Overlap topological charge and axial anomaly for lattice fermions with overlap-Dirac operator

David H. Adams

Dept. of Pure Mathematics, University of Adelaide, S.A. 5005, Australia.
email: dadams@maths.adelaide.edu.au

Abstract

We study the classical continuum limit of the overlap topological charge $Q_{\text{lat}}^{(m)}$. For $m = r = 1$, and with either periodic or anti-periodic boundary conditions on the spinor fields, we prove that $Q_{\text{lat}}$ reduces to the continuum topological charge $Q$ when the continuum gauge field is in a singular gauge, vanishes outside a bounded region, and is pure gauge in a neighbourhood of the singularity. In the non-singular ($Q = 0$) case this result is established for general values of the parameters $m$ and $r$. The calculation also determines the continuum limit of the density function for $Q_{\text{lat}}$, which is seen to reproduce the density function for $Q$ when $0 < m < 2$. This can be viewed as an exact derivation of the classical continuum limit of the axial anomaly for lattice fermions with overlap-Dirac operator.

1 Introduction

The overlap formalism [1] leads to a topological charge $Q_{\text{lat}}^{(m)}$ for lattice gauge fields, defined as the spectral flow of a hermitian Dirac–Wilson operator $H$ ((2.5) below) as the mass parameter increases from $-\infty$ to $m$.\(^1\) The spectrum of $H(m)$ is known to be symmetric and without zero for $m < 0$, hence $Q_{\text{lat}}^{(m)}$ is half the spectral asymmetry of $H(m)$:

$$Q_{\text{lat}} = \frac{1}{2} \text{Tr} \left( \frac{H}{\sqrt{H^2}} \right)$$ (1.1)

\(^1\)In the notation we are suppressing a parameter $r$ which also appears in $H$. 

1
$Q_{\text{lat}}$ also arises as an index:

$$Q_{\text{lat}} = \text{index}(D)$$

(1.2)

where

$$D = \frac{1}{a} \left( 1 + \gamma_5 \frac{H}{\sqrt{H^2}} \right)$$

(1.3)

is Neuberger’s overlap-Dirac operator [2]. This operator satisfies the Ginsparg–Wilson relation [3]. In [4], Hasenfratz, Laliena and Niedermayer noted that such operators have a well-defined index (the nullspace of the operator is invariant under $\gamma_5$), and eq.(1.2) follows from an index formula derived there. $Q_{\text{lat}}$ (and related quantities) have been studied numerically by Edwards, Heller and Narayanan (see [5] and the ref.’s therein), and by Chiu [6, 7]. In [8], Edwards et. al. studied the spectral flow of $H$ for smooth SU(2) instantons on the lattice. They found that under suitable conditions, e.g. anti-periodic boundary conditions with the instanton in a singular gauge, $Q_{\text{lat}}^{(m=2)}$ coincided with the continuum topological charge $Q$ (as expected from the overlap formalism). In this paper the first analytic result in this direction is presented:

**Theorem 1.** Let $A$ be a smooth SU(N) gauge field on $\mathbb{R}^4 - \{0\}$ which (i) vanishes outside a bounded region of $\mathbb{R}^4$, and (ii) is pure gauge in a neighbourhood of 0. It may be singular at 0, corresponding to a non-trivial topological charge $Q$. Then, with either periodic or anti-periodic boundary conditions on the spinor fields (on which $H$ acts), $Q_{\text{lat}}$ reduces to $Q$ in the classical continuum limit when $m = r = 1$, i.e.

$$\lim_{a \to 0} \lim_{N_1, \ldots, N_4 \to \infty} Q_{\text{lat}}^{(m=r=1)}(m=r=1) = Q$$

(1.4)

Here $Q_{\text{lat}}$ is the overlap topological charge of the lattice transcript of $A$ (with the lattice arranged as in §2 to avoid singularities in the resulting lattice field), $a$ is the lattice spacing and $2N_\mu$ is the number of sites along the $\mu$-axis. In addition, when $A$ is smooth at the origin (the topologically trivial case), eq. (1.4) holds (with $Q = 0$) for all $r > 0$ and $m \in \mathbb{R}$, $m \neq 0, 2, 4, 6, 8$.

Examples of topologically non-trivial gauge fields of this kind are readily obtained as follows. Take a smooth map $\phi : S^3 \to \text{SU}(N)$ with degree $Q$ and define $\tilde{\phi} :$
\( \mathbb{R}^4 - \{0\} \to \text{SU}(N) \) by \( \tilde{\phi}(y, t) = \phi(y), \ y \in S^3, \ t \in \mathbb{R}_+ \), where we are identifying \( \mathbb{R}^4 - \{0\} \) with \( S^3 \times \mathbb{R}_+ \) in the obvious way. Choose a smooth function \( \lambda: \mathbb{R}^4 \to \mathbb{R} \) with \( \lambda = 1 \) in a neighbourhood of 0 and vanishing outside a bounded region. Then the field

\[
A_\mu(x) = \lambda(x) \tilde{\phi}(x) \partial_\mu \tilde{\phi}^{-1}(x)
\]

is singular at \( x = 0 \) with topological charge \( Q \), and is pure gauge in a neighbourhood of 0 and vanishes outside the bounded region.

The overlap topological charge \( Q_{\text{lat}}^{(m)} \) is defined for all lattice gauge fields for which \( H(m) \) has no zero-modes (see (1.1)). In [9], Hernández, Jansen and Lüscher showed that \( H \) does not have zero-modes when \( m = r = 1 \) and the curvature of the lattice field satisfies a certain smoothness condition ((2.7) below). Thus the lattice fields satisfying this condition split into topological sectors labelled by \( Q_{\text{lat}}^{(m=r=1)} \). Theorem 1 is then a statement on how this topological structure captures topological structure of continuum gauge fields in the classical continuum limit. This is analogous to a result established many years ago by Lüscher [10]. The present setup has some advantages compared to Lüscher’s: the expression for his lattice topological charge – a direct discretisation of a geometric definition of the continuum charge \( Q \) – is quite messy, in contrast to the elegant definition of \( Q_{\text{lat}} \) in terms of spectral flow. Also, it has no apparent interpretation as the index of an operator on lattice spinor fields; in contrast, \( Q_{\text{lat}} \) is given by the index formula (1.2).

The formula (1.1) leads to a density function for \( Q_{\text{lat}} \):

\[
Q_{\text{lat}}(x) = \frac{1}{2} \text{Tr} \left( \frac{\hat{\delta}_x H}{a^4 \sqrt{H^2}} \right)
\]

where \((\hat{\delta}_x \psi)(y) = \delta_{xy} \psi(x)\).

**Theorem 2.** When \( A \) is smooth at the origin (the topologically trivial case),

\[
\lim_{a \to 0} \lim_{N_1, \ldots, N_4 \to 0} Q_{\text{lat}}^{(m)}(x) = -\frac{I(m)}{32 \pi^2} \epsilon_{\mu\nu\rho\sigma} \text{tr}(F_{\mu\nu}(x)F_{\rho\sigma}(x))
\]

where

3
\[ I(m) = \begin{array}{|c|c|c|c|c|c|}
\hline
m & 0 < m < 2 & 2 < m < 4 & 4 < m < 6 & 6 < m < 8 & m \notin [0, 8] \\
\hline
1 & -3 & 3 & -1 & 0 & \ \\
\hline
\end{array} \] (1.8)

(The limit is ill-defined for \( m = 0, 2, 4, 6, 8 \).) This result also holds in the topologically non-trivial case when \( m = r = 1 \).

Thus \( Q_{\text{lat}}^{(m)}(x) \) reduces to the continuum topological charge density in the classical continuum limit under the above conditions when \( 0 < m < 2 \). We note that the table (1.8) reflects previously derived properties of \( Q_{\text{lat}}^{(m)} \) as a function of \( m \) : \( Q_{\text{lat}}^{(m)} = 0 \) for \( m \notin [0, 8] \) [1] and has the reflection symmetry \( Q_{\text{lat}}^{(m)} = -Q_{\text{lat}}^{(8-m)} \) [7].

\( Q_{\text{lat}}(x) \) also arises as the local axial anomaly (modulo the usual factor \( 2i \)) for lattice fermions with action specified by the overlap-Dirac operator (1.3), either via the jacobian for Lüscher’s exact lattice-deformed chiral symmetry transformation [11], or via the standard derivation [12] of the axial anomaly in the lattice theory (as was done in [4])\(^2\). Thus theorem 2 says that the lattice axial anomaly exactly reproduces the continuum anomaly in the classical continuum limit under suitable conditions. This result has previously been obtained in a perturbative framework by Kikukawa and Yamada [13].\(^3\) The corresponding exact result for Dirac-Wilson fermions was established a long time ago by Kerler [17] and Seiler and Stamatescu [18] and our arguments have some similarities with theirs.

In \S\ 2 the relevant definitions are recalled and \( Q_{\text{lat}} \) is shown to be well-defined in the classical continuum limit considered in the theorems above. In \S\ 3 the proof of theorems 1, 2 is sketched, and in \S\ 4 we make some concluding remarks. A more general version of this work will be given in a forthcoming paper [19].

\(^2\)I thank Prof. W. Kerler for pointing this out to me.

\(^3\)General arguments relating the axial anomaly for solutions to the GW relation to the continuum anomaly expression were also given in [4], drawing on the older work of Ginsparg and Wilson [14]. See [15] for a detailed exposition of the Ginsparg–Wilson paper in the context of current developments. The details of how the axial anomaly arises in the overlap formalism have been given in [16].
2 The setup

We take the finite lattice in Euclidean $\mathbf{R}^4$ to have spacing $a$ and sites \( \{ x = a(n_1 - 1/2, n_2 - 1/2, n_3 - 1/2, n_4 - 1/2) \}, -N_\mu < n_\mu \leq N_\mu \). Note that there is no site at the origin; this allows us to consider the lattice transcript

\[ U_\mu(x) = T \exp \left( \int_0^1 aA_\mu(x + tae_\mu) \, dt \right) \tag{2.1} \]

of a continuum gauge field $A$ with singularity at the origin. (T=t-ordering, $e_\mu$=unit vector in the positive $\mu$-direction and for simplicity the coupling constant has been set to unity.)

With suitable boundary conditions on the lattice gauge fields $U_\mu(x)$ and spinor fields $\psi(x)$, the covariant finite difference operators

\[ \nabla^+_\mu \psi(x) = U_\mu^{-1}(x)\psi(x + ae_\mu) - \psi(x) \tag{2.2} \]
\[ \nabla^-_\mu \psi(x) = \psi(x) - U_\mu(x - ae_\mu)\psi(x - ae_\mu) \tag{2.3} \]

are well-defined and $(\nabla^\pm_\mu)^* = -\nabla^{\mp}_\mu$ with respect to the inner product

\[ \langle \psi_1, \psi_2 \rangle = a^4 \sum_x \psi_1(x)^* \psi(x). \tag{2.4} \]

For example, one can impose the conditions $U_\mu(x) = 1$ at the boundary and either periodic or anti-periodic boundary conditions on $\psi$ – this is the case we will be mainly dealing with in this paper. The $\gamma$-matrices are taken to be hermitian, then $\nabla := \gamma^{1/2}_5(\nabla^+_\mu + \nabla^-_\mu)$ is skew-hermitian and $\Delta := -\nabla^-\nabla^+_\mu$ is hermitian and positive. The hermitian Dirac–Wilson operator is\(^4\)

\[ H = \frac{1}{a} X \gamma_5 \tag{2.5} \]

where

\[ X = \nabla + r\left( \frac{1}{2}\Delta - m \right), \quad m \in \mathbf{R}, \quad r > 0. \tag{2.6} \]

\(^4\)One can also set $H = \gamma_5 \frac{1}{a} X$ (the usual definition in the literature); this has the same spectrum as (2.5) since the two operators are related by conjugation by $\gamma_5$. 

5
The overlap topological charge $Q_{\text{lat}}$ given by (1.1) is well-defined provided $X$ does not have zero-modes. In [9] it was shown that when $m=r=1$ and the curvature of the lattice gauge field satisfies the smoothness condition

$$|1 - U_{\mu\nu}(x)| \leq \epsilon$$

(2.7)

then

$$X^*X \geq 1 - 30\epsilon.$$  

(2.8)

Thus $X$ has no zero-modes when $\epsilon < 1/30$. Here $U_{\mu\nu}(x)$ is the holonomy around the plaquette specified by $x, \mu, \nu$.

Consider the case where the lattice field $U_\mu(x)$ is determined by a continuum gauge field $A$ via (2.1). Expanding

$$U_\mu(x) = 1 + aA_\mu(x) + O(a^2)(x)$$

(2.9)

$$U_{\mu\nu}(x) = 1 + a^2F_{\mu\nu}(x) + O(a^3)(x)$$

(2.10)

we see that if $A$ is smooth and $A_\mu(x)$ and $F_{\mu\nu}(x)$ are bounded at infinity, then for small lattice spacing $a$ there are bounds

$$|1 - U_\mu(x)| \leq K_1a$$

(2.11)

$$|1 - U_{\mu\nu}(x)| \leq K_2a^2$$

(2.12)

where $K_1, K_2$ depend on $A$ but are independent of $a$ and the $N_\mu$’s. The first bound breaks down if $A$ has singularities. However, when $A$ satisfies the conditions of theorem 1 the second bound continues to hold if $A$ is singular at the origin (since the holonomy around plaquettes contained in the neighbourhood of the origin where $A$ is pure gauge is trivial). It follows that in this case $X^{(m=r=1)}$ has no zero-modes and $Q_{\text{lat}}^{(m=r=1)}$ is well-defined for all sufficiently small $a$. In the non-singular case, the bound (2.11) leads to a similar result holding for all $r > 0$ and $m \neq 0, 2, 4, 6, 8$ as we now show. First, when $A$ is trivial the operator $X^*X$ is diagonal in a plane wave basis; its elements can be calculated to be

$$X_0^*X_0(ak) = \sum_\mu \sin^2(ak_\mu) + r^2 \left(-m + \sum_\mu 1 - \cos(ak_\mu)\right)^2$$

(2.13)
\(X_0 = X^{(U=1)}\). From this we see that when \(m \neq 0,2,4,6,8\) there is a \(\delta > 0\) independent of \(a\) and the \(N_\mu\)'s such that

\[
X_0^*X_0 \geq \delta \tag{2.14}
\]

Now, noting that \(\nabla^\pm = \nabla^\pm (U=1) + \chi_\mu^\pm\) where \(\chi_\mu^+\psi(x) = (U_\mu^{-1}(x) - 1)\psi(x + ae_\mu)\) and \(\chi_\mu^-\psi(x) = (1 - U_\mu(x - ae_\mu))\psi(x - ae_\mu)\), and combining \(||\chi_\mu^\pm|| \leq K_1a\) (due to (2.11)) with standard bounds (e.g. \(||\nabla_\mu^\pm|| \leq 2\)) it is easy to establish a bound

\[
||X^*X - X_0^*X_0|| \leq K_3a \tag{2.15}
\]

with \(K_3\) independent of small \(a\) and the \(N_\mu\)'s. Combining this with (2.14) we see that

\[
X^*X \geq \delta/2 \quad \text{for} \quad a \leq \delta/2K_3, \tag{2.16}
\]

proving that also in this case \(X\) has no zero-modes for sufficiently small \(a\) as claimed.

### 3 Proof of theorems 1,2

From now on, the lattice gauge field \(U_\mu(x)\) is determined via (2.1) by a continuum field \(A\) satisfying the conditions of theorem 1. Noting that \(X^* = \gamma_5X\gamma_5\) we have

\[
Q_{lat} = \frac{1}{2} \text{Tr}(\gamma_5 \frac{X^*}{\sqrt{X^*X}}). \tag{3.1}
\]

The strategy is to carry out certain power series expansions of the inverse square root which in the end lead to an expansion in the lattice spacing \(a\). A calculation gives

\[
X^*X = -\nabla^2 + r^2(\frac{1}{2}\Delta - m)^2 + V \tag{3.2}
\]

where

\[
V = r^2\gamma_\mu V_\mu + \frac{1}{4}[\gamma_\mu, \gamma_\nu]V_{\mu\nu} \tag{3.3}
\]

\[
V_\mu = \frac{1}{2}[(\nabla_\mu^+ + \nabla_\mu^-) \cdot \sum_\nu(\nabla_\nu^+ - \nabla_\nu^-)] \tag{3.4}
\]

\[
V_{\mu\nu} = -\frac{1}{4}[(\nabla_\mu^+ + \nabla_\mu^-) \cdot (\nabla_\nu^+ + \nabla_\nu^-)] \tag{3.5}
\]
Noting that
\[ \left[ \nabla_\mu^+, \nabla_\nu^+ \right] \psi(x) = (U_\nu(x) - 1)U_\mu^{-1}(x)a \psi(x + a \mu + a \nu) \] (3.6)
and similar expressions for the other commutators, the bound (2.12) gives
\[ \| \left[ \nabla_\mu^+, \nabla_\nu^+ \right] \| \leq K_2 a^2 \] (3.7)
for small \( a \). It follows that there is a constant \( K_4 \) independent of small \( a \) and the \( N_\mu \)'s such that
\[ \| V \| \leq K_4 a^2 \] (3.8)
The bounds (2.8) and (2.12) imply \( X^*X \geq 1/2 \) for \( a \leq 1/\sqrt{15K_2} \). Set
\[ L := -\nabla^2 + r^2(\frac{1}{2} \Delta - m)^2 =: G^{-1} \] (3.9)
so that \( X^*X = L + V \). Then \( L \geq 1/4 \) for \( a \leq \min\{ \frac{1}{\sqrt{4K_4}}, \frac{1}{\sqrt{15K_2}} \} \), so \( G = L^{-1} \) exists for such \( a \) and has the bound \( \| G \| \leq 4 \) for sufficiently small \( a \). Then \( \| GV \| \leq 4K_4 a^2 \) and for small \( a \) we can expand
\[ \frac{1}{\sqrt{X^*X}} = \sqrt{G} \frac{1}{\sqrt{1 + GV}} \] (3.10)
as a power series in \( GV \). Noting that the \( \gamma \)-matrices only appear in \( V \), and using the fact that only terms involving products of at least 4 \( \gamma \)-matrices contribute to the trace over spinor indices in (3.1), we get
\[ Q_{lat} = \frac{1}{2} \text{Tr} \left[ \gamma_5 X^* \sqrt{G} \left( \sum_{n=0}^{\infty} c_n (GV)^n \right) \right] \] (3.11)
\[ = \frac{3}{16} \text{Tr} (\gamma_5 X^* G^{3/2} VGV) + \frac{1}{2} \sum_{n=3}^{\infty} c_n \text{Tr} (\gamma_5 X^* \sqrt{G} (GV)^n) \] (3.12)
where \( c_n = (-1)^n (1 \cdot 3 \cdot 5 \cdots (2n - 1))/2^n n! \) is the n’th coefficient in the Taylor expansion of \( 1/\sqrt{1+x} \) around zero. The sum in (3.12) vanishes in the limit \( a \to 0 \) as we now show. Here and in the following we use the correspondence between operators \( O \) and functions \( O(x, y) \) via
\[ O\psi(x) = \sum_{y} a^4 O(x, y) \psi(y) \] (3.13)
and exploit the bound

\[ a^4 |\mathcal{O}(x,y)| \leq ||\mathcal{O}||. \] (3.14)

Let \( R \subset \mathbb{R}^4 \) denote the bounded region outside which \( A \) vanishes. It is clear from (3.6) that we can choose a bounded region \( R_1 \) containing \( R \) such that for \( a < 1 \) \( V(x,y) \) vanishes unless \( x,y \in R_1 \). Let \( \delta_x \) denote the lattice field (without gauge or spinor indices) given by \( \delta_x(y) = \delta_{xy} \). Then \( \{a^4 \delta_x\} \) is an o.n.b. for the space of such fields, and the sum in (3.12) can be written as

\[
\frac{1}{2} \text{tr} \sum_{n=3}^\infty c_n \sum_x \langle \left( \frac{1}{a^4} \delta_x \right), \gamma_5 X^* \sqrt{G} (G V)^n \left( \frac{1}{a^4} \delta_x \right) \rangle.
\] (3.15)

The \( n \)'th term can be written as

\[
\frac{1}{2} c_n \text{tr} \sum_{x,y} (a^4)^2 \left( \gamma_5 X^* \sqrt{G} (G V)^{n-1} G \right) (x,y) V(x,y).
\] (3.16)

Since \( V(x,y) \) vanishes for \( x \neq R_1 \) we see that the sum over \( x \) in (3.15) can be restricted to \( x \in R_1 \). Using the Cauchy-Schwartz inequality (and \(|c_n| \leq 1\)) we now get

\[
|\text{tr}_{(3.15)}| \leq d_{g,s} \sum_{n=3}^\infty \sum_{x \in R_1} ||\gamma_5 X^* \sqrt{G} (G V)^n||
\leq d_{g,s} \sum_{n=3}^\infty ||X^*|| ||G||^{n+1/2} (K_4 a^2)^n \left( \sum_{x \in R_1} 1 \right)
= a^2 d_{g,s} ||X^*|| ||G||^{7/2} K_4^3 \left( \sum_{x \in R_1} a^4 \right) \sum_{n=0}^\infty (||G|| K_4 a^2)^n
\]

where \( d_{g,s} = \) the product of the dimensions of the gauge group and spinor representation. For sufficiently small \( a \), \( \sum_{x \in R_1} a^4 \) is bounded by a constant \( K_{R_1} > \) volume of \( R_1 \), and we get

\[
|\text{tr}_{(3.15)}| \leq a^2 d_{g,s} ||X^*|| ||G||^{7/2} K_4^3 K_{R_1} (1 - ||G|| K_4 a^2)^{-1}
\] (3.17)

which vanishes for \( a \to 0 \), proving the claim. (Recall \( ||G|| \leq 4 \) for small \( a \) and note \( ||X^*|| \leq 8 + r(8 + |m|) \) by standard bounds.)

We now turn to the first term in (3.12),

\[
\frac{3}{16} \text{Tr}(\gamma_5 X^* G_A^3 V G_A V)
\] (3.18)
where we have explicitly indicated the dependence of \( G = G_A \) on the continuum gauge field \( A \). We wish to expand \( G_A \) as a power series in \( a \) (or equivalently in \( A \), since they enter as a product \( aA \) in (2.1)). This cannot be done a priori though, since the field \( A \) has a singularity at the origin. To get around this, we show that \( G_A \) may be replaced by \( G_{\tilde{A}} \) in (3.18) (modulo terms which vanish in the \( a \to 0 \) limit) where \( \tilde{A} \) is a smooth field vanishing outside a bounded region like \( A \), but without any singularities. A power series expansion of \( G_{\tilde{A}} \) in \( a \) is then permissible by the following reasoning: For \( a < 1 \) there is a bound

\[
||L_{\tilde{A}} - L_0|| \leq K_5 a
\]  

obtained in an analogous way to the bound (2.15) on \( X^*X - X_0^*X_0 \) in the non-singular case. Since \( L_0 = X_0^*X_0 \geq \delta \) cf. (2.14), it follows that \( L_{\tilde{A}} \geq \delta/2 \) for \( a \leq \delta/2K_5 \), so for such \( a \)

\[
G_{\tilde{A}} = L_{\tilde{A}}^{-1} = \frac{1}{X_0^*X_0 + (L_{\tilde{A}} - L_0)}
\]  

exists, has a bound \( ||G_{\tilde{A}}|| \leq 2/\delta \) independent of \( a \) and the \( N_\mu \)'s, and admits a power series expansion in \( L_{\tilde{A}} - L_0 \). In turn, \( L_{\tilde{A}} - L_0 \) has a power series expansion in \( a \) (obtained by expanding \( 1 - U_{\tilde{A}} \) in \( a \)) which is \( O(a) \) for \( a \to 0 \). Thus \( G_{\tilde{A}} \) has an expansion in \( a \) of the form

\[
G_{\tilde{A}} = (X_0^*X_0)^{-1} + O(a)
\]  

with \( O(a) \) having a bound

\[
||O(a)|| \leq K_6 a
\]  

with \( K_6 \) independent of small \( a \) and the \( N_\mu \)'s.

To show that \( G_A \) can be replaced by \( G_{\tilde{A}} \) we first write (3.18) as

\[
\frac{3}{16} \text{tr} \left( \gamma_5 \sum_{x,y_1,\ldots,y_4} (a^4)^5 X^*(x, y_1) G_{\tilde{A}}^{3/2}(y_1, y_2) V(y_2, y_3) G_A(y_3, y_4) V(y_4, x) \right)
\]  

Let \( B \subset \mathbb{R}^4 \) denote the neighbourhood of the origin where \( A \) is pure gauge. The holonomy of \( A \) around a loop in \( B \) is trivial. This observation together with (3.6)
implies that we can choose a neighbourhood $B_1$ of the origin, with $B_1 \subset B$, such that when $a$ is sufficiently small $V(x, y)$ vanishes. With $R_1 \supset R$ as before (so $V(x, y) = 0$ unless $x, y \in R_1$) we exploit the fact that $X^*$ is local to choose a bounded region $R_2 \supset R_1$ such that when $a$ is sufficiently small, $\forall x \in R_1 \ X^*(x, y) = 0$ if $y$ is outside $R_2$. Then, for small $a$, the summand in (3.23) vanishes unless $x, y_1, \ldots, y_4 \in R_2 \setminus B_1$. Now choose a neighbourhood $B_2$ of the origin, with $B_2 \subset B_1$, such that there is a minimum distance $b > 0$ between the boundaries of $B_2$ and $B_1$ with respect to the norm

$$|x|_1 = \sum_{\mu} |x_{\mu}|. \quad (3.24)$$

Choose a smooth function $\lambda : \mathbb{R}^4 \to \mathbb{R}$ such that $\lambda = 1$ outside of $B_2$ and $\lambda = 0$ in a neighbourhood of the origin. Set $\tilde{A} = \lambda A$. Then the key to replacing $G_A$ by $G_{\tilde{A}}$ in (3.18) is the bound

$$|G_A(x, y) - G_{\tilde{A}}(x, y)| \leq \left( \frac{2\kappa}{1 - e^{-\theta}} \right) \frac{1}{a^4} e^{-\theta b^2/2a} \quad \text{for } x, y \notin B_1 \quad (3.25)$$

where $\kappa$ and $\theta > 0$ are constants defined independent of small $a$ and the $N_\mu$’s as follows. Recall that $L_A \geq 1/4$ for small $a$, and note that $L_0 = X^*_0 X_0 \geq \delta$, together with (3.19), implies $L_{\tilde{A}} \geq \delta/2$ for sufficiently small $a$. Thus there are constants $u = (\min\{1/4, \delta/2\})^2 > 0$ and $v < \infty$ such that $u \leq L^2_A \leq v$ and $u \leq L^2_{\tilde{A}} \leq v$ for sufficiently small $a$. (The existence of $v$ follows from standard bounds.) Then (3.25) holds with $\theta = \cosh^{-1}\left( \frac{v}{v-u} \right)$ and $\kappa = (\frac{4\kappa}{v-u})$. We prove the bound in an appendix. It is similar to a result of Hernández et. al. in [9], where it was shown that $(X^*X)^{-1/2}$ is local modulo an exponentially decaying tail with cut-off rate proportional to $1/a$. Our argument is a variant of theirs (with $(X^*X)^{-1/2}$ replaced by $G = L^{-1} = (L^3)^{-1/2}$) and uses the fact that $A$ and $\tilde{A}$ coincide except in a neighbourhood of 0 contained in $B_2$.

A bound similar to (3.25) holds for $|G_A^{3/2}(x, y) - G_{\tilde{A}}^{3/2}(x, y)|$ (this follows from the argument in the appendix after writing $G^{3/2} = (L^3)^{-1/2}$). After substituting $G_{\tilde{A}}(x, y) + (G_A(x, y) - G_{\tilde{A}}(x, y))$ for $G_A(x, y)$ in (3.23), and making a similar substitution for $G_{\tilde{A}}^{3/2}(x, y)$, we see that the terms involving $G_A(x, y) - G_{\tilde{A}}(x, y)$ and
$G_{A}^{3/2}(x, y) - G_{A}^{3/2}(x, y)$ vanish in the $a \to 0$ limit. For example, using (3.14) we have

$$\left| \sum_{x, y_1, \ldots, y_4 \in R_2 \backslash B_1} (a^4)^5 X^*(x, y) G_{A}^{3/2}(y_1, y_2) V(y_2, y_3)(G_A(y_3, y_4) - G_{A}^{3/2}(y_3, y_4)) V(y_4, x) \right|$$

$$\leq \sum_{x, y_1, \ldots, y_4 \in R_2 \backslash B_1} ||X^*|| ||G_{A}||^{3/2} ||V||^2 \left( \frac{2\kappa}{1 - e^{-\theta}} \right) e^{-\theta b/2a}$$

$$\leq \frac{1}{a^{20}} e^{-\theta b/2a} \left( \frac{2\kappa}{1 - e^{-\theta}} \right) ||X^*|| ||G_{A}||^{3/2} K^2_1 \left( \sum_{x \in R_2} a^4 \right) \left( \sum_{y_1 \in R_2} a^4 \right) \cdots \left( \sum_{y_4 \in R_2} a^4 \right)$$

This vanishes in the limit $a \to 0$ since $R_2$ has finite volume. It follows that

$$\text{Tr}(\gamma_5 X^* G_{A}^{3/2} V G_A V) = \text{Tr}(\gamma_5 X^* G_{A}^{3/2} V G_A V) + O(a) \quad (3.26)$$

as claimed. Substituting this in (3.12), and recalling that the sum there vanishes in the $a \to 0$ limit, we arrive at

$$Q_{\text{lat}} = \frac{3}{16} \text{Tr}(\gamma_5 X^* G_{A}^{3/2} V G_A V) + O(a) \quad (3.27)$$

We now expand $G_{A}$ in powers of $a$ as discussed above (recall (3.21)--(3.22)). Only the lowest order term contributes to (3.27) in the $a \to 0$ limit; the others can be seen to vanish by a similar argument to the one leading to (3.17) above. For the same reason, we can replace $X^*$ by $X^*_0$ in (3.27). This gives

$$Q_{\text{lat}} = \frac{3}{16} \text{Tr}(\gamma_5 X^*_0 G_{0}^{3/2} V G_0 V_A) + O(a) \quad (3.28)$$

where the dependence of $V$ on $A$ is now explicitly indicated. Since $V_A(x, y)$ vanishes for $x, y$ outside $R_1$ or inside $B_1$, the first term can be replaced by

$$\frac{3}{16} \text{Tr}(\gamma_5 X^*_0 G_{0}^{3/2} V G_0 V \hat{1}_{R_1 \backslash B_1}) \quad (3.29)$$

where $\hat{1}_{R_1 \backslash B_1}(x) = \psi(x)$ for $x \in R_1 \backslash B_1$ and zero otherwise. The trace over spinor indices in this expression can be evaluated to give

$$\text{Tr}(Q \hat{1}_{R_1 \backslash B_1}) \quad (3.30)$$

where

$$Q = \frac{3\epsilon_{\mu\nu\rho\sigma}}{16} \left[ (\frac{1}{2} \Delta^{(0)} - m) G_{0}^{3/2} \tilde{V}_{\mu\nu} G_0 V_{\rho\sigma} - \nabla_\mu^{(0)} G_{0}^{3/2} (V_{\rho}^{A} G_0 V_{\sigma}^{\tilde{A}} + V_{\rho}^{\tilde{A}} G_0 V_{\sigma}^{A}) \right] \quad (3.31)$$
Recall that $V_{\tilde{A}}$ is given in terms of commutators of the $\tilde{\nabla}_{\mu}^\pm$’s (see (3.3)–(3.5)) where $\tilde{\nabla}_{\mu}^\pm$ denotes the covariant finite difference operator specified by $\tilde{A}$. These have expansions

\begin{align}
[\tilde{\nabla}_{\mu}^+ , \tilde{\nabla}_{\nu}^+] \psi(x) &= -a^2 \tilde{F}_{\mu\nu}(x) \psi(x + ae_\mu + ae_\nu) + \mathcal{O}_1(a^3) \psi(x) \tag{3.32} \\
[\tilde{\nabla}_{\mu}^+ , \tilde{\nabla}_{\nu}^-] \psi(x) &= -a^2 \tilde{F}_{\mu\nu}(x) \psi(x + ae_\mu - ae_\nu) + \mathcal{O}_2(a^3) \psi(x) \tag{3.33} \\
[\tilde{\nabla}_{\mu}^- , \tilde{\nabla}_{\nu}^-] \psi(x) &= -a^2 \tilde{F}_{\mu\nu}(x) \psi(x - ae_\mu + ae_\nu) + \mathcal{O}_3(a^3) \psi(x) \tag{3.34} \\
[\tilde{\nabla}_{\mu}^- , \tilde{\nabla}_{\nu}^+] \psi(x) &= -a^2 \tilde{F}_{\mu\nu}(x) \psi(x - ae_\mu - ae_\nu) + \mathcal{O}_4(a^3) \psi(x) \tag{3.35}
\end{align}

Since $\tilde{A}$ is smooth the operators $\mathcal{O}_j(a^3)$ can be shown to have bounds $||\mathcal{O}_j(a^3)|| \leq K a^3$ for some constant $K$ independent of small $a$ and the $N_\mu$’s. Using this, the contributions of the $\mathcal{O}_j(a^3)$’s to (3.30) can be seen to vanish in the $a \to 0$ limit. We now evaluate the trace in (3.30)–(3.31) using the plane wave o.n. basis \{\$\psi_k(x) = \frac{1}{\sqrt{N}} e^{i k \cdot x}\$\} where $N = \prod_\mu (2N_\mu)$. With the lattice as in §2, and with periodic or anti-periodic boundary conditions, the values of $k$ can be chosen as follows:

\begin{align}
\text{periodic : } & k_\mu \in \frac{2 \pi}{a N_\mu} \{-N_\mu, -N_\mu + 1, \ldots, N_\mu - 1\} \tag{3.36} \\
\text{anti-periodic : } & k_\mu \in \frac{\pi}{a N_\mu} \left(2\{-N_\mu, -N_\mu + 1, \ldots, N_\mu - 1\} + \frac{1}{2}\right) \tag{3.37}
\end{align}

Note that the volume per $k$ in both cases is

\begin{equation}
\Delta^4 k = \prod_\mu \frac{2 \pi}{a N_\mu} = \frac{(2 \pi)^4}{a^4} \frac{2^4}{N}. \tag{3.38}
\end{equation}

A calculation using (3.32)–(3.35) now gives

\begin{equation}
\langle \psi_k, (\mathcal{Q} \mathcal{1}_{R_1 \setminus R_1}) \psi_k \rangle = J(a) \tilde{Q}(ak) + \langle \psi_k, \mathcal{O}(a) \psi_k \rangle \tag{3.39}
\end{equation}

where

\begin{align}
J(a) &= \sum_{x \in R_1 \setminus R_1} a^4 \epsilon_{\mu\nu\rho\sigma} \text{tr}(F_{\mu\nu}(x) F_{\rho\sigma}(x)) \tag{3.40} \\
\tilde{Q}(k) &= \frac{3 \pi}{16} G_0^{5/2}(k) \prod_\mu \cos k_\mu \left[ -m + \left( \sum_\nu 1 - \cos k_\nu \right) - \sum_\rho \frac{\sin^2 k_\rho}{\cos k_\rho} \right] \tag{3.41}
\end{align}

$G_0(k) = 1/(X_0^x X_0^k)$ where $X_0^x X_0^k$ is given by (3.18), and $\mathcal{O}(a)$ involves the operators $\mathcal{O}_j(a^3)$ in (3.32)–(3.35). Using by now familiar arguments it can be checked
that $\text{Tr}(O(a))$ vanishes in the $a \to 0$ limit; it then follows from (3.39) that

\[
Q_{\text{lat}} = J(a) \sum_k \frac{1}{N} \hat{Q}(ak) + O(a)
\]

\[
= J(a) \frac{1}{24} \sum_k \frac{a^4 \Delta^4 k}{(2\pi)^4} \hat{Q}(ak) + O(a)
\]

where we have used (3.38). In both the periodic and anti-periodic cases, after changing summation variable from $k$ to $ak$, we get

\[
\lim_{N_1, \ldots, N_4 \to \infty} Q_{\text{lat}} = J(a) \int \frac{d^4 k}{(2\pi)^4} \hat{Q}(k) + O(a)
\]

The integral over $k$ is independent of $a$. Taking the $a \to 0$ limit we get

\[
\lim_{a \to 0} \lim_{N_1, \ldots, N_4 \to \infty} Q_{\text{lat}} = -\frac{I(r, m)}{32\pi^2} \int_{R_1 \setminus B_1} \epsilon_{\mu\nu\rho\sigma} \text{tr}(F_{\mu\nu}(x) F_{\rho\sigma}(x)) \, dx_1 \cdots dx_4
\]

where

\[
I(r, m) = -\frac{3r}{8} \int_{-\pi}^{\pi} d^4 k \frac{\prod_{\nu=1}^4 \cos k_{\nu} \left[ -m + \sum_{\mu}(1 - \cos k_{\mu}) - \sum_{\mu} \sin^2 k_{\mu} \cos k_{\mu} \right]}{\left[ \sum_{\mu} \sin^2 k_{\mu} + r^2 \left( -m + \sum_{\mu}(1 - \cos k_{\mu}) \right)^2 \right]^{5/2}}
\]

Since $A$ vanishes at $R_1$ and is pure gauge at $B_1$, (3.45) reduces to $I(r, m)Q$. We will show that the value of $I(r, m)$ is $I(m)$, given by the table (1.8). In particular, $I(1, 1) = 1$, proving theorem 1.

The integral $I(r, m)$ is similar to the integral (A.17) of [18], although the exponents and numerical factor are different and the parameter $m$ did not appear there. To evaluate it we exploit the symmetries of the integrand (as in [17, 18]) and change variables to $s_{\nu} \equiv \sin k_{\nu}$ to write

\[
I(r, m) = \sum_{\epsilon_{\nu} = \pm 1} \left( \prod_{\mu=1}^4 \text{sign}(\epsilon_{\mu}) \right) I(r, m, \epsilon)
\]

where

\[
I(r, m, \epsilon) = -\frac{3r}{8\pi^2} \int_{-1}^{1} d^4 s \frac{-m + \sum_{\mu}(1 - \epsilon_{\mu} \sqrt{1 - s_{\mu}^2}) - \sum_{\mu} \frac{s_{\mu}^2}{\epsilon_{\mu} \sqrt{1 - s_{\mu}^2}}}{\left[ s^2 + r^2 \left( -m + \sum_{\mu}(1 - \epsilon_{\mu} \sqrt{1 - s_{\mu}^2}) \right)^2 \right]^{5/2}}
\]

This diverges for $m = m_{\epsilon} \equiv \sum_{\mu}(1 - \epsilon_{\mu}) \in \{0, 2, 4, 6, 8\}$ but is finite for all other values of $m$. It is a constant function of $m$ with a jump at $m_{\epsilon}$; to see this set
\[ \tilde{\Delta}(s, m) = -m + \sum_\mu (1 - \epsilon_\mu \sqrt{1 - s_\mu^2}), \]  

then

\[ -8\pi^2 \frac{d}{dm} I(r, m, \epsilon) = -r \int_{-1}^{1} \frac{d^4 s}{(s^2 + r^2 \Delta^2)^{5/2}} + 5r^3 \int_{-1}^{1} d^4 s \frac{\tilde{\Delta}(1 - \sum_\nu s_\nu \frac{\partial}{\partial s_\nu}) \tilde{\Delta}}{(s^2 + r^2 \Delta^2)^{7/2}} \]  

(3.49)

Inspired by the identity eq. (A.19) of [18] (which was originally due to Karsten and Smit) we rewrite the second integral as

\[ 5r \int_{-1}^{1} d^4 s \frac{(1 - \frac{1}{2} \sum_\nu s_\nu \frac{\partial}{\partial s_\nu})(s^2 + r^2 \tilde{\Delta}^2)}{(s^2 + r^2 \Delta^2)^{7/2}} = 5r \int_{-1}^{1} d^4 s \left(1 + \frac{1}{5} \sum_\nu s_\nu \frac{\partial}{\partial s_\nu}\right)(s^2 + r^2 \tilde{\Delta}^2)^{-5/2} \]  

(3.50)

Integration by parts now gives the first integral in (3.49) except with the opposite sign, so (3.49) vanishes. \( I(r, m) \) can now be determined by evaluating \( I(r, m, \epsilon) \) in the limits \( m \to m_\epsilon \) from above and below. For \( m > m_\epsilon \), after setting \( m = m_\epsilon + \hat{m} \) and changing variables to \( \tilde{s}_\nu = s_\nu / \hat{m} \) in (3.48), we get

\[ I(r, m, \epsilon) = \frac{-3r}{8\pi^2} \int_{-1/\hat{m}}^{1/\hat{m}} d^4 \tilde{s} \hat{m}^4 \frac{-\hat{m} + O(\hat{m}^2)}{\hat{m}^2 + r^2(-1 + O(\hat{m}^2))^{5/2}} \]  

\[ \hat{m} \to 0^+ \quad \text{and} \quad \frac{-3r}{8\pi^2} \int_{-\infty}^{\infty} d^4 \tilde{s} \frac{-r}{(\tilde{s}^2 + r^2)^{3/2}} = 1/2 \]  

(3.51)

For \( m < m_\epsilon \) a similar calculation gives \( I(r, m, \epsilon) = -1/2 \). \( I(r, m) \) can now be calculated from (3.47), leading to the value \( I(m) \) in table (1.8) as promised. This completes the proof of theorem 1. Theorem 2 follows from essentially the same argument after replacing \( \gamma_5 \frac{X^*}{\sqrt{X^*X}} \) by \( \gamma_5 \frac{\delta_X}{\sqrt{X^*X}} \) in the trace in (3.1).

4 Concluding remarks

This paper is a first step in the analytic investigation of the overlap topological charge \( Q_{\text{lat}} \), and its density \( Q_{\text{lat}}(x) \), and leads on to a number of interesting problems for future work.

(i) The extension to arbitrary even dimensions is straightforward and will be given in [19]. The results will also be extended there to allow for a general riemannian background metric.

(ii) The restriction \( m = r = 1 \) in the topologically non-trivial case should be relaxed.
The problem is to find a bound $X^*X \geq \delta$ (with $\delta > 0$ independent of $a$ and the $N_\mu$’s) holding for all sufficiently small $a$. The argument of Hernández et. al. in Appendix C of [9] (together with (2.12) above) establish this for $m=r=1$; their result can be extended to $m,r$ close to 1 (as noted in [9]) but not to general $m,r$; in fact not even to all $m \in [0,2]$.

(iii) Theorem 2, concerning the density $Q_{lat}(x)$ of the overlap topological charge, has been established here in the case where the continuum gauge field vanishes outside a bounded region. However, it is plausible that this condition can be relaxed so that $A_\mu(x)$ is only required to vanish at a certain rate for $|x| \to \infty$.

(iv) The conditions in theorem 1, that the continuum field $A$ be pure gauge in a neighbourhood of the singularity and vanish outside a bounded region, excludes instantons. However, instantons can be approximated by fields satisfying these conditions, and by an approximation process it may be possible to extend theorem 1 to allow for instantons.

(v) As mentioned in the introduction, in [8] Edwards et. al. found that the spectral flow for SU(2) instantons on a finite lattice, with anti-periodic boundary conditions on $H$ and the instanton in a singular gauge, reproduced the continuum topological charge $Q$ as the parameter $m$ was varied from 0 to 2. It would be very interesting to prove this analytically. This together with the results of this paper suggests the following picture: For given continuum gauge field $A$, the values of $m$ in $[0,2]$ for which $H(m)$ has zero-modes (i.e. the points at which an eigenvalue $\lambda_j(m)$ of $H(m)$ crosses the origin) converge to the left of $[0,2]$ as the lattice spacing decreases to 0 and the $N_\mu$’s become large. An analytic verification of this picture would also be interesting.

(vi) In [8] $Q_{lat}^{(m=2)}$ was also seen to reproduce $Q$ when the instanton was in a regular gauge (i.e. smooth, with topological charge associated with “the sphere at infinity”) and with Dirichlet boundary conditions on the spinor fields on which $H$ acts. A natural question is whether the results here can be extended to this case. Different calculations will be required since, e.g., the plane wave basis can no longer be used (it is excluded by Dirichlet boundary conditions).
Q_{lat} has also been studied numerically by Chiu [6, 7]. His approach is to implement the overlap-Dirac operator D on the lattice and extract its zero-modes. These have definite chirality (since the nullspace is invariant under $\gamma_5$) and determine $Q_{lat}^{(m)} = \text{index}(D^{(m)})$ via the index formula (1.2). Chiu investigated $Q_{lat}^{(m)}$ as a function of m for U(1) gauge fields on the 2-torus [6] and some SU(2) fields on the 4-torus [7]. He found a phase structure which included a phase where $Q_{lat}^{(m)}$ coincided with the continuum topological charge $Q$. In a sequel to this paper we determine analytically the classical continuum limit of $Q_{lat}^{(m)}$ for U(1) gauge fields on the 2d-torus for arbitrary even dimension 2d. The answer is found to be $I(m)Q$ where $I(m)$ is a locally constant function with $I(m) = 1$ for $0 < m < d$. E.g., in the 4-dimensional case $I(m)$ is given by the table (1.8) in §1. The non-abelian case is currently under investigation.

It would be interesting to prove directly (i.e. without invoking the continuum index theorem) that the index of the overlap-Dirac operator D reduces to the index of the continuum Dirac operator in the classical continuum limit. (This is easy to verify when the gauge field is trivial; the general case is currently under investigation.) Then the lattice index formula (1.2) together with the results here would imply a new lattice-based proof of the continuum index theorem for the (relatively simple) situation considered in this paper. One could then go on and attempt to develop a new lattice-based proof of the index theorem in more general situations. An obstacle in this context is that closed manifolds do not admit cubic cell decompositions in general. Therefore it would be desirable to find a construction of $Q_{lat}$ and D using a triangulation rather than a cubic lattice (as was done for Lüscher’s topological charge in [20]).

A previous version of this paper (hep-lat/9812003 v1) was circulated and submitted for publication on 14 Nov.’98. Shortly afterwards, papers by K. Fujikawa [21] and H. Suzuki [22] appeared on the hep-th archive where the axial anomaly associated with the overlap-Dirac operator was also considered. The arguments of the latter paper are similar to ours, while the former paper uses a different, more general scheme.
Acknowledgements. I am indebted to a referee of a previous version of this paper for pointing out subtleties concerning topologically non-trivial gauge fields overlooked in the previous version, and for other helpful comments. I thank Prof. W. Kerler for several helpful comments, and Prof.’s T-W. Chui and H. Neuberger for encouraging correspondence. The support of a postdoctoral fellowship from the Australian Research Council is gratefully acknowledged.

A Appendix: Derivation of the bound (3.25).

Recall that \( u \leq L^2 \leq v \) \((u > 0 \text{ and } v < \infty)\) for \( L = L_A \) and \( L = L_{\tilde{A}} \). Following [9], we consider the generating functional

\[
(1 - 2tz + t^2)^{-1/2} = \sum_{k=0}^{\infty} t^k P_k(z)
\]

(A.1)

for the Legendre polynomials \( P_k(z) \). Setting

\[
z = (v + u - 2L^2)/(v - u)
\]

(A.2)

we have \(|z| \leq 1\), which implies \(|P_k(z)| \leq 1\) for all \(k\), hence the expression (A.1) is norm convergent when \(|t| < 1\). Now define \( \theta > 0 \) by

\[
cosh \theta = (v + u)/(v - u)
\]

(A.3)

and set \( t = e^{-\theta} \). Then (A.1) reduces to

\[
G = (L^2)^{-1/2} = \kappa \sum_{k=0}^{\infty} t^k P_k(z) \quad , \quad \kappa = [4t/(v - u)]^{1/2}
\]

(A.4)

This leads to

\[
G_A(x,y) - G_{\tilde{A}}(x,y) = \kappa \sum_{k=0}^{\infty} t^k (G_k^A(x,y) - G_k^{\tilde{A}}(x,y))
\]

(A.5)

where \( G_k = P_k(z) \). The sum is norm convergent \( \forall x, y \) since

\[
a^4|G_k(x,y)| \leq ||G_k|| \leq 1.
\]

(A.6)

Note that \( G_k \) is a polynomial of degree \( k \) in \( L^2 \) and \( L \) is a combination of next to nearest neighbour operators. Recall that \( A = \tilde{A} \text{ outside of } B_2 \) and that \( B_2 \subset B_1 \).
with minimum distance $b > 0$ between the boundaries of $B_1$ and $B_2$ w.r.t. the “taxi driver” norm $| \cdot |_1$ ($|x|_1 = \sum_{\mu} |x_{\mu}|$). It follows that

$$x, y \notin B_1 \implies G^A_k(x, y) = G^\Lambda_k(x, y) \text{ for all } k < 2(b/4a). \quad (A.7)$$

Taking this into account in (A.5) and using (A.6) we now get the postulated bound (3.25):

$$a^4|G_A(x, y) - G^\Lambda_A(x, y)| \leq \kappa \sum_{k \geq b/2a} t^k a^4|G^A_k(x, y) - G^\Lambda_k(x, y)|$$

$$\leq 2\kappa t^{b/2a} \sum_{k=0}^{\infty} t^k = \left( \frac{2\kappa}{1 - e^{-\theta}} \right) e^{-\theta b/2a}.$$ 

Apart from the observation (A.7), all the steps in the argument have been taken from [9].

References


