$$
\begin{aligned}
\delta r^* &= \min\{\delta r, |1-\delta r|\} \\
\delta s^* &= \min\{\delta s, |1-\delta s|\}.
\end{aligned}
$$

In the case when k_1 and k_2 are reasonably small (on the order of 10) and (r_+, s_+) is bounded away from the corners of the square, the coercivity constant is reasonably sized. However, in the case that k_1 or k_2 is very large, c_k may be very small, since

the numerator is bounded above by 8π and the denominator can grow without bound. Thankfully, it is our experience that the constant portions of gauge fields generated using standard methods tend to be fairly small (< 15 for all of our test cases).

Case 2: $\mathcal{B}_k$ Singular

Let $k_1 = 2\pi r_0$ and $k_2 = 2\pi s_0$ for some integers r_0 and s_0. In this case, $\mathcal{B}_k$ is singular on $\mathcal{V}_{\mathbb{C}}$, and, thus, not coercive. In this case, let $\mathcal{U}_{L}$ be the orthogonal complement of $\phi_0 = e^{i(k_1 x + k_2 y)}$ in $\mathcal{V}_{\mathbb{C}}$. That is,

$$\mathcal{U}_{L} = \left\{ \xi \in \mathcal{V}_{\mathbb{C}} \text{ s.t. } \left\langle \xi, e^{i(k_1 x + k_2 y)} \right\rangle = 0 \right\},$$

or, in terms of Fourier modes,

$$\mathcal{U}_{L} = \text{Span} \left\{ e^{i2\pi(rx+sy)} \text{ s.t. } r, s \in \mathbb{Z} \text{ and } r \neq \frac{k_1}{2\pi}, s \neq \frac{k_2}{2\pi} \right\}.$$

In this case, the integer minimizers are located on the boundary of the larger square in the rs-plane given by $r_m \leq r \leq r_M$, $s_m \leq s \leq s_M$, where

$$\begin{aligned} r_m &= \frac{k_1}{2\pi} - 1, \qquad r_M = \frac{k_1}{2\pi} + 1, \\ s_m &= \frac{k_2}{2\pi} - 1, \qquad s_M = \frac{k_2}{2\pi} + 1. \end{aligned}$$

This boundary contains eight potential minimizers of $f(r, s)$ (See Figure 5.10). Again, the minimizing boundary point varies depending on the relative size of k_1 and k_2. A lower bound on c_k is given by

$$c_k := \min_{r,s \in \mathbb{Z}, (r,s) \neq (r_+, s_+)} f(r, s) = \geq \frac{4\pi^2}{1 + 4\pi^2 \left[(r_m + 1)^2 + (s_m + 1)^2 \right]}, \quad (5.47)$$

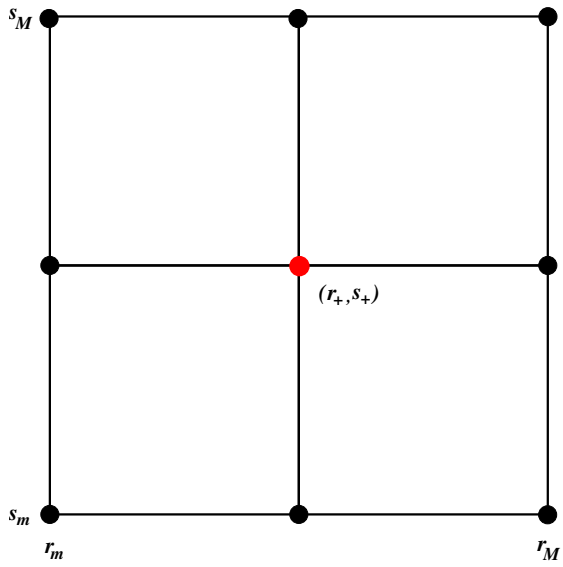


Figure 5.10: Constants k_1 and k_2 integer multiples of 2π.

Again, if k_1 and k_2 are of reasonable size, then so too is c_k.

Case 3: $\mathcal{B}_k$ Nearly Singular

Here we reexamine the case when $\mathcal{B}_k$ is nonsingular, but $k_1 \approx 2\pi r_0$ and $k_2 \approx 2\pi s_0$, for integers r_0 and s_0, to develop a more precise result. Consider the lower bound on c_k given in (5.46). It is clear that, as k_1 and k_2 approach integer multiples of 2π, c_k approaches 0. Restricting $\mathcal{U}_L$ to the orthogonal complement of $\phi_0 = e^{i2\pi(r_0 x + s_0 y)}$ in $\mathcal{V}_{\mathbb{C}}$, then the integers that minimize $f(r,s)$ lie on the boundary of the square in the rs-plane given by $r_m \le r \le r_M$, $s_m \le s \le s_M$, where

$$r_m = r_0 - 1, \qquad r_M = r_0 + 1,$$
$$s_m = s_0 - 1, \qquad s_M = s_0 + 1.$$

The boundary again contains eight potential minimizers (see Figure 5.11). In this case, a lower bound for c_k is given by

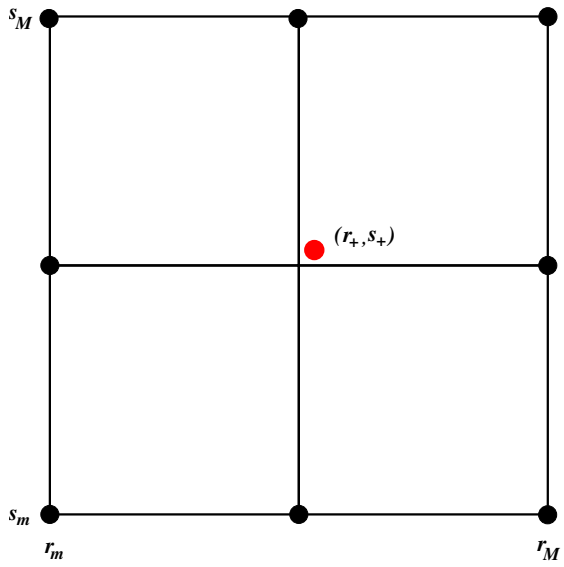


Figure 5.11: Constants k_1 and k_2 approaching integer multiples of 2π.

$$c_k := \min_{r,s\in\mathbb{Z},(r,s)\neq(r_0,s_0)} f(r,s) = \geq \frac{4\pi^2 (\delta^*)^2}{1+4\pi^2\left[(r_m+1)^2+(s_m+1)^2\right]}, \quad (5.48)$$

where

$$(\delta^*)^2 = \min\left\{\delta r^2+(1-\delta s)^2, \delta s^2+(1-\delta r)^2, (1-\delta r)^2+(1-\delta s)^2\right\}. \quad (5.49)$$

Then, there exists $c_k > 0$ such that

$$\|\mathcal{B}_k\xi\|_0^2 \geq c\|\xi\|_1^2 \quad \forall \xi \in \mathcal{U}_L,$$

where c_k is bounded from below by one of the expressions given in (5.46)-(5.48) above.

□

Lemma 5.2.6. *If $\mathcal{A}$ is not an exceptional configuration then $\mathcal{B}_k^*$ is coercive on $\mathcal{U}_R := \mathcal{V}_{\mathbb{C}}$. That is, there exists $c_k^* > 0$ such that*

$$\|\mathcal{B}_k^*\xi\|_0^2 \;\geq\; c_k^*\|\xi\|_1^2,$$

for all $\xi \in \mathcal{U}_{_R}$*, where* $\mathcal{U}_{_R} := \mathcal{V}_{\mathbb{C}}$*. If* $\mathcal{A}$ *is an exceptional configuration then* $\mathcal{B}_k^*$ *is coercive on* $\mathcal{U}_{_R} := \mathcal{N}\,(\mathcal{B}_k^*)^{\perp}$*. That is, there exists* $c_k^* > 0$ *such that*

$$\|\mathcal{B}_k^*\xi\|_0^2 \;\geq\; c_k^*\|\xi\|_1^2,$$

for all $\xi \in \mathcal{U}_{_R}$*, where* $\mathcal{U}_{_R} := \mathcal{N}\,(\mathcal{B}_k^*)^{\perp}$*.*

Proof. The proof is very similar to that of Lemma 5.2.5. In fact, the lower bounds on c_k are valid for c_k^* as well. □

Theorem 5.2.7. *If* $m > 0$ *or* $\mathcal{A}$ *is not an exceptional configuration then there exist positive constants* $c_{_\mathcal{A}}$ *and* $C_{_\mathcal{A}}$*, which depend on the gauge field,* $\mathcal{A}$*, such that,*

$$c_{_\mathcal{A}}\|\psi\|_1^2 \;\leq\; G(\psi,\mathcal{A};0) \;\leq\; C_{_\mathcal{A}}\|\psi\|_1^2,$$

for all $\psi \in \mathcal{U}$*, where* $\mathcal{U} := \mathcal{V}_{\mathbb{C}}^2$*. If* $m = 0$ *and* $\mathcal{A}$ *is an exceptional configuration then there exist positive constants* $c_{_\mathcal{A}}$ *and* $C_{_\mathcal{A}}$*, which depend on the gauge field,* $\mathcal{A}$*, such that,*

$$c_{_\mathcal{A}}\|\psi\|_1^2 \;\leq\; G(\psi,\mathcal{A};0) \;\leq\; C_{_\mathcal{A}}\|\psi\|_1^2,$$

for all $\psi \in \mathcal{U}$*, where* $\mathcal{U} := \mathcal{N}\,(\mathcal{D})^{\perp}$*.*

Proof. The upper-bound is verified first. By repeated use of the triangle and Cauchy's inequalities,

$$
\begin{aligned}
G(\psi,\mathcal{A};0) &= \|\mathcal{D}\psi\|_0^2 \\
&= \|m\psi_{_R} + (\nabla_x - i\nabla_y)\,\psi_{_L}\|_0^2 + \|m\psi_{_L} + (\nabla_x + i\nabla_y)\,\psi_{_R}\|_0^2 \\
&\leq 2\left(\|m\psi_{_R}\|_0^2 + \|\nabla_x\psi_{_L}\|_0^2 + \|\nabla_y\psi_{_L}\|_0^2\right. \\
&\qquad \left. + \|m\psi_{_L}\|_0^2 + \|\nabla_x\psi_{_R}\|_0^2 + \|\nabla_y\psi_{_R}\|_0^2\ \right) \\
&= 2\left(\|m\psi_{_R}\|_0^2 + \|\left(\partial_x - \mathcal{A}_1\right)\psi_{_L}\|_0^2 + \|\left(\partial_y - \mathcal{A}_2\right)\psi_{_L}\|_0^2\right. \\
&\qquad \left. + \|m\psi_{_L}\|_0^2 + \|\left(\partial_x - \mathcal{A}_1\right)\psi_{_R}\|_0^2 + \|\left(\partial_y - \mathcal{A}_2\right)\psi_{_R}\|_0^2\ \right) \\
&\leq 2\left[\|m\psi_{_R}\|_0^2 + 2\left(\|\partial_x\psi_{_L}\|_0^2 + \|\mathcal{A}_1\psi_{_L}\|_0^2 + \|\partial_y\psi_{_L}\|_0^2 + \|\mathcal{A}_2\psi_{_L}\|_0^2\right)\right. \\
&\qquad \left. + \|m\psi_{_L}\|_0^2 + 2\left(\|\partial_x\psi_{_R}\|_0^2 + \|\mathcal{A}_1\psi_{_R}\|_0^2 + \|\partial_y\psi_{_R}\|_0^2 + \|\mathcal{A}_2\psi_{_R}\|_0^2\right)\ \right] \\
&\leq 2\left[\|m\psi_{_R}\|_0^2 + \|m\psi_{_L}\|_0^2 + 2\left(\|\partial_x\psi_{_R}\|_0^2 + \|\partial_y\psi_{_R}\|_0^2 + \|\partial_x\psi_{_L}\|_0^2 + \|\partial_y\psi_{_L}\|_0^2\right)\right. \\
&\qquad \left. + 2\left(\|\mathcal{A}_1\|_\infty^2\|\psi_{_R}\|_0^2 + \|\mathcal{A}_2\|_\infty^2\|\psi_{_R}\|_0^2 + \|\mathcal{A}_1\|_\infty^2\|\psi_{_L}\|_0^2 + \|\mathcal{A}_2\|_\infty^2\|\psi_{_L}\|_0^2\right)\right] \\
&\leq \left[2m^2 + 4\left(\|\mathcal{A}_1\|_\infty^2 + \|\mathcal{A}_2\|_\infty^2\right)\right]\|\psi\|_0^2 + 4\|\nabla\psi\|_0^2 \\
&\leq C_{\mathcal{A}}\|\psi\|_1^2,
\end{aligned}
$$

where

$$
C_{\mathcal{A}} = \max\left\{2m^2 + 4\left(\|\mathcal{A}_1\|_\infty^2 + \|\mathcal{A}_2\|_\infty^2\right), 4\right\}.
$$

From the proof of Lemma 5.2.5 it is clear that, when verifying the coercivity condition, care must be taken when the gauge field is an exceptional configuration. Clearly, if $\mathcal{B}_k$ is singular, so too will be the massless Dirac operator, $\mathcal{D}_0$. As such, the cases when $\mathcal{B}_k$ is nonsingular and when $\mathcal{B}_k$ is singular are handled separately.

Case 1: $\mathcal{B}_k$ Nonsingular

Assume that k_1 and k_2 are bounded comfortably away from integer multiples of 2π. Then, define $\mathcal{U} := \mathcal{V}_{\mathbb{C}}^2$ and $\mathcal{U}_{_L} := \mathcal{U}_{_R} := \mathcal{V}_{\mathbb{C}}$. From Lemmas 5.2.5 and 5.2.6, there exist positive constants c_k and c_k^* such that

$$\begin{aligned} \|\mathcal{B}_k\psi_{_L}\|_0^2 &\geq c_k\|\psi_{_L}\|_1^2 \quad \forall\psi_{_L}\in\mathcal{U}_{_L} \\ \|\mathcal{B}_k^*\psi_{_R}\|_0^2 &\geq c_k^*\|\psi_{_R}\|_1^2 \quad \forall\psi_{_R}\in\mathcal{U}_{_R} \end{aligned}$$

Note now that, for any periodic function, z,

$$\begin{aligned} \hat{\mathcal{U}}_{_R} = e^{-\bar{z}}\mathcal{U}_{_R} &= \mathcal{V}_{\mathbb{C}}, \\ \hat{\mathcal{U}}_{_L} = e^{z}\mathcal{U}_{_L} &= \mathcal{V}_{\mathbb{C}}. \end{aligned}$$

Thus, from Lemmas 5.2.3 and 5.2.4, there exist positive constants $c_{_L}$ and $c_{_R}$ such that

$$\begin{aligned} \|\mathcal{B}\psi_{_L}\|_0^2 &\geq c_{_L|}\|\psi_{_L}\|_1^2 \quad \forall\psi_{_L}\in\mathcal{U}_{_L}, \\ \|\mathcal{B}^*\psi_{_R}\|_0^2 &\geq c_{_R}\|\psi_{_R}\|_1^2 \quad \forall\psi_{_R}\in\mathcal{U}_{_R}. \end{aligned}$$

Then, for all $\psi\in\mathcal{U}$,

$$\begin{aligned} \|\mathcal{D}_0\psi\|_0^2 &= \left\| \begin{bmatrix} 0 & \mathcal{B} \\ -\mathcal{B}^* & 0 \end{bmatrix} \begin{bmatrix} \psi_{_R} \\ \psi_{_L} \end{bmatrix} \right\|_0^2 \\ &= \|\mathcal{B}\psi_{_L}\|_0^2 + \|\mathcal{B}^*\psi_{_R}\|_0^2 \\ &\geq c_{_L}\|\psi_{_L}\|_1^2 + c_{_R}\|\psi_{_R}\|_1^2 \\ &\geq c_0\|\psi\|_1^2, \end{aligned}$$

where

$$c_0 = \min\{c_{_R}, c_{_L}\}.$$

Finally, from the coercivity of $\mathcal{D}_0$ on $\mathcal{U}$ and Lemma 5.2.2, there exists some positive constant $c_{\mathcal{A}}$ such that

$$\|\mathcal{D}\psi\|_0^2 \;\geq\; c_{\mathcal{A}}\|\psi\|_1^2,$$

where

$$\begin{aligned} c_{\mathcal{A}} \;=\; \min\Big\{ & c_k\left[\left(\|e^{-z}\|_\infty^2\right)\max\left\{\|e^{z}\|_\infty^2+2\|\nabla e^{z}\|_\infty^2, 2\|e^{z}\|_\infty^2\right\}\right]^{-1}, \\ & c_k^*\left[\left(\|e^{\bar{z}}\|_\infty^2\right)\max\left\{\|e^{-\bar{z}}\|_\infty^2+2\|\nabla e^{-\bar{z}}\|_\infty^2, 2\|e^{-\bar{z}}\|_\infty^2\right\}\right]^{-1}\Big\}, \end{aligned}$$

and c_k and c_k^* are both bounded from below by (5.46).

Case 2: $\mathcal{B}_k$ Singular

In the case that the gauge field is an exceptional configuration, the problem is formulated in the orthogonal complement of $\phi_0 = e^{i(k_1x+k_2y)}$. Define both $\mathcal{U}_{_L}$ and $\mathcal{U}_{_R}$ according to

$$\mathcal{U}_{_L} = \mathcal{U}_{_R} \;:=\; \left\{\xi \in \mathcal{V}_{\mathbb{C}} \text{ s.t. } \left\langle \xi, e^{i(k_1x+k_2y)}\right\rangle = 0\right\}.$$

Again, from Lemmas 5.2.5 and 5.2.6, there exist positive constants $c_{_L}$ and $c_{_R}$ such that

$$\begin{aligned} \|\mathcal{B}_k\psi_{_L}\|_0^2 &\geq c_{_{L|}}\|\psi_{_L}\|_1^2 \quad \forall\psi_{_L}\in\mathcal{U}_{_L} \\ \|\mathcal{B}_k^*\psi_{_R}\|_0^2 &\geq c_{_R}\|\psi_{_R}\|_1^2 \quad \forall\psi_{_R}\in\mathcal{U}_{_R}. \end{aligned}$$

Define $\hat{\mathcal{U}}_{_L}$ and $\hat{\mathcal{U}}_{_R}$ according to

$$\begin{aligned} \hat{\mathcal{U}}_{_L} &:= \left\{\xi \in \mathcal{V}_{\mathbb{C}} \text{ s.t. } \left\langle \xi, e^{z+i(k_1x+k_2y)}\right\rangle = 0\right\} \\ \hat{\mathcal{U}}_{_R} &:= \left\{\xi \in \mathcal{V}_{\mathbb{C}} \text{ s.t. } \left\langle \xi, e^{-\bar{z}+i(k_1x+k_2y)}\right\rangle = 0\right\}. \end{aligned}$$

Then, from Lemmas 5.2.3 and 5.2.4, there exist positive constants $\hat{c}_{_L}$ and $\hat{c}_{_R}$ such that

$$\begin{aligned} \|\mathcal{B}\psi_{_L}\|_0^2 &\geq \hat{c}_{_L|}\|\psi_{_L}\|_1^2 \quad \forall\psi_{_L} \in \hat{\mathcal{U}}_{_L}, \\ \|\mathcal{B}^*\psi_{_R}\|_0^2 &\geq \hat{c}_{_R}\|\psi_{_R}\|_1^2 \quad \forall\psi_{_R} \in \hat{\mathcal{U}}_{_R}. \end{aligned}$$

Finally, define $\mathcal{U}$ by

$$\mathcal{U} = \hat{\mathcal{U}}_{_R} \otimes \hat{\mathcal{U}}_{_L}.$$

The remainder of the proof is identical to that in Case 1.

Case 3: $\mathcal{B}_k$ Nearly Singular

In the case that the gauge field is near an exceptional configuration, that is, $k_1 \approx 2\pi r_0$ and $k_2 \approx 2\pi s_0$ for some integers r_0 and s_0, the problem is formulated in the orthogonal complement of $\phi_0 = e^{i2\pi(r_0x+s_0y)}$. Define both $\mathcal{U}_{_L}$ and $\mathcal{U}_{_R}$ according to

$$\mathcal{U}_{_L} = \mathcal{U}_{_R} := \left\{\xi \in \mathcal{V}_{\mathbb{C}} \text{ s.t. } \left\langle \xi, e^{i2\pi(r_0x+s_0y)} \right\rangle = 0\right\}.$$

Then, the remainder of the proof case is identical to Case 2.

□

5.2.2 Implications

H^1-ellipticity of the least-squares functional has two implications for the least-squares solution process. First, because the functional is H^1-elliptic, it can be shown that an optimal $\mathcal{O}(N)$ multilevel iterative method exists that can solve the resulting linear system [54].

Second, it suggests an alternate argument that the discrete least-squares operator, $\mathbb{D}_{LS}$, does not suffer from species doubling. The application of the least-squares operator involves multiplication by matrix $\mathbb{L}$. The absence of species doubling requires that this

application does not map oscillatory modes onto smooth modes. Recall that, from the coercivity condition on the least-squares functional,

$$\begin{aligned} \|\mathcal{D}\psi\|_0^2 &\geq c_{\mathcal{A}}\|\psi\|_1^2 \\ &= c_{\mathcal{A}}\left(\|\psi\|_0^2 + \|\nabla\psi\|_0^2\right) \\ &\geq c_{\mathcal{A}}\|\nabla\psi\|_0^2. \end{aligned}$$

Thus,

$$\frac{\|\mathcal{D}\psi\|_0^2}{\|\nabla\psi\|_0^2} \geq c_{\mathcal{A}} \quad \forall\psi \in \mathcal{V}_{\mathbb{C}}^2. \tag{5.50}$$

Recall, from (5.32), that

$$\|\mathcal{D}\psi^h\|_0^2 = \left\langle \mathbb{L}\underline{\psi}, \underline{\psi}\right\rangle_{l^2}. \tag{5.51}$$

Similarly,

$$\|\nabla\psi^h\|_0^2 = \left\langle \mathbb{A}\underline{\psi}, \underline{\psi}\right\rangle_{l^2}. \tag{5.52}$$

where $\mathbb{A}$ is the standard Galerkin discretization of the constant coefficient Laplacian. Then, combining (5.50) - (5.52) yields a discrete coercivity condition

$$\frac{\left\langle \mathbb{L}\underline{\psi}, \underline{\psi}\right\rangle}{\left\langle \mathbb{A}\underline{\psi}, \underline{\psi}\right\rangle} \geq c_{\mathcal{A}} \quad \forall\underline{\psi} \in \mathcal{N}_{\mathbb{C}}^2. \tag{5.53}$$

Clearly, the discrete Laplacian, $\mathbb{A}$, maps oscillatory modes to the high end of the spectrum. The coercivity bound, (5.53), implies that $\mathbb{L}$ must map oscillatory modes to the high end of the spectrum as well. Thus, the application of $\mathbb{L}$ cannot result in species doubling.

5.3 Numerical Experiments

In this section, the use of classical algebraic multigrid (AMG) is investigated as a preconditioner for the solution of the linear system appearing in Step 3 of Algorithm 5.3. That is,

$$\mathbb{L}\underline{\zeta} \;=\; \mathbb{K}\underline{g}. \tag{5.54}$$

The conjugate gradient (CG) algorithm is used to accelerate the solution process [33], [47]. Later, we compare the use of AMG preconditioned CG applied to (5.54) to adaptive smoothed aggregation (αSA) preconditioned CG, when applied to the Dirac-Wilson operator. A brief introductions to general multigrid theory is given. This is then extended to AMG and αSA.

5.3.1 Multigrid Methods

Multigrid methods are a large class of *iterative* methods used to solve linear systems of equations of the form

$$Ax = b. \tag{5.55}$$

Let x be the true solution of (5.55) and y be some approximations to x. Two quantities of interest are the residual and error of the approximation, given by

$$\begin{aligned} r &= b - Ay, \\ e &= x - y, \end{aligned}$$

respectively. Some simple algebra provides a relationship between the error and the residual. This relationship, called the *residual-error equation*, is given by

$$Ae = r.$$

Naturally, if the error of the current iterate was known exactly, then the true solution is obtained by correcting the current iterate according to

$$x = y + e.$$

The error is generally not known in practice. However, if an approximation to the current error is available, call it $\widetilde{e}$, it can be used to correct the current iterate. An iteration based on this correction would set

$$y \leftarrow y + \widetilde{e},$$

at each step in the iteration. In multigrid methods, a sequence of coarse grids are used to compute an approximation to the current error in an inexpensive manner.

At the heart of all multigrid methods are two complementary processes: *relaxation* and *coarse-grid correction*. Relaxation is a *local* process that effectively reduces some portion of the error in the current iterate. Error that is not effectively reduced by relaxation is called *algebraically smooth*. A general step in a relaxation method sets

$$y \leftarrow y + M^{-1}r, \tag{5.56}$$

where M is some approximation of the matrix A. The relaxation step can also be formulated in terms of its action on the error of the current approximation. That is, in a single step of relaxation, the error is updated according to

$$e \leftarrow e - M^{-1}Ae. \tag{5.57}$$

This provides a useful characterization of algebraically smooth error. Suppose the error is such that

$$M^{-1}Ae \approx 0,$$

relative to e. Then, the relaxation step given in (5.57) fails to reduce a significant portion of the error. Thus, algebraically smooth error is often said to be *near-kernel*.

Coarse-grid correction is a *global* process that uses a coarse grid to obtain an approximation to algebraically smooth error. This approximation of the error is then used to correct the current approximation. Suppose that the original grid, or the *fine* grid, has N points. Now, suppose some *coarse* grid exists with N_c points, where $N_c << N$. Let $P \in \mathbb{C}^{N \times N_c}$ be the so called *interpolation* operator, that maps objects from the coarse grid to the fine grid. The so called *restriction* operator maps objects on the fine grid to the coarse grid. There are many choices for the restriction operator. For the methods in this thesis, we take P^t to be the restriction operator, where P is the previously introduced interpolation operator. Now, suppose that some approximation to the fine-grid error is computed on the coarse grid; call it e_c. Then, a correction to the fine-grid approximation can be made by setting

$$y \leftarrow y + Pe_c.$$

Note that the error in this iteration is updated according to

$$e \leftarrow e - Pe_c. \tag{5.58}$$

The coarse-grid error is obtained by solving a coarse-grid version of the residual-error equation:

$$A_c\, e_c \;=\; r_c.$$

There are many choices for the coarse-grid operator, A_c. The methods employed in this thesis use a variational formulation of the coarse-grid operator given by

$$A_c \;=\; P^t A P.$$

The coarse-grid version of the residual, r_c, is simply the fine-grid residual restricted to the coarse grid. That is,

$$r_c \;=\; R\, r.$$

Then, the error update in the coarse-grid correction, (5.58), can be written as

$$e \;\leftarrow\; e - P(P^t A P)^{-1} P^t A e. \tag{5.59}$$

Now, suppose that the fine-grid error lies completely in the range of interpolation; that is, $e = Pe_c$, for some coarse-grid vector, e_c. Then,

$$\begin{aligned} e &\leftarrow e - P(RAP)^{-1} RAe \\ &\leftarrow Pe_c - P(RAP)^{-1} RAPe_c \\ &\leftarrow 0. \end{aligned}$$

Thus, the coarse-grid correction completely eliminates error in the range of interpolation. Of course, most fine-grid error will not have this quality. However, if interpolation is designed so that much of the algebraically smooth error is in its range, coarse-grid

correction can be used in tandem with relaxation to eliminate much of the error in the current iterate, at a relatively inexpensive cost. Using these ideas together defines the two-grid correction scheme, given in Algorithm 5.4 [21].

ALGORITHM 5.4: Two-Grid Correction Scheme

Input: Matrix A, right-hand side vector b, initial guess y_0.

Output: Approximation solution y.

1. Relax ν_1 times on the fine-grid problem, $Ay = b$.
2. Restrict residual to fine gird by $r_c = Rr$.
3. Solve coarse-grid problem, $A_c e_c = r_c$.
4. Correct iterate with interpolated error by $y \leftarrow y + Pe_c$.
5. Relax ν_2 times on the fine-grid problem, $Ay = b$.

In practice, two-level methods are not optimal because, if the initial grid is large, solving the coarse-grid problem via a direct method may be intractable. Thus, a multi-level method is obtained by applying the two-grid correction scheme recursively on the coarse level. Such a method is known as a *V-cycle*. If ν_1 and ν_2 are the number of pre- and post-relaxation steps, respectively, the method is known as a $V(\nu_1, \nu_2)$-*cycle*.

To recap, the V-cycle, as described above, is completely defined by the choice of relaxation schemes, interpolation operator P, and the number of pre- and post-relaxation steps, ν_1, and ν_2. The goal is to choose these objects so that relaxation leaves the algebraically smooth error in a subspace that can be well approximated by P. There are many different flavors of multigrid, each characterized by a different choice of P. For further details, an excellent reference for standard multigrid methods is [21]. In the remainder of this section, we *briefly* describe two such flavors, AMG, and αSA .

In AMG, the coarse grid points are chosen to be a small subset of the fine grid

points. (Here, *point* is used ambiguously. In general, a point refers to a single unknown associated with a specific spatial-point on the physical grid.) Furthermore, the points chosen for the coarse grid should be those that strongly influence many other points. This concept of *influence* is defined by looking at the rows of the matrix, and using a strength-of-connection measure to determine if the value of an unknown greatly affects the values of other unknowns. Let C and F be the collection of coarse- and fine-grid points, respectively. In general, let e characterize the algebraically smooth error. This can be known *a priori* or it can be identified by repeated relaxation on $Ax = 0$. Then, interpolation operator, P, is defined by its action on e. That is,

$$[\,Pe\,]_i \;=\; \begin{cases} e_i & \text{if } i \in C, \\ \displaystyle\sum_{j \in C_i} w_{ij} e_i & \text{if } i \in F, \end{cases} \tag{5.60}$$

where C_i is the set of coarse-grid points that strongly influence point x_i. The weights, w_{ij}, are chosen so that P represents the algebraically smooth error well. In general, e could be any algebraically smooth error components. In practice, e is usually chosen to be the constant vector. For a more in-depth description of this method, see [12] or [21].

The discussion of αSA begins with an introduction to smoothed aggregation multigrid (SA). Standard geometric multigrid methods, and standard AMG, are designed to be efficient when the algebraically smooth error is geometrically smooth as well. When the algebraically smooth error is *not* geometrically smooth, as is often the case in QED, these methods can perform poorly. SA was designed to remedy this. In SA, the coarse grid is not a subset of the fine grid. Instead, collections of points (or unknowns) are grouped together based on their relative strength of connection. These groups are called *aggregates*. Each aggregate then becomes a point on the coarse grid. Again, let e be a vector characterizing the algebraically smooth error (which is defined relative to the chosen relaxation scheme). Let $\mathcal{A}_j$ be the collection of fine-grid points in the j^{th} aggregate.

Then interpolation operator P is defined according to

$$[\, P \,]_{ij} \;=\; \begin{cases} e_i & \text{if } i \in \mathcal{A}_j, \\ 0 & \text{otherwise.} \end{cases} \tag{5.61}$$

Note that, with P defined in this manner, error vector e is *exactly* in the range of interpolation. Thus, a coarse-grid correction completely eliminates e. Often, this choice of P leads to an effective multigrid method. More often, though, e is only a characterization of *some* of the algebraically smooth error. If the remainder of the error is not captured well by P, then convergence of the method will be poor. To remedy this, the columns of P are *smoothed* by applying one step of relaxation to homogeneous equation with each column of P as the initial guess. (Note that the relaxation scheme used to smooth P is often different than the one used in the V-cycle. Usually weighted-Jacobi is chosen.) This smoothing allows P to capture algebraically smooth error vector e, as well as other error that is locally similar to e. Thus, smoothing enriches the space of vectors that can be well represented by P.

We have said nothing about how to obtain e. In many problems, such as linear elasticity, the algebraically smooth error modes are known *a priori*. In the case that they are not known, they can be obtained by relaxing on $Ax = 0$. This adaptive exposure of the algebraically smooth error is the basis for αSA. In the setup phase, relaxation is applied to the homogeneous problem until convergence stalls. Then, the resulting smooth error component is used to form P. Note that it is possible that the algebraically smooth error is spanned by more than one distinct smooth error vector. In this case, a V-cycle would effectively reduce error similar to the vector used to build P, while leaving the remaining error largely untouched; P can then be constructed to accurately interpolate more than one error component. Note that if P is built to handle two distinct error components, then P will have twice as many columns. This leads to

a coarse-grid operator that is twice as large.

To build an interpolation operator to handle more than one *prototype error component*, we first relax on $Ax = 0$ until convergence stalls. P is then constructed based on this component and used to define a V-cycle. Then, we iterate on $Ax = 0$ with the *V-cycle*. When this stalls, the resulting prototype error vector will be algebraically smooth with respect to the current *method*, instead of relaxation. P is then extended to accurately capture the new error component as well. This continues until convergence of the current method, applied to the homogeneous system of equations, is satisfactory. The setup phase is then exited and the resulting V-cycle is used to solve the target (inhomogeneous) problem. Note that this process leads to a much higher setup cost than that of AMG, where only the coarse-fine partitioning of the grid and the computation of the interpolation weights are needed. For a more in-depth description of SA and αSA, see [18] or [49].

Finally, we introduce some terminology that allows us to compare the computational cost of these multilevel methods. Let *operator complexity*, σ, be the ratio of total number of nonzeros in the operators on all grids in the multigrid hierarchy to the number of nonzeros in the fine-grid operator alone. Then σ gives an indication of how much work is required to preform a computation, such as one residual calculation, on all grid levels compared to performing that computation on just the fine grid [21]. Since a multigrid V(ν_1, ν_2)-cycle performs $(\nu_1 + \nu_2)$ relaxation steps on each grid level, and one residual calculation on each grid level, the operator complexity allows us to estimate the total cost of such a cycle relative to the cost of one relaxation step on the fine grid. Let one *work unit* (WU) be the cost of doing one matrix-vector multiply on the fine grid. Then, the cost of a single V(ν_1, ν_2)-cycle is approximately $\sigma(\nu_1 + \nu_2 + 1)$ WUs. Finally, define η to be the number of work units needed to improve the current iterate by one digit of accuracy. Then

$$\eta \quad = \quad \sigma\left(\nu_1 + \nu_2 + 1\right) \frac{\log .1}{\log \rho}, \tag{5.62}$$

where ρ is the convergence factor observed in the method [20].

5.3.2 Numerical Results

Linear system (5.54) clearly contains complex entries. To avoid working in complex arithmetic, the following equivalent real formulation (ERF) of (5.54) is solved instead:

$$\begin{bmatrix} \mathbb{X} & -\mathbb{Y} \\ \mathbb{Y} & \mathbb{X} \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix} \quad = \quad \begin{bmatrix} \underline{a} \\ \underline{b} \end{bmatrix}, \tag{5.63}$$

where $\mathbb{X}, \mathbb{Y}$ are real-valued matrices satisfying $\mathbb{L} = \mathbb{X} + i\mathbb{Y}$, $\underline{\zeta} = \underline{x} + i\underline{y}$, and $\mathbb{K}\underline{g} = \underline{a} + i\underline{b}$. Note that $\mathbb{Y}$ is skew-Hermitian so that (5.63) is a symmetric real system. Moreover, since the complex matrix is Hermitian positive semidefinite, the real system is symmetric positive semidefinite. In particular, we are interested in how the performance of AMG preconditioned CG (AMG-PCG) and αSA preconditioned CG (αSA-PCG) varies with fermion mass, m, gauge field temperature parameter, β, and lattice size, N.

In the following tests, AMG-PCG and αSA-PCG are applied to (5.63). For AMG-PCG, a single V(1,1)-cycle with Gauss-Seidel relaxation is used as the preconditioner in each step of CG. For αSA-PCG, a single V(2,2)-cycle used as the preconditioner is constructed using 2 prototype error components to build the interpolation operator. The relaxation method used is Gauss-Seidel. Aggregation is done algebraically. Each method is applied to (5.63) with a zero right-hand side and random initial guess. The iteration is terminated when the relative residual in the iteration has been decreased by a factor of 10^{-6}. Average convergence factors are computed by applying the method to operators constructed using 20 distinct gauge fields.

Table 5.1 reports average convergence factors for AMG preconditioned CG (AMG-PCG) and αSA preconditioned CG (αSA-PCG) for various values of particle mass m, gauge field temperature β, and lattice sizes N.

β/m	.001	.01	.1
2	.26	.26	.25
3	.27	.27	.28
5	.27	.27	.25

β/m	.001	.01	.1
2	.28	.26	.25
3	.28	.25	.27
5	.28	.28	.26

β/m	.001	.01	.1
2	.24	.28	.26
3	.28	.28	.26
5	.27	.27	.24

β/m	.001	.01	.1
2	.26	.26	.25
3	.25	.25	.23
5	.28	.25	.26

β/m	.001	.01	.1
2	.28	.20	.21
3	.27	.25	.25
5	.22	.24	.23

β/m	.001	.01	.1
2	.25	.24	.25
3	.25	.25	.23
5	.23	.24	.22

Table 5.1: Average convergence factors for AMG-PCG (left) and αSA-PCG (right) applied to (5.63) on 64×64 (top), 128×128 (middle), and 256×256 (bottom) lattices with varying choices of mass parameter m and temperature β. In all tests, operator complexity, σ, with AMG-PCG is approximately 1.8 and with αSA-PCG is approximately 1.2.

In Table 5.1, all tests with AMG-PCG have operator complexities of approximately 1.80, and with αSA-PCG have operator complexities of approximately 1.20. Note that, as mass parameter m decreases, the performance of both solvers is essentially unchanged. This is important because it means that the problem of critical slowing down has been eliminated. Also, the performance of the methods remain unchanged as the gauge field becomes more disordered. Finally, there is no decrease in performance as the size of the lattice grows. Thus, the two multigrid methods, applied to (5.63), appear to be scalable.

Table 5.1 seems to suggest that the performance of AMG-PCG and αSA-PCG on (5.63) are almost identical. However, we must take the computational cost of each

type of V-cycle used in the two methods into consideration. Recall that the AMG V(1,1)-cycle uses two fewer relaxation steps on each level than the V(2,2)-cycle used in αSA. On the other hand, the smaller operator complexity associated with αSA means that the relaxation and residual computations done in αSA are cheaper than those done in AMG. To accurately compare the performance of the two methods we look at the computational cost required by each to reduce the error in the current iteration by one digit of accuracy. Table 5.2 reports average η-values for each method applied to (5.63) on a 64×64 lattice, where η is defined in (5.62).

β/m	.001	.01	.1
2	9.2	9.2	9.0
3	9.5	9.5	9.8
5	9.5	9.5	9.0

β/m	.001	.01	.1
2	10.9	10.3	10.0
3	10.9	10.0	10.5
5	10.9	10.9	10.3

Table 5.2: Average η-values for AMG-PCG (left) and αSA-PCG (right) applied to (5.63) on a 64×64 lattice with varying choices of mass parameter m and temperature β.

From Table 5.2 we see that in overall accuracy per computational cost, AMG-PCG performs roughly 10% better than αSA-PCG. Furthermore, these comparisons do not take into consideration the setup cost of the two methods. Because αSA requires the computation of prototype error components the setup cost for αSA is inherently more expensive than for AMG. Although this is minimized in the present case because only 2 prototype error components are used, the αSA setup phase is still considerably more expensive than that of AMG.

It is interesting to compare the performance of an algebraic multigrid method on the discrete least-squares operator, $\mathbb{D}_{LS}$, and the traditional Dirac-Wilson operator, $\mathbb{D}_W$, given by (3.26). Since AMG-PCG performed better in our tests than αSA-PCG we will use it here. Unfortunately, AMG-PCG is not an effective solver for the Dirac-Wilson operator because of its rich near-kernel space [13]. Instead, αSA-PCG is used. The non-Hermiticity of $\mathbb{D}_W$ makes the problem difficult to solve with standard αSA

methods. As such, it is typical to solve the normal equations of (3.26) instead. That is,

$$\mathbb{D}_W^* \, \mathbb{D}_W \, \underline{\psi} \;\; = \;\; \mathbb{D}_W^* \, \underline{f}. \tag{5.64}$$

Again, (5.64) has complex entries, so its equivalent real formulation is solved instead. This is given by

$$\begin{bmatrix} \mathbb{U} & -\mathbb{V} \\ \mathbb{V} & \mathbb{U} \end{bmatrix} \begin{bmatrix} \underline{u} \\ \underline{v} \end{bmatrix} \;\; = \;\; \begin{bmatrix} \underline{c} \\ \underline{d} \end{bmatrix}, \tag{5.65}$$

where $\mathbb{U}$ and $\mathbb{V}$ are real-valued matrices satisfying $\mathbb{D}_W^* \, \mathbb{D}_W = \mathbb{U} + i\mathbb{V}$, $\underline{\psi} = \underline{u} + i\underline{v}$, and $\mathbb{D}_W^* \, \underline{f} = \underline{c} + i\underline{d}$. Table 5.3 compares the performance of AMG-PCG applied to (5.63) and αSA-preconditioned CG (αSA-PCG) applied to (5.65). Again, an AMG V(1,1)-cycle is used as the preconditioner for (5.63). For (5.65), an αSA V(2,2)-cycle is used, with 8 prototype error components used to construct the method.

β/m	.01	.1	.3
2	.26	.25	.23
3	.27	.28	.25
5	.27	.25	.25

β/m	.01	.1	.3
2	.33	.31	.31
3	.42	.40	.31
5	.31	.29	.28

Table 5.3: Average convergence factors for AMG-PCG applied to the least-squares formulation (left) and αSA-PCG applied to the normal equations of the Dirac-Wilson operator (right) on a 64×64 lattice with varying choices of mass parameter m and temperature β. In the least-squares case, operator complexity, σ, is approximately 1.8. In the Dirac-Wilson case, σ is approximately 3.0

Based on Table 5.3, AMG-PCG applied to the least-squares operator exhibits mildly better convergence than αSA-PCG applied to the Dirac-Wilson operator. However, the computational cost required in each cycle must be taken into consideration as well. The cycle based on (5.65) is more costly than that based on (5.63) for three

reasons. First, the operator in (5.65) has approximately 88% more nonzeros than that in (5.63). Thus, one work unit with respect to (5.65) involves approximately 1.88 times as much computation as a work unit with respect to (5.63). Second, the V(2,2)-cycle used for αSA-PCG requires 2 more relaxations on each level than the V(1,1)-cycle used for AMG. Finally, the reported operator complexity of the cycle for (5.65) is approximately 3.0, versus 1.8 for (5.63). To directly compare the efficiency of the two methods, η-values are computed for each, cost in terms of a work unit with respect to (5.63). For clarity, define η_{LS} and η_W by

$$\eta_{LS} = \sigma_{LS}\left(\nu_1+\nu_2+1\right)\frac{\log .1}{\log \rho_{LS}}, \tag{5.66}$$

$$\eta_W = 1.88\times\ \sigma_W\left(\nu_1+\nu_2+1\right)\frac{\log .1}{\log \rho_W}, \tag{5.67}$$

where σ_{LS} and ρ_{LS} are the operator complexity and convergence factor of the method applied to (5.63), respectively, and σ_W and ρ_W are the operator complexity and convergence factor of the method applied to (5.65), respectively. Then, η_W specifies the number of work units with respect to the least-squares operator that are needed to improve the accuracy of the current iterate by one digit.

Table 5.4 gives the previously define η-values for the results provided in Table 5.3. Furthermore, the ratio η_W/η_{LS} gives an estimate of the speedup obtained in using the least-squares discretization over the Dirac-Wilson discretization. These ratios are given in Table 5.5.

β/m	.01	.1	.3
2	9.2	9.0	8.5
3	9.5	9.8	9.0
5	9.5	9.0	9.0

β/m	.01	.1	.3
2	58.6	55.4	55.4
3	72.8	70.9	55.4
5	55.4	52.5	51.0

Table 5.4: Average η_{LS} and η_W-values for AMG-PCG applied to (5.63) and αSA-PCG applied to (5.65) on a 64×64 lattice with varying choices of mass parameter m and temperature β.

β/m	.01	.1	.3
2	6.3	6.2	6.6
3	7.9	7.3	6.2
5	5.9	5.8	5.7

Table 5.5: Average speedup factors for AMG-PCG applied to (5.63) over αSA-PCG applied to (5.65) on a 64×64 lattice with varying choices of mass parameter m and temperature β.

Table 5.5 indicates that AMG-PCG, with the specified cycle type, applied to the least-squares discretization attains between 5 and 8 times the accuracy per computational cost that αSA-PCG applied to the normal equations of the Dirac-Wilson discretization attains. Furthermore, using 8 prototype error components to build the V-cycle, the setup cast for αSA-PCG applied to the Dirac-Wilson operator is *much* greater than the setup cost of AMG.

Chapter 6

Least-Squares Finite Elements for a Transformed Schwinger Model

In this chapter a discrete approximation to the 2D Schwinger model is formulated by applying the least-squares methodology to the transformed system described in Section 2.2.3. Again, naively applying least-squares directly to the governing equations results in an algorithm that is not gauge covariant. To remedy this, the same gauge fixing concept used in Chapter 5 is applied. It is demonstrated that the resulting discrete solution process satisfies gauge covariance and chiral symmetry, and does not suffer from species doubling. Next, the least-squares functional for the transformed system is shown to be H^1-elliptic. Finally, numerical experiments are carried out, using adaptive smoothed aggregation multigrid, as well as classic algebraic multigrid, as preconditioners for solving the resulting linear system. The results show that the discretization of the transformed system can be solved very efficiently by algebraic multigrid methods.

6.1 Discretization of the Transformed System

Again, the first objective is to discretize the 2D Schwinger model, given by

$$\begin{bmatrix} mI & \mathcal{B} \\ -\mathcal{B}^* & mI \end{bmatrix} \begin{bmatrix} \psi_R \\ \psi_L \end{bmatrix} = \begin{bmatrix} f_R \\ f_L \end{bmatrix}, \tag{6.1}$$

where $\mathcal{B}(\mathcal{A}) = \nabla_x - i\nabla_y = (\partial_x - i\mathcal{A}_1) - i(\partial_y - i\mathcal{A}_2)$. Recall that, if the gauge field has the usual Helmholtz decomposition, $\mathcal{A} = \nabla^\perp u + \nabla v + \underline{k}$, then $\mathcal{B}$ and $\mathcal{B}^*$ transform

according to

$$\mathcal{B}(\mathcal{A})\, e^{z} = e^{z}\mathcal{B}_k, \tag{6.2}$$

$$\mathcal{B}^{*}(\mathcal{A})\, e^{-\bar{z}} = e^{-\bar{z}}\mathcal{B}_k^{*}, \tag{6.3}$$

where $z = u + iv$ and $\mathcal{B}_k = (\partial_x - ik_1) - i(\partial_y - ik_2)$. Define a transformation, Q, according to

$$Q = \begin{bmatrix} e^{-\bar{z}}I & 0 \\ 0 & e^{z}I \end{bmatrix}. \tag{6.4}$$

Setting $\psi = Q\xi$ in (6.1) yields

$$\begin{bmatrix} me^{-\bar{z}}I & e^{z}\mathcal{B}_k \\ -e^{-\bar{z}}\mathcal{B}_k^{*} & me^{z}I \end{bmatrix} \begin{bmatrix} \xi_R \\ \xi_L \end{bmatrix} = \begin{bmatrix} f_R \\ f_L \end{bmatrix}. \tag{6.5}$$

Denote this transformed operator in (6.5) by $\hat{\mathcal{D}}(\mathcal{A})$. Note that $\hat{\mathcal{D}}$ still depends on gauge field $\mathcal{A}$ because the gauge data is contained in the exponential terms. Then, if an efficient method of discretizing and solving the transformed system (6.5) exists, the original system, (6.1), can be solved by first solving (6.5) and then setting $\psi = Q\xi$. Thus, a solution process for the continuum problem, based on this transformation, is given in Algorithm 6.1.

Note that care must be taken with the boundary conditions prescribed to the auxiliary function ξ. Requiring ψ to be periodic on $\mathcal{R}$ is, of course, equivalent to requiring $Q\xi$ to be periodic on $\mathcal{R}$. But, since $z \in \mathcal{W}_{\mathbb{R}}$ is periodic by definition, $Q\xi$ will be periodic as long as ξ is. Thus, in Algorithm 6.1, the periodicity of ξ alone is enforced.

The auxiliary continuum equation in Step 1 of Algorithm 6.1 is discretized using least-squares finite elements by formulating the solution of (6.5) as a minimization problem:

ALGORITHM 6.1: Transformed Continuum Dirac Solve

Input: Gauge field $\mathcal{A}$, source term f.

Output: Wavefunction ψ.

1. Solve $\hat{\mathcal{D}}(\mathcal{A})\,\xi = f$.
2. Set $\psi = Q\xi$.

$$\xi \;=\; \arg\min_{\varphi\in\mathcal{V}_{\mathbb{C}}^2} \|\hat{\mathcal{D}}\varphi - f\|_0^2, \tag{6.6}$$

where $\mathcal{V}_{\mathbb{C}}$ is the usual space of continuous, periodic, complex-valued functions in $H^1(\mathcal{R})$. Minimization principle (6.6) is equivalent to the following weak form:

$$\text{Find } \xi \in \mathcal{V}_{\mathbb{C}}^2 \text{ s.t. } \left\langle \hat{\mathcal{D}}\xi, \hat{\mathcal{D}}w \right\rangle \;=\; \left\langle f, \hat{\mathcal{D}}w \right\rangle \quad \forall w \in \mathcal{V}_{\mathbb{C}}^2. \tag{6.7}$$

The formal normal for this weak form is given by

$$\hat{\mathcal{D}}^*\hat{\mathcal{D}} \;=\; \begin{bmatrix} me^{-z}I & -\mathcal{B}_k\, e^{-z} \\ \mathcal{B}_k^*\, e^{\bar{z}} & me^{\bar{z}}I \end{bmatrix} \begin{bmatrix} me^{-\bar{z}}I & e^{z}\mathcal{B}_k \\ -e^{-\bar{z}}\mathcal{B}_k^* & me^{z}I \end{bmatrix} \tag{6.8}$$

$$= \begin{bmatrix} m^2e^{-2u}I + \mathcal{B}_k^*\, e^{-2u}\mathcal{B}_k & 0 \\ 0 & m^2e^{2u}I + \mathcal{B}_k\, e^{2u}\mathcal{B}_k^* \end{bmatrix}. \tag{6.9}$$

Notice that v, which is associated with the gradient portion of the gauge field, vanishes from the formal normal. Moreover, $\hat{\mathcal{D}}^*\hat{\mathcal{D}}$ is block diagonal with each diagonal block containing a zeroth-order term and a second-order term resembling a diffusion operator with variable coefficients. This block diagonal structure is, in fact, what drove our choice of the particular transformation because it is easier to design multigrid solvers for scalar equations, especially those of diffusion type [23], [24].

The least-squares solution is obtained by restricting the minimization problem in (6.6) and, thus, the weak form in (6.7), to the usual finite-dimensional space, $\mathcal{V}_{\mathbb{C}}^h \subset \mathcal{V}_{\mathbb{C}}$. That is, the solution must satisfy the following weak form:

$$\text{Find } \xi^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2 \text{ s.t. } \left\langle \hat{\mathcal{D}}\xi^h, \hat{\mathcal{D}}w^h \right\rangle = \left\langle f^h, \hat{\mathcal{D}}w^h \right\rangle \quad \forall w^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2. \tag{6.10}$$

Step 2 of Algorithm 6.1 obtains the solution by setting $\psi = Q\xi$, which can be formulated as a weak form as well:

$$\text{Find } \psi^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2 \text{ s.t. } \left\langle \psi^h, w^h \right\rangle = \left\langle Q\xi^h, w^h \right\rangle \quad \forall w^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2. \tag{6.11}$$

Note that this is just the L^2-projection of $Q\xi^h$ onto $\left(\mathcal{V}_{\mathbb{C}}^h\right)^2$. These processes define the following potential algorithm for the solution of (6.1).

ALGORITHM 6.2: Transformed Least-Squares Dirac Solve

Input: Gauge field $\underline{A}$, source term $\underline{f}$.

Output: Wavefunction $\underline{\psi}$.

1. Compute $\underline{u}$ and $\underline{v}$ such that $\underline{A} = \mathbb{C}\underline{u} + \mathbb{G}\underline{v} + \underline{k}$.
2. Map $\underline{u} \mapsto u^h$ and $\underline{v} \mapsto v^h$.
3. Map $\underline{f} \mapsto f^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2$.
4. Find $\xi^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2$ s.t. $\left\langle \hat{\mathcal{D}}\xi^h, \hat{\mathcal{D}}w^h \right\rangle = \left\langle f^h, \hat{\mathcal{D}}w^h \right\rangle \quad \forall w^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2$,
5. Find $\psi^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2$ s.t. $\left\langle \psi^h, w^h \right\rangle = \left\langle Q\xi^h, w^h \right\rangle \quad \forall w^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2$.
6. Map $\psi^h \mapsto \underline{\psi} \in \left(\mathcal{N}_{\mathbb{C}}\right)^2$.

6.1.1 Gauge Covariance

Unfortunately, Algorithm 6.2, as it stands, does not satisfy the principle of gauge covariance. The same difficulty discovered in the development of the algorithm in Chapter 5 is present here. That is, gauge covariance is thwarted because the proper relationship between $\underline{f}$ and $\mathbb{T}^*_\omega \underline{f}$ is lost when these vectors are projected into the finite element space. The transformed algorithm suffers a similar problem in the weak form that relates the auxiliary solution, ξ^h, and the solution to the original problem, ψ^h, appearing in Step 5 of Algorithm 6.2. Fortunately, both problems are remedied using the same gauge fixing idea employed in Chapter 5.

As before, gauge covariance is retained by transforming the input data into an equivalent set corresponding to a divergence free gauge field. That is, given discrete source term, $\underline{f}$, and gauge data,

$$\underline{A} \;=\; \underline{A}_0 + \mathbb{G}\underline{v},$$

an auxiliary problem is solved based on

$$\widetilde{\underline{A}} \;=\; \underline{A}_0$$

and

$$\widetilde{\underline{f}} \;=\; \mathbb{T}^*_{\underline{v}}\, \underline{f}.$$

The solution is then obtained by setting

$$\underline{\psi} \;=\; \mathbb{T}_{\underline{v}}\, \widetilde{\underline{\psi}},$$

where $\widetilde{\underline{\psi}}$ is the solution based on the transformed data. Note the implication of this for the transformed system. Since the gauge field does not contain a gradient, transformation Q only involves the smoother curl portion of the gauge field. That is,

$$Q \;=\; Q(u) \;\;=\;\; \begin{bmatrix} e^{-u}I & 0 \\ 0 & e^{u}I \end{bmatrix}. \tag{6.12}$$

A gauge covariant algorithm based on the transformed system is formulated in Algorithm 6.3.

ALGORITHM 6.3: Transformed Gauge Covariant Least-Squares Dirac Solve

---

Input: Gauge field $\underline{A}$, source term $\underline{f}$.

Output: Wavefunction $\underline{\psi}$.

1. Compute $\underline{u}$ and $\underline{v}$ such that $\underline{A} = \mathbb{C}\underline{u} + \mathbb{G}\underline{v} + \underline{k}$.
2. Set $\underline{g} = \mathbb{T}^*_{\underline{v}}\, \underline{f}$
3. Map $\underline{u} \mapsto u^h$.
4. Map $\underline{g} \mapsto g^h \in \left(\mathcal{V}^h_{\mathbb{C}}\right)^2$.
5. Find $\xi^h \in \left(\mathcal{V}^h_{\mathbb{C}}\right)^2$ s.t. $\left\langle \hat{\mathcal{D}}\xi^h, \hat{\mathcal{D}}w^h \right\rangle = \left\langle f^h, \hat{\mathcal{D}}w^h \right\rangle \quad \forall w^h \in \left(\mathcal{V}^h_{\mathbb{C}}\right)^2$,
   where $\hat{\mathcal{D}} = \hat{\mathcal{D}}(\nabla^{\perp}u^h + \underline{k})$
6. Find $\zeta^h \in \left(\mathcal{V}^h_{\mathbb{C}}\right)^2$ s.t. $\left\langle \zeta^h, w^h \right\rangle = \left\langle Q\xi^h, w^h \right\rangle \quad \forall w^h \in \left(\mathcal{V}^h_{\mathbb{C}}\right)^2$,
   where $Q = Q(u^h)$
7. Map $\zeta^h \mapsto \underline{\zeta} \in (\mathcal{N}_{\mathbb{C}})^2$.
8. Set $\underline{\psi} = \mathbb{T}_{\underline{v}}\, \underline{\zeta}$

To avoid serious complication of notation, the same symbols used to represent discrete operators in Chapter 5 ($\mathbb{L}$, $\mathbb{K}$ and $\mathbb{D}_{LS}$) are used again here. However, their use in this chapter is entirely self contained, so there should be no confusion. Using the

nodal basis for $\mathcal{V}_{\mathbb{C}}^h$, the following matrix equation for Step 5 of Algorithm 6.3 can be established:

$$\mathbb{L}\underline{\xi} = \mathbb{K}\underline{g}, \tag{6.13}$$

where $\underline{\xi}$ and $\underline{g}$ are the coefficients in the expansions of ξ^h and g^h, respectively. Note that, since the discrete values of $\underline{\xi}$ and $\underline{g}$ naturally coincide with the expansion coefficients of ξ^h and g^h, respectively, it is convenient to represent them using the same notation. Matrices $\mathbb{L}$ and $\mathbb{K}$ are given according to

$$\mathbb{L} := \begin{bmatrix} \mathbb{L}_{11} & 0 \\ 0 & \mathbb{L}_{22} \end{bmatrix}, \tag{6.14}$$

$$\mathbb{L}_{11} := m^2\mathbb{M}^- + \mathbb{L}_{xx}^- + \mathbb{L}_{yy}^- + i(\mathbb{L}_{xy}^- - \mathbb{L}_{yx}^-), \tag{6.15}$$

$$\mathbb{L}_{22} := m^2\mathbb{M}^+ + \mathbb{L}_{xx}^+ + \mathbb{L}_{yy}^+ + i(\mathbb{L}_{xy}^+ - \mathbb{L}_{yx}^+), \tag{6.16}$$

$$\mathbb{K} = \begin{bmatrix} m\mathbb{M}^+ & \mathbb{B}_x^+ - i\mathbb{B}_y^+ \\ \mathbb{B}_x^- + i\mathbb{B}_y^- & m\mathbb{M}^- \end{bmatrix}, \tag{6.17}$$

where

$$
\begin{aligned}
[\,\mathbb{L}^{\pm}_{xx}\,]_{j,k} &= < e^{\pm u}(\partial_x - ik_1)\phi_k, e^{\pm u}(\partial_x - ik_1)\phi_j >, \\
[\,\mathbb{L}^{\pm}_{xy}\,]_{j,k} &= < e^{\pm u}(\partial_x - ik_1)\phi_k, e^{\pm u}(\partial_y - ik_2)\phi_j >, \\
[\,\mathbb{L}^{\pm}_{yx}\,]_{j,k} &= < e^{\pm u}(\partial_y - ik_2)\phi_k, e^{\pm u}(\partial_x - ik_1)\phi_j >, \\
[\,\mathbb{L}^{\pm}_{yy}\,]_{j,k} &= < e^{\pm u}(\partial_y - ik_2)\phi_k, e^{\pm u}(\partial_y - ik_2)\phi_j >, \\
[\,\mathbb{B}^{\pm}_{x}\,]_{j,k} &= < \phi_k, e^{\pm u}(\partial_y - ik_2)\phi_j >, \\
[\,\mathbb{B}^{\pm}_{y}\,]_{j,k} &= < \phi_k, e^{\pm u}(\partial_x - ik_1)\phi_j >, \\
[\,\mathbb{M}^{\pm}\,]_{j,k} &= < e^{\pm u}\phi_k, e^{\pm u}\phi_j > .
\end{aligned}
$$

Here, ϕ_j is the usual bilinear finite element basis function associated with the j^{th} lattice site. Similarly, a linear system can be developed to replace the weak form in Step 6 of Algorithm 6.3:

$$\mathbb{P}\underline{\zeta} = \mathbb{Q}\underline{\xi}. \tag{6.18}$$

where

$$\mathbb{P} = \begin{bmatrix} \mathbb{P}_0 & 0 \\ 0 & \mathbb{P}_0 \end{bmatrix}, \qquad \mathbb{Q} = \begin{bmatrix} \mathbb{Q}^- & 0 \\ 0 & \mathbb{Q}^+ \end{bmatrix}, \tag{6.19}$$

and

$$[\,\mathbb{P}_0\,]_{j,k} = < \phi_k, \phi_j >, \quad [\,\mathbb{Q}^{\pm}\,]_{j,k} = < e^{\pm u}\phi_k, \phi_j > . \tag{6.20}$$

Using arguments similar to those employed in Chapter 5, it is easy to see that $\mathbb{L}$ and $\mathbb{K}$ are singular only when $m = 0$ and A_0^h is an exceptional configuration. Furthermore, matrices $\mathbb{P}$ and $\mathbb{Q}$ are nonsingular for all masses and gauge configurations.

Using the above matrix representations, Algorithm 6.3 can be formulated completely in the discrete setting. This is summarized in Algorithm 6.4.

ALGORITHM 6.4: Discrete Transformed Gauge Covariant Least-Squares Dirac Solve

---

Input: Gauge field $\underline{A}$, source term $\underline{f}$.

Output: Wavefunction $\underline{\psi}$.

1. Compute $\underline{u}$ and $\underline{v}$ such that $\underline{A} = \mathbb{C}\underline{u} + \mathbb{G}\underline{v} + \underline{k}$.
2. Set $\underline{g} = \mathbb{T}^*_{\underline{v}}\, \underline{f}$
3. Compute $\underline{\xi} = \mathbb{L}^{-1}\mathbb{K}\underline{g}$
4. Compute $\underline{\zeta} = \mathbb{P}^{-1}\mathbb{Q}\underline{\xi}$
5. Set $\underline{\psi} = \mathbb{T}_{\underline{v}}\, \underline{\zeta}$

Combining Steps 2-5 in Algorithm 6.4 yields an expression for the discrete least-squares propagator, hereafter denoted by $\mathbb{D}^{-1}_{LS}$:

$$\mathbb{D}^{-1}_{LS} \;=\; \mathbb{T}_{\underline{v}}\, \mathbb{P}^{-1}\mathbb{Q}\, \mathbb{L}^{-1}\mathbb{K}\, \mathbb{T}^*_{\underline{v}}. \tag{6.21}$$

From (6.21), the least-squares representation of the Dirac operator is given by

$$\mathbb{D}_{LS} \;=\; \mathbb{T}_{\underline{v}}\, \mathbb{K}^{-1}\mathbb{L}\mathbb{Q}^{-1}\mathbb{P}\, \mathbb{T}^*_{\underline{v}}. \tag{6.22}$$

The operators given in (6.21) and (6.22) are ill-posed only if $m = 0$ and A^h_0 is an exceptional configuration.

The discrete gauge covariance of Algorithm 6.4 is established in the following theorem.

Theorem 6.1.1. *The least-squares solution process defined in Algorithm 6.4 satisfies discrete gauge covariance as described in Definition 3.1.1.*

Proof. We need to show that, given the least-squares propagator, $\mathbb{D}_{LS}^{-1}$, and associated discrete gauge transformation $\underline{\Omega}_{\underline{\omega}}$ and $\mathbb{T}_{\underline{\omega}}$, there exists a modified solution process, denoted by $\widetilde{\mathbb{D}}_{LS}^{-1}$, such that

$$\mathbb{D}_{LS}^{-1}\,\mathbb{T}_{\underline{\omega}}\,\underline{\xi} \;=\; \mathbb{T}_{\underline{\omega}}\,\widetilde{\mathbb{D}}_{LS}^{-1}\,\underline{\xi}.$$

Then, from (6.21),

$$\begin{aligned}\mathbb{D}_{LS}^{-1}\,\mathbb{T}_{\underline{\omega}}\,\underline{\xi} &= \mathbb{T}_{\underline{v}}\,\mathbb{P}^{-1}\mathbb{Q}\mathbb{L}^{-1}\mathbb{K}\mathbb{T}_{\underline{v}}^{*}\,\mathbb{T}_{\underline{\omega}}\,\underline{\xi}\\ &= \mathbb{T}_{\underline{\omega}}\,\left[\mathbb{T}_{\underline{v}-\underline{\omega}}\,\mathbb{P}^{-1}\mathbb{Q}\mathbb{L}^{-1}\mathbb{K}\mathbb{T}_{\underline{v}-\underline{\omega}}^{*}\,\right]\underline{\xi}.\end{aligned}$$

Finally, defining $\widetilde{\mathbb{D}}_{LS}^{-1}$ according to

$$\widetilde{\mathbb{D}}_{LS}^{-1} \;=\; \mathbb{T}_{\underline{v}-\underline{\omega}}\,\mathbb{P}^{-1}\mathbb{Q}\mathbb{L}^{-1}\mathbb{K}\mathbb{T}_{\underline{v}-\underline{\omega}}^{*}$$

proves the result.

□

6.1.2 Chiral Symmetry

The proof that the solution process in Algorithm 6.4 satisfies chiral symmetry is similar to that used to show that Algorithm 5.3 does. Again, recall Definition 3.1.2 regarding the discrete chiral symmetry of a discrete Dirac operator. Given $\omega_{_R}$ and $\omega_{_L} \in \mathbb{R}$, and associated transformation $\underline{\Lambda}$, the least-squares analogue of the *massless* transformed Dirac operator, $\mathbb{D}_{LS}$, must satisfy

$$\left\langle \underline{\Lambda}\underline{\xi}, \underline{\Gamma}_1\mathbb{D}_{LS}\underline{\Lambda}\underline{\xi}\right\rangle \;=\; \left\langle \underline{\xi}, \underline{\Gamma}_1\mathbb{D}_{LS}\underline{\xi}\right\rangle. \tag{6.23}$$

Recall, from (6.22), that, in the massless case, $\mathbb{D}_{LS}^{-1}$ does not exist if the gauge field is an exceptional configuration. Instead, the discrete matrix equation can be represented as

$$\mathbb{L}\mathbb{Q}^{-1}\mathbb{P}\,\mathbb{T}_{\underline{v}}^{*}\,\underline{\psi} \;=\; \mathbb{K}\mathbb{T}_{\underline{v}}^{*}\,\underline{f}.$$

This representation is valid since $\mathbb{Q}$ is nonsingular independently of m and A_0^h. The following lemma demonstrates that the transformed least-squares process satisfies chiral symmetry.

Lemma 6.1.2. *(Chiral symmetry for the discrete least-squares operator). Given any* $\lambda_{_R},\ \lambda_{_L} \in \mathbb{R}$, *and any* $\underline{\psi},\ \underline{f} \in \mathcal{N}_{\mathbb{C}}^2$ *such that*

$$\mathbb{L}\mathbb{Q}^{-1}\mathbb{P}\,\mathbb{T}_{\underline{v}}^{*}\,\underline{\psi} \;=\; \mathbb{K}\mathbb{T}_{\underline{v}}^{*}\,\underline{f} \tag{6.24}$$

for $m = 0$, *then*

$$\begin{aligned} \underline{\hat{\psi}} &= \underline{\Lambda}\underline{\psi}, \\ \underline{\hat{f}} &= \underline{\Gamma}_1\underline{\Lambda}\underline{\Gamma}_1\underline{f}, \end{aligned} \tag{6.25}$$

satisfy

$$\mathbb{L}\mathbb{Q}^{-1}\mathbb{P}\,\mathbb{T}_{\underline{v}}^{*}\,\underline{\hat{\psi}} \;=\; \mathbb{K}\mathbb{T}_{\underline{v}}^{*}\,\underline{\hat{f}},$$

where $\underline{\Lambda}$ *is as defined in (3.14).*

Proof. In the massless case, matrices $\mathbb{L}$ and $\mathbb{K}$ have the following block form:

$$\mathbb{L} = \begin{bmatrix} \mathbb{L}_{11} & 0 \\ 0 & \mathbb{L}_{22} \end{bmatrix},$$
$$\mathbb{K} = \begin{bmatrix} 0 & \mathbb{K}_{12} \\ \mathbb{K}_{21} & 0 \end{bmatrix}.$$

Noting that the individual diagonal blocks of $\mathbb{Q}$ are themselves nonsingular, we write

$$\mathbb{Q}^{-1} = \begin{bmatrix} (\mathbb{Q}^-)^{-1} & 0 \\ 0 & (\mathbb{Q}^+)^{-1} \end{bmatrix}.$$

Then, recalling the block form of $\mathbb{P}$, matrices $\mathbb{L}\mathbb{Q}^{-1}\mathbb{P}\,\mathbb{T}^*_{\underline{v}}$ and $\mathbb{K}\mathbb{T}^*_{\underline{v}}$ appear as

$$\mathbb{L}\mathbb{Q}^{-1}\mathbb{P}\,\mathbb{T}^*_{\underline{v}} = \begin{bmatrix} \mathbb{L}_{11}\,(\mathbb{Q}^-)^{-1}\,\mathbb{P}_0\,\underline{\Omega}^*_{\underline{v}} & 0 \\ 0 & \mathbb{L}_{22}\,(\mathbb{Q}^+)^{-1}\,\mathbb{P}_0\,\underline{\Omega}^*_{\underline{v}} \end{bmatrix}, \tag{6.26}$$
$$\mathbb{K}\mathbb{T}^*_{\underline{v}} = \begin{bmatrix} 0 & \mathbb{K}_{12}\,\underline{\Omega}^*_{\underline{v}} \\ \mathbb{K}_{21}\,\underline{\Omega}^*_{\underline{v}} & 0 \end{bmatrix}. \tag{6.27}$$

From the block structure of (6.26) and the constant nature of the diagonal blocks of $\underline{\Lambda}$, it is clear that

$$\mathbb{L}\mathbb{Q}^{-1}\,\mathbb{P}\mathbb{T}^*_{\underline{v}}\,\underline{\Lambda} = \underline{\Lambda}\mathbb{L}\mathbb{Q}^{-1}\,\mathbb{P}\mathbb{T}^*_{\underline{v}}. \tag{6.28}$$

Then, from the block structure of (6.27), and the constant nature of $\underline{\Lambda}$, it follows that

$$\mathbb{K}\mathbb{T}^*_{\underline{v}}\,\underline{\Gamma}_1\,\underline{\Lambda}\,\underline{\Gamma}_1 = \underline{\Lambda}\,\mathbb{K}\mathbb{T}^*_{\underline{v}}. \tag{6.29}$$

Thus,

$$\begin{aligned}\mathbb{L}\mathbb{Q}^{-1}\,\mathbb{P}\mathbb{T}^*_{\underline{v}}\hat{\psi} &= \mathbb{L}\mathbb{Q}^{-1}\mathbb{Q}\mathbb{T}^*_{\underline{v}}\underline{\Lambda}\psi \\ &= \underline{\Lambda}\mathbb{L}\mathbb{Q}^{-1}\,\mathbb{P}\mathbb{T}^*_{\underline{v}}\psi,\end{aligned}$$

and

$$\begin{aligned}\mathbb{K}\mathbb{T}^*_{\underline{v}}\,\hat{f} &= \mathbb{K}\mathbb{T}^*_{\underline{v}}\,\underline{\Gamma}_1\,\underline{\Lambda}\,\underline{\Gamma}_1\underline{f} \\ &= \underline{\Lambda}\,\mathbb{K}\mathbb{T}^*_{\underline{v}}\,\underline{f},\end{aligned}$$

which completes the proof.

□

6.1.3 Species Doubling

One way to see that the solution process given by Algorithm 6.4 does not suffer from species doubling is to recall that the doubling analysis done for Algorithm 5.3 in Chapter 5 is based on the spectrum of the gauge-free propagator. Then, note that in the gauge free case, Algorithm 6.4 coincides exactly with Algorithm 5.3. Since Algorithm 5.3 does not suffer from species doubling, neither does Algorithm 6.4. Note also, that from (6.14), it is easy to see that the principle part of $\mathbb{L}$ is again based on a 9-point Laplacian-like stencil. This operator clearly does not suffer from red-black instability, and thus does not suffer from species doubling.

6.2 H^1-Ellipticity

The least-squares functional for the transformed system can be shown to satisfy H^1-ellipticity, but in a scaled version of the H^1-norm. From the definition of $\hat{\mathcal{D}}$ given in (6.5) and minimization principle (6.6), the least-squares functional is defined as

$$\hat{G}\left(\psi, \mathcal{A}; f\right) \;=\; \|me^{-\bar{z}}\xi_{R} + e^{z}\mathcal{B}_{k}\,\xi_{L} - f_{R}\|_{0}^{2} + \|me^{z}\xi_{L} - e^{-\bar{z}}\mathcal{B}_{k}^{*}\,\xi_{R} - f_{L}\|_{0}^{2}. \quad (6.30)$$

The proof of the ellipticity of $\hat{G}\left(\psi, \mathcal{A}; 0\right)$ follows almost immediately from that of the Theorem 5.2.7.

6.2.1 Main Theorem

Theorem 6.2.1. *For any $\xi \in \hat{\mathcal{U}}$, a closed subspace of $\mathcal{V}_{\mathbb{C}}^{2}$, there exist positive constants $c_{\mathcal{A}}$ and $C_{\mathcal{A}}$, which depend on the gauge field, $\mathcal{A}$, such that*

$$c_{\mathcal{A}}|||\xi|||_{1}^{2} \;\leq\; \hat{G}(\xi, \mathcal{A}; 0) \;\leq\; C_{\mathcal{A}}|||\xi|||_{1}^{2},$$

where $|||\cdot|||_{1}$ is a scaled version of the H^{1}-norm defined by

$$|||\xi|||_{1} \;=\; \|Q\xi\|_{1}. \quad (6.31)$$

If $m > 0$ or $\mathcal{A}$ is not an exceptional configuration then coercivity holds on $\hat{\mathcal{U}} = \mathcal{V}_{\mathbb{C}}^{2}$. If $m = 0$ and $\mathcal{A}$ is an exceptional configuration then $\hat{\mathcal{D}}$ is singular and and coercivity holds on $\hat{\mathcal{U}} = \mathcal{N}(\hat{\mathcal{D}})^{\perp}$.

Proof. Recall, from Theorem 5.2.7, that for any $\psi \in \mathcal{U}$, a closed subspace of $\mathcal{V}_{\mathbb{C}}^{2}$, there exist positive constants $c_{\mathcal{A}}$ and $C_{\mathcal{A}}$ such that

$$c_{\mathcal{A}}\|\psi\|_{1}^{2} \;\leq\; \|\mathcal{D}\psi\|_{0}^{2} \;\leq\; C_{\mathcal{A}}\|\psi\|_{1}^{2}. \quad (6.32)$$

Writing (6.32) as

$$c_{\mathcal{A}}\|\psi\|_{1}^{2} \;\leq\; \|\mathcal{D}\left(QQ^{-1}\right)\psi\|_{0}^{2} \;\leq\; C_{\mathcal{A}}\|\psi\|_{1}^{2}. \quad (6.33)$$

Then, recalling that $\hat{\mathcal{D}} = \mathcal{D}Q$, and setting $\xi = Q^{-1}\psi$, then (6.32) becomes

$$c_{\mathcal{A}}\|Q\xi\|_1^2 \;\leq\; \|\hat{\mathcal{D}}\xi\|_0^2 \;\leq\; C_{\mathcal{A}}\|Q\xi\|_1^2, \tag{6.34}$$

as desired. Note that, $c_{\mathcal{A}}$ and $C_{\mathcal{A}}$, here are exactly the same constants defined in the proof of Theorem 5.2.7. Furthermore, if $\mathcal{U} = \mathcal{U}_{R} \otimes \mathcal{U}_{L}$, then $\hat{\mathcal{U}} = e^{\bar{z}}\mathcal{U}_{R} \otimes e^{-z}\mathcal{U}_{L}$.

□

6.3 Numerical Experiments

In this section, we consider the numerical solution of the equations that appear in Step 3 of Algorithm 6.4:

$$\mathbb{L}\underline{\xi} \;=\; \mathbb{K}\underline{g}. \tag{6.35}$$

Note that the problem in Step 4 of Algorithm 6.4, $\mathbb{P}\zeta = \mathbb{Q}\xi$, is not considered here because $\mathbb{P}$ is a mass-like matrix. Thus, $\mathbb{P}$ can be inverted easily using a number of standard methods. In the experiments that follow, standard algebraic multigrid (AMG) and adaptive smoothed aggregation multigrid (αSA) are each used as preconditioners for CG.

6.3.1 Numerical Results

Like the system of interest in Chapter 5, problem (6.35) contains complex entries and it can be rewritten in equivalent real formulation (ERF):

$$\begin{bmatrix} \mathbb{X} & -\mathbb{Y} \\ \mathbb{Y} & \mathbb{X} \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix} \;=\; \begin{bmatrix} \underline{a} \\ \underline{b} \end{bmatrix}, \tag{6.36}$$

where $\mathbb{X}, \mathbb{Y}$ are real-valued matrices satisfying $\mathbb{L} = \mathbb{X} + i\mathbb{Y}$, $\underline{\xi} = \underline{x} + i\underline{y}$, and $\mathbb{K}\underline{g} = \underline{a} + i\underline{b}$. Again, we are interested in how the performance of AMG varies with fermion mass, m, gauge field temperature parameter, β, and lattice size, N.

In the following tests, AMG-PCG and αSA-PCG are applied to (6.36). For AMG-PCG, a single V(1,1)-cycle with Gauss-Seidel relaxation is used as the preconditioner in each step of CG. For αSA-PCG, a single V(2,2)-cycle is used as the preconditioner and is constructed using 4 prototype error components to build the interpolation operator. The relaxation method is Gauss-Seidel. Aggregation is done algebraically. Each method is applied to (6.36) with a zero right-hand side and random initial guess. The iteration is terminated when the relative residual in the iteration has been decreased by a factor of 10^{-6}. Average convergence factors are computed by applying the method to operators constructed using 20 distinct gauge fields.

Table 6.1 reports average convergence factors for AMG-PCG and αSA-PCG, applied to (6.36). Various values of the particle mass m, and gauge field temperature β are considered on varying lattice sizes.

β/m	.001	.01	.1
2	.40	.39	.33
5	.29	.29	.27
3	.24	.24	.23

β/m	.001	.01	.1
2	.22	.20	.12
3	.18	.18	.10
5	.15	.15	.14

β/m	.001	.01	.1
2	.35	.34	.27
3	.29	.29	.27
5	.27	.27	.26

β/m	.001	.01	.1
2	.22	.22	.13
3	.19	.20	.18
5	.16	.15	.11

β/m	.001	.01	.1
2	.43	.42	.34
3	.40	.41	.40
5	.38	.38	.34

β/m	.001	.01	.1
2	.22	.21	.12
3	.22	.22	.17
5	.18	.17	.12

Table 6.1: Average convergence factors for AMG-PCG (left) and αSA-PCG (right) applied to (6.36) on 64×64 (top), 128×128 (middle), and 256×256 (bottom) lattices with varying choices of mass parameter, m, and temperature, β. In each case, operator complexity, σ, with AMG-PCG was approximately 1.8 and with αSA-PCG was approximately 1.4.

In Table 6.1, all tests with AMG-PCG have operator complexities of approxi-

mately 1.80, and with αSA-PCG have operator complexities of approximately 1.40. For both solvers, the method performs better for the largest mass tested and slightly worse for the two smaller masses. However, since performances remains fairly static between $m = .01$ and $m = .001$ it appears that critical slowing down has been eliminated. The value of β seems to affect the performance on both solvers. As the value of β increases the matrix becomes easier to invert with both methods. This is not surprising, since a larger β implies less disorder in the background gauge field. Finally, both solvers appear to be scalable with respect to the lattice size.

To determine which solver is the most efficient we must consider convergence rates, operator complexities, and types of V-cycle. The convergence rates reported in Table 6.1 are better for αSA-PCG than for AMG-PCG. Furthermore, the αSA V-cycle has an operator complexity smaller than that of the AMG V-cycle, meaning that relaxation sweeps and residual calculations are cheaper for αSA. However, αSA performs two more relaxation steps on each level than AMG. To directly compare the two methods we again compute their respective computational cost per digit of error reduction. These η-values are given in Table 6.2.

β/m	.001	.01	.1
2	13.6	13.2	11.2
3	10.0	10.0	9.5
5	8.7	8.7	8.5

β/m	.001	.01	.1
2	10.6	10.0	7.6
3	9.4	9.4	7.0
5	8.5	8.5	8.2

Table 6.2: Average η-values for AMG-PCG (left) and αSA-PCG (right) applied to (6.36) on a 64×64 lattice with varying choices of mass parameter m and temperature β.

Table 6.2 indicates that αSA-PCG performs better on (6.36) than AMG-PCG, but only slightly. In fact, for simulations in which the discrete system need only be solved with a few right-hand side vectors per gauge field, the larger setup costs of αSA may render the method less efficient than AMG-PCG.

Again, it is interesting to compare the performance of an algebraic multigrid

method applied to both the least-squared discretization and the Dirac-Wilson operator. Since αSA-PCG is slightly more effective at solving (6.36) than AMG-PCG, we compare it to αSA-PCG applied to the equivalent real formulation of the normal equations of $\mathbb{D}_W$. For convenience, this operator is reiterated below. It is

$$\begin{bmatrix} \mathbb{U} & -\mathbb{V} \\ \mathbb{V} & \mathbb{U} \end{bmatrix} \begin{bmatrix} \underline{u} \\ \underline{v} \end{bmatrix} = \begin{bmatrix} \underline{c} \\ \underline{d} \end{bmatrix}, \tag{6.37}$$

where $\mathbb{U}$ and $\mathbb{V}$ are real-valued matrices satisfying $\mathbb{D}_W^* \mathbb{D}_W = \mathbb{U} + i\mathbb{V}$, $\underline{\psi} = \underline{u} + i\underline{v}$, and $\mathbb{D}_W^* \underline{f} = \underline{c} + i\underline{d}$. Table 6.3 compares the performance of αSA-PCG applied to (6.36) and αSA-PCG applied to (6.37). For the least-squares operator, (6.36), an αSA V(2,2)-cycle is used as a preconditioner for CG, where 4 prototype error components are constructed in the setup phase and used to define the method. For (6.37), an αSA V(2,2)-cycle is used, with 8 prototype error components used to construct the method. Average convergence factors for these methods are compared in Table 6.3.

β/m	.01	.1	.3
2	.20	.12	.11
3	.18	.10	.10
5	.15	.14	.12

β/m	.01	.1	.3
2	.33	.31	.31
3	.42	.40	.31
5	.31	.29	.28

Table 6.3: Average convergence factors for AMG-PCG applied to the least-squares formulation (left) and αSA-PCG applied to the normal equations of the Dirac-Wilson operator (right) on a 64×64 lattice with varying choices of mass parameter m and temperature β. In the least-squares formulation, the operator complexity, σ, is approximately 1.4. In the Dirac-Wilson case, σ is approximately 3.0.

From Table 6.3 it is clear that αSA-PCG applied to (6.36) converges significantly faster than αSA-PCG applied to (6.37). Here, the computational cost of αSA-PCG applied to (6.37) is much greater than when applied to (6.36) for multiple reasons. First, the fine-grid normal equations of the Dirac-Wilson operator again have approximately 88% more nonzeros than the fine-grid operator from the least-squares formulation. Thus,

all fine-grid computations are more expensive in the Dirac-Wilson case. Second, the operator complexity associated with the Dirac-Wilson operator is 3.0 while the operator complexity associated with the least-squares operator is only 1.4. The discrepancy in operator complexity occurs for two reasons. First, the use of 8 prototype error components for the Dirac-Wilson operator (versus 4 for least-squares) naturally yields coarse grids with a greater number of unknowns. Second, because the normal equations of the Dirac-Wilson operator must be solved, the stencil of the fine-grid operator is much larger than the stencil of the least-squares operator. This leads to coarse-grid operators with more connections between unknowns and, thus, more nonzeros in the coarse-grid operators. To make a direct comparison between the two cases the appropriate η-values, defined in (5.66) and (5.67), are compared. The results are reported in Table 6.4. The associated speedup is reported in Table 6.5.

β/m	.01	.1	.3
2	10.0	7.6	7.3
3	9.4	7.0	7.0
5	8.5	8.2	7.6

β/m	.01	.1	.3
2	58.6	55.4	55.4
3	72.8	70.9	55.4
5	55.4	52.5	51.0

Table 6.4: Average η_{LS} and η_W-values for αSA-PCG applied to the least-squares discretization and αSA-PCG applied to the normal equations of the Dirac-Wilson discretization on a 64×64 lattice with varying choices of mass parameter, m, and temperature, β.

β/m	.01	.1	.3
2	5.8	7.3	7.6
3	8.0	10.1	7.9
5	6.5	6.4	6.7

Table 6.5: Average speedup factors for αSA-PCG applied to the least-squares discretization over αSA-PCG applied to the normal equations of the Dirac-Wilson discretization on a 64×64 lattice with varying choices of mass parameter, m, and temperature, β.

The speedups reported in Table 6.5 show that αSA-PCG applied to the least-squares discretization attains between 5 and 10 times the reduction in error per work unit

than αSA-PCG applied to the normal equations of the Dirac-Wilson operator.

Chapter 7

Conclusions and Future Work

We conclude by reiterating the accomplishments of this thesis. We discuss the long term goals using least-squares finite element methods to discretize the governing equations of QED and QCD.

Chapters 1-3 serve as a general introduction to the Dirac equation for the applied mathematician, both in the continuum and on the lattice. The full physical models of QED and QCD are introduced, as well as the simpler 2D Schwinger model. Finally, an alternate formulation of the 2D Schwinger model is introduced. Physical properties of interest are defined and discussed for the continuum model, including gauge covariance of the fermion propagator and chiral symmetry. These concepts are then extended to the lattice setting. Two traditional discretizations of the governing equations are introduced based on covariant finite-differences. These are the Naive and Dirac-Wilson discretizations. Through their development, the concept of species doubling is discovered and discussed. We argue that both discretizations have serious problems, the Naive's being species doubling, and Wilson's loss of chiral symmetry and very poor approximation of the low modes of the continuum operator. A brief introduction to the least-squares finite element methodology is presented in Chapter 4.

Chapter 5 develops a discretization of the 2D Schwinger model by applying the least-squares methodology to the governing equations. It is quickly found that the resulting discretization does not satisfy gauge covariance. To remedy this, the method

is altered, based on gauge fixing, to satisfy gauge covariance.

There are several advantages to discretizing the Schwinger model by least-squares finite elements. First, the original governing equations are non-Hermitian. Thus, straightforward discretization methods produce similarly non-Hermitian discrete operators that are complicated to invert using multilevel methods. Using the least-squares methodology avoids this problem because, by construction, the resulting discrete operator that must be inverted in the solution process is Hermitian positive semidefinite. This allows the application standard algebraic multigrid techniques without serious modification. Second, the principle part of the resulting operator has a 9-point Laplacian-like stencil. Such an operator cannot have red-black instability and, thus, cannot suffer from species doubling. This is perhaps the most fundamental benefit of the least-squares methodology. By naturally avoiding species doubling, there is no need to add artificial diffusion terms to the operator that break chiral symmetry and seriously damage the spectrum of the discrete operator in the process. To see this, we compare the spectra of the continuum, least-squares, and Dirac-Wilson operators. We show that the least-squares operator, unlike the Dirac-Wilson operator, approximates the relevant parts of the continuum spectrum very well.

Additionally, the least-squares functional is proved to be H^1-elliptic. This allows an additional argument that the discrete least-squares operator does not suffer from species doubling. It also implies that an optimal multilevel method exists that can efficiently solve the resulting system of equations. We investigate the use of algebraic multigrid preconditioned conjugate gradients (AMG-PCG) and adaptive smoothed aggregation multigrid preconditioned conjugate gradients (αSA-PCG) as solution processes for the resulting discrete system. Numerical experiments show that the discrete least-squares operator can be solved efficiently using these methods. Furthermore, neither AMG-PCG nor αSA-PCG suffer from critical slowing down, and both scale well with lattice size.

In Chapter 6, the least-squares methodology is applied to a transformation of the 2D Schwinger operator. The transformation, based on a Helmholtz decomposition of the gauge field, effectively removes the gauge field from the differential operator. Then, applying least-squares yields a discrete operator similar to those obtained by discretizing a variable-coefficient diffusion operator. This is very promising because algebraic multigrid methods have proved to be highly effective at solving these types of systems.

Again, gauge fixing is necessary to produce a gauge covariant solution process. As in Chapter 5, we demonstrate that the discretization based on the transformed system satisfies chiral symmetry while not suffering from species doubling. We also show that the functional based on the transformed continuum operator satisfies H^1-ellipticity in a scaled H^1-norm. We solve the resulting linear system using AMG-PCG and αSA-PCG. Both methods perform very well, attaining convergence rates comparable to those associated with Laplacian operators with variable coefficients. Finally, we compare the solution of the discrete least-squares operator by AMG-PCG with the solution of the Dirac-Wilson operator by αSA-PCG. We show that the solution of the former is roughly seven times faster than the solution of the latter.

In summary, the main accomplishment of this thesis is the development of two discretizations of the Schwinger model based on least-squares finite elements. Both discretizations yield solution processes that retain gauge covariance and chiral symmetry, while avoiding the problem of species doubling. Furthermore, both discretizations agree very well spectrally with the continuum operator. Finally, both methods produce Hermitian positive semidefinite linear systems of equations that can be solved very effectively by algebraic multigrid methods. This is the first result to date that accomplishes these goals without extending the theory to a costly extra dimension.

Although this thesis demonstrates that the least-squares methodology yields a discretization of the Schwinger model with many nice properties, it is still based on

a simplified model. Extending it to the full physical model of either QED or QCD is much more complicated. Though the form of the operators for the full QED model are the same, extending a finite-element discretization to four dimensions is extremely complicated. In QCD, gauge transformation are represented by objects in SU(3), which, unlike objects in U(1), do not commute with one another. The lack of cancellation in the formal normal means that a least-squares discretization of QCD yields a discrete operator with many more nonzero entries. This, in and of itself, is not a serious problem because the least-squares operator will not be any *more* dense than that of traditional discretizations, but it may hinder the performance of multigrid solvers.

Finally, since there is no analytic solution of the governing equations for an arbitrary gauge field, the only method of testing the viability of a discretization is to incorporate it in a Monte Carlo simulation of the theory. Since this requires much more knowledge of the physical theory, it is beyond the scope of this thesis. We are confident, though, based on our faithfulness to the original partial differential equation and the agreement between the spectra of the discrete and continuum operators, that our discretization methodology would be successful.

Bibliography

[1] J. Adler. Nested Iteration and FOSLS for Incompressible Resistive Magnetohydro-. dynamics. Ph.D. thesis, University of Colorado at Boulder, 2009.

[2] K. Atkinson. An Introduction to Numerical Analysis. John Wiley and Sons, 2nd edition, 1989.

[3] T. Austin and T. Manteuffel. A least-squares finite element method for the linear Boltzmann equation with anisotropic scattering. SIAM J. Numer. Anal., 44:540–560, 2006.

[4] M. Berndt, T. Manteuffel, and S. McCormick. Analysis of first-order system least squares (FOSLS) for elliptic problems with discontinuous coefficients: Part I. SIAM J. Numer. Anal., 43:386, 2005.

[5] M. Berndt, T. Manteuffel, and S. McCormick. Analysis of first-order system least squares (FOSLS) for elliptic problems with discontinuous coefficients: Part II. SIAM J. Numer. Anal., 43:409, 2005.

[6] Z. Bochev, Z. Cai, T. Manteuffel, and S. McCormick. First-order system least squares principles for the Navier-Stokes Equations: Part I. SIAM J. Numer. Anal., 35:990–1009, 1998.

[7] Z. Bochev, Z. Cai, T. Manteuffel, and S. McCormick. First-order system least squares principles for the Navier-Stokes Equations: Part II. SIAM J. Numer. Anal., 36:1125–1144, 1999.

[8] D. Braess. Towards algebraic multigrid for elliptic problems of second order. Computing, 55:379–393, 1995.

[9] D. Braess. Finite Elements. Cambridge Univerity Press, Cambridge MA, 2nd edition, 2001.

[10] A. Brandt. personal communication, 2008.

[11] A. Brandt, S. McCormick, and J. Ruge. Algebraic multigrid (amg) for solution with application to geodetic computations. Sparsity and Its Applications, 1984.

[12] A. Brandt, S. McCormick, and J. Ruge. Algebraic multigrid (AMG) for sparse matrix equations. Sparsity and Its Applications, 1984.

[13] J. Brannick, M. Brezina, D. Keyes, O. Livne, I. Livshits, S. MacLachlan, T. Manteuffel, S. McCormick, J. Ruge, and L. Zikatanov. Adaptive Smoothed Aggregation in Lattice QCD. Lecture Notes in Computational Science and Engineering, 55:507, 2007.

[14] J. Brannick, J. C. Brower, M. A. Clark, T. Manteuffel, S. McCormick, J. C. Osborn, and C. Rebbi. The removal of critical slowing down. Submitted in Proceedings of Science (PoS), Lat2008, 2008.

[15] J. Brannick, RC Brower, MA Clark, JC Osborn, and C. Rebbi. Adaptive Multigrid Algorithm for Lattice QCD. Physical Review Letters, 100(4):41601, 2008.

[16] J. Brannick, C. Ketelsen, T. Manteuffel, and S. McCormick. Least-squares finite element methods for quantum electrodynamics. SIAM J. Sci. Comp., 2009. Submitted.

[17] S.C. Brenner and L.R. Scott. Mathematical Theory of Finite Element Methods. Springer, 2nd edition, 2002.

[18] M. Brezina, R. Falgout, S. MacLachlan, T. Manteuffel, S. McCormick, and J. Ruge. Adaptive smoothed aggregation (alphasa) multigrid. SIAM Review, 47(2):317, 2005.

[19] M. Brezina, R. Falgout, S. MacLachlan, T. Manteuffel, S. McCormick, and J. Ruge. Adaptive algebraic multigrid. SIAM J. on Sci. Comp. (SISC), 27:1261–1286, 2006.

[20] M. Brezina, T. Manteuffel, S. McCormick, J. Ruge, and G. Sanders. Towards adaptive smoothed aggregation (αsa) for nonsymmetric problems. SIAM J. Sci. Comp., 2008. Submitted.

[21] W. L. Briggs, V. E. Henson, and S. F. McCormick. A Multigrid Tutorial. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[22] R. C. Brower, C. Edwards, C. Rebbi, and E. Vicari. Projective multigrid for wilson fermions. Nucl. Phys., B336(689), 1993.

[23] Z. Cai, R. Lazarov, T. Manteuffel, and S. McCormick. First-order system least squares for second-order partial differential equations: Part I. SIAM J. Numer. Anal., 31:1785, 1994.

[24] Z. Cai, T. Manteuffel, and S. McCormick. First-order system least squares for second-order partial differential equations: Part II. SIAM J. Numer. Anal., 34:425, 1997.

[25] M. Clark. Quantum chromodynamics multigrid (qcdmg) software package. website: http://lattice.bu.edu/qcdmg. 2008.

[26] A. Codd, T. Manteuffel, and S. McCormick. Multilevel first-order system least squares for nonlinear elliptic partial differential equations. SIAM J. Numer. Anal., 41:2197–2209, 2003.

[27] M. Creutz. Quarks, Gluons and Lattices. Cambridge University Press, 1983.

[28] M. Creutz. Fun with Dirac eigenvalues. Arxiv preprint hep-lat/0511052, 2005.

[29] T. DeGrand and C. DeTar. Lattice Methods for Quantum Chromodynamics. World Scientific, 2006.

[30] R. G. Edwards. Numerical Simulations in Lattice Gauge Theories and Statistical. Mechanics. Ph.D. thesis, New York University, 1989.

[31] R. Friedberg, TD Lee, Y. Pang, and HC Ren. Noncompact lattice formulation of gauge theories. Physical Review D, 52(7):4053–4081, 1995.

[32] D. Griffiths. Introduction to Elementary Particles. Wiley-VCH, 2004.

[33] M. Hestenes and E. Steifel. Methods of conjugate gradients for solving linear systems. J. Research NBS, 49:400–436, 1952.

[34] J. K. Hunter and B. Nachtergaele. Applied Analysis. World Scientific, 2001.

[35] D. B. Kaplan. A method for simulating chiral fermions on the lattice. Phys. Lett. B, 288(342), 1992.

[36] C. Ketelsen, T. Manteuffel, S. McCormick, and J. Ruge. Finite element methods for quantum electrodynamics based on a helmholtz decomposition of the gauge field. Num. Lin. Alg., 2009. Submitted.

[37] M. Lusher. Local coherence and deflation of low quark modes in lattice QCD. Phys. Lett. B, 417(141), 1998.

[38] M. Lusher. Local coherence and deflation of low quark modes in lattice QCD. High Energy Phys., 2007(81):1126–6708, 2007.

[39] S. MacLachlan. Improving Robustness in Multiscale Methods. Ph.D. thesis, University of Colorado at Boulder, 2004.

[40] T. Manteuffel and K. Ressel. A least-squares finite element solution of the neutron transport equation in diffusive regimes. SIAM J. Numer. Anal., 85:806–835, 1998.

[41] S. McCormick. Multilevel first-order system least squares (FOSLS) for Helmholtz equations. Procs. Conf. Maxwell Equations, 1993.

[42] J. C. Nédélec. A new family of mixed finite elements in R^3. Numerische Mathematik, 50(1):57–81, 1986.

[43] C. W. Osterlee, A. Schuller, and U. Trottenberg. Multigrid. Academic Press, 2000.

[44] P.A. Raviart and J.M. Thomas. A mixed finite element method for second order elliptic problems. Aspects of the Finite Element Method, Lectures Notes in Math., 606, 1977.

[45] C. Rebbi. Chiral-invariant regularization of fermions on the lattice. Physics Letters B, 186(2):200–204, 1987.

[46] O. Roehrle. Multilevel FOSLS Quasilinear Elliptic Partial Differential Equations. Ph.D. thesis, University of Colorado at Boulder, 2004.

[47] Y. Saad. Iterative Methods for Sparse Linear Systems. SIAM Books, 2003.

[48] Y. Shamir. Chiral fermions from lattice boundaries. Nucl. Phys., B406(90), 1993.

[49] P. Vaněk, J. Mandel, and M. Brezina. Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. Computing, 56(3):179–196, 1996.

[50] P. Wesseling. Principles of Computational Fluid Dynamics. Springer, 2001.

[51] C. Westphal. FOSLS for Geometrically-Nonlinear Elasticity in Nonsmooth. Domains. Ph.D. thesis, University of Colorado at Boulder, 2004.

[52] K.G. Wilson. Confinement of quarks. Physical Review D, 10(8):2445–2459, 1974.

[53] A. Zee. Quantum Field Theory in a Nutshell. Princeton University Press, 2003.

[54] X. Zhang. Multilevel Schwarxz Methods. Numer. Math, 63:521–539, 1992.