Introduction to Supersymmetry: Fixing Chapter 1.

Alan Reynolds
September 4, 2017

1 Introduction

The book “Introduction to Supersymmetry” by Harald J. W. Müller-Kirsten and Armin Wiedemann [1] is notable in its inclusion of more detailed calculations than are often found in supersymmetry texts. This is certainly the case in the first chapter, which introduces the various types of spinor that are used both in the standard model and in supersymmetry. Given the number of different possibilities for conventions (metric signature, choice of $\epsilon$ matrices, etc., etc.) and the numerous opportunities for introducing sign errors into calculations, the aim of producing such a text is certainly worthwhile.

Alas, chapter one, at least, suffers from a large number of errors and inconsistencies. These vary from minor, isolated sign errors, through to the deliberate choice of conventions that ultimately turn out to be inconsistent with results that the authors wish to prove. Subsequently, the authors seem contort calculations into (at first glance) reasonable looking forms, obscuring errors, in order to produce the desired outcome. Unfortunately, a reader such as myself who appreciates the details will find these errors and contortions to be extremely off-putting and a source of confusion. The consequence is that fixing chapter one requires considerably more work than merely the correction of a few typographical errors. While I write these notes primarily for my own benefit, I hope that, at some stage, they might be useful to other readers of Introduction to Supersymmetry and might save such readers the considerable effort of fixing the chapter themselves. While I still consider the aim of the book to be a worthy one, a supplementary source that pointed out some of the errors would certainly have been useful to me. Hopefully, this document might be such a source for future students.

This document focuses on the larger issues of chapter one. In particular, it analyses the convention used for index position for the complex conjugate representation of $SL(2,\mathbb{C})$ and suggests an alternative that allows the desired results to be proved without calculational contortions. It also explores some issues regarding the ‘index structure’ of various objects described in the chapter, examines the use of the various objects denoted by $\epsilon$ and provides a valid proof of a proposition whose ‘proof’ in the text seems insufficient. This document does not give a comprehensive list of errata and ignores many typographical errors\(^1\). It does, however spend more time discussing why corrections need to be made that a typical errata list.

\(^1\) I may, however, transcribe all the errata I have noted on my copious collection of paper bookmarks into \LaTeX at some future date.
2 A note on matrices and indices

Before attempting to fix any issues with the book, I would like to describe my attitude to the indexing notation, the summation convention and to the treatment of tensors and tensor-like objects (such as spinors) as matrices.

2.1 Matrix notation

Introduction to Supersymmetry uses matrix notation a great deal. However, it is important, in my opinion, not to think of, for example, $\phi_A$ as being a column vector. Rather, take the more flexible line that $\phi_A$ may be considered as a column if appropriate to the calculation being performed. Of course, if a calculation may more easily be converted to matrix form with $\phi_A$ written as a row, then feel free to do so! Index position — i.e. whether the index is up or down — really has nothing to do with rows or columns.

Matrix notation is really rather limited when compared with index notation, and we don’t need to work with objects as complicated as spinors to see this. Any tensor with more than two indices clearly cannot be written as a matrix. Tensors with exactly two indices come in different forms, depending on the index location — converting all of these to matrices clearly loses information about the index structure. Now consider what happens if we try to consider vectors $a^\mu$ to be column vectors and $a_\mu$ to be row vectors. A linear transformation such as $M^{\mu}_{\nu}$ works fine in matrix notation — $a^{\mu} = M^{\mu}_{\nu} b^\nu$ becomes $a = M b$. But now consider the metric $g_{\mu\nu}$, which is also often written as a matrix. If we write $b_\mu = g_{\mu\alpha} a^\alpha$ in matrix form as $b = g a$, then we get $b$ as a column vector, despite having its index down. Similarly, $a^\mu g^{\mu\nu} b_\nu$ can only be written in matrix form if $a^{\mu}$ is written as a row.

In summary, matrix notation is a convenient way to write down the details of a spinorial expression, but it is limited, loses information regarding index structure and can easily lead to misconceptions if one mistakenly thinks that the vector or matrix written down is the spinor. However, one cannot avoid matrices when reading Introduction to Supersymmetry — indeed, to include all the details, the authors had to write down all the numbers represented by something such as $e^{AB}$, and writing them in some form of grid (i.e. a matrix) is really the only sensible way. But, don’t think that the matrix is the spinor and don’t conflate rows and columns with index location up or down.

2.2 Index notation

Introduction to Supersymmetry attempts, not entirely successfully, to ensure that the index notation works smoothly. This sometimes involves using an ‘identity matrix in disguise’, e.g. $(\sigma^0)_{AB}$, to convert an index from one type to another, e.g. from an upper dotted index to a lower undotted one. Numerically, $(\sigma^0)_{AB}$ is just the iden-
introduction to supersymmetry

3 Complex conjugates and index placement

My first major stumbling block on reading Introduction to Supersymmetry occurred at equation (1.63), i.e.

\[(\psi_A)^* = \overline{\psi}_A \in \tilde{F},\]

on page 34. Three issues with this definition, in increasing order of severity, are examined below.

1. It contradicts the description of the complex conjugate representation that precedes it on page 34, which states that \(\overline{\psi}_A \in \tilde{F}\). This is essentially just an issue with the naming of the spaces \(\tilde{F}\) and \(\tilde{F}^*\). However, there is subsequent confusion between these two spaces in the text, in particular on pages 45–47, where it leads to some unnecessary complications.

2. Combining (1.63) with the definition of the \(\epsilon\) matrices on pages 36–38 and with (1.200) on page 110, i.e.

\[\overline{\psi}_A = (\psi_A^\dagger)^*,\]

leads us to a contradiction — an inconsistency in our definitions. To see this, we consider what (1.63) and the definition of the \(\epsilon\) matrices mean for \(\overline{\psi}_A\). We repeat the definitions of the matrices

\[\text{if necessary, put the summation signs in explicitly.}\]

\[\text{In other words, when } B = 1 \text{ we must have } B = 1.\]

\[\text{If you feel that this is being overly fussy and you are comfortable with the calculation above, despite the issues with index structure, then feel free to skip these sections.}\]
here.

\[ \epsilon^{AB} = e^{AB} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \]

\[ \epsilon_{AB} = e_{AB} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \]

Raising or lowering an index of a spinor involves contracting it with an appropriate \( \epsilon \) matrix, summing over the second index, e.g. \( \psi^A = e^{AB} \phi_B \). So\(^{10}\)

\[ \overline{\psi}_A = \epsilon_{AB} \psi^B = \epsilon_{AB} (\psi_B)^* = e_{AB} e_{BC} (\psi^C)^* = - (\psi^A)^*, \]

contradicting (1.200) due to the overall minus sign.\(^{11}\)

Given this, it appears impossible to correct the mathematical contortions of pages 110–111\(^{12}\) to get the desired relationship between the Dirac spinors and the Weyl spinors. In particular, one cannot get Dirac conjugation to work as desired. Ignoring the index structure and starting with

\[ \Psi_W = \left( \begin{array}{c} \phi_A \\ \overline{\psi}^A \end{array} \right) \]

we get\(^{13}\)

\[ \Psi_W = \psi^*_W \gamma^0_W \\
= \left( \begin{array}{c} \phi_A^* \\ \overline{\psi}^A \end{array} \right) \left( \begin{array}{cc} 0 & \mathbb{I}_{2 \times 2} \\ \mathbb{I}_{2 \times 2} & 0 \end{array} \right) \\
= \left( \begin{array}{c} \overline{\psi}^A \\ \phi_A \end{array} \right) \\
= \left( \begin{array}{c} \psi_A \\ \overline{\psi}^A \end{array} \right), \]

which is not (even numerically) the desired \( \left( \begin{array}{c} \psi^A \\ \overline{\psi}_A \end{array} \right) \).

Note that equation (1.63) does not appear to be a typographical error, but rather a deliberate choice by the authors. This is evidenced by the notation about the index structure that follows and by the careful way in which the index structure is fixed in equations (1.90) and (1.91) on page 42 and later in (1.199) on page 110. The unfortunate result is that any fix will have a number of repercussions later in the book.

So how do we fix this? First we handle the definition of \( \vec{F} \), which by inference will also define \( \vec{F}^* \). Here we will simply choose that a star implies indices up, while no star means indices down\(^{14}\), i.e. \( \psi_A \in F, \psi^A \in F^*, \overline{\psi}_A \in F \) and \( \overline{\psi}^A \in F^* \). Even if we make no other changes, this means that (1.63) becomes

\[ (\psi_A)^* =: \overline{\psi}^A \in F^*. \]

Furthermore, about two thirds of the way down page 58, we note that the Pauli matrices \( \sigma^i \) now map \( F^* \) into \( F \), while the \( \sigma^i \) matrices map \( F \) into \( F^* \).

\(^{10}\) Cue another example of squiffy index dotting and positioning. One can use an ‘identity matrix in disguise’ to fix the index structure, as follows:

\[ \overline{\psi}_A = \epsilon_{AB} \psi^B = \epsilon_{AB} (\sigma^a)^{BC} (\psi^c)^* = = \epsilon_{AB} (\sigma^a)^{BC} e_{CD} (\psi^D)^* = - (\psi^A)^*, \]

where we have used (1.90) and (1.106b) and have obtained the first equation on page 143, but with the unfortunate addition of an overall minus sign.

\(^{11}\) At this point, it would seem that there are multiple options for fixing the problem: changing (1.63) to

\[ (\psi_A)^* =: \overline{\psi}_A \in F; \]

keeping (1.63) as is and changing the \( \epsilon \)-matrices so that

\[ \epsilon^{AB} = e_{AB} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \]

\[ \epsilon_{AB} = e^{AB} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}; \]

or changing (1.200) (and two equations on pages 42–43) to

\[ \psi^A = - \overline{\psi}_A. \]

However, only the first option deals satisfactorily with issue 3.

\(^{12}\) Yes, it took 77 pages of reading (and a few weeks) before I could be sure that it was (1.63) that was wrong. Even then, we could perhaps fix this third issue by using a different representation of the Dirac matrices to get the desired form of the Dirac conjugate.

\(^{13}\) See appendix A for a description of how complex conjugation and transposition affect index structure. In particular, neither operation raises or lowers indices.

\(^{14}\) This is merely a naming convention. This choice is made solely because most of the book seems to follow this approach. There is no mathematical reason why we could not switch \( \vec{F} \) and \( \vec{F}^* \), but this would require more changes to the book. (Of course, one must be consistent, which the book fails to be.)
To fix the calculational issues, we replace (1) with

$$(\psi_A)^* = \bar{\Psi}_A \in \bar{F}. \quad (2)$$

We will consider the subsequent changes required to the text in some detail. In this section we will describe the simple changes that must be made solely due to the change in (1.63). Consideration of two sections of the book that are further complicated by additional errors and misunderstandings, on pages 45–47 and 110–111, is deferred to later sections of this document.

The first discrepancy with this new form of (1.63) appears to be the paragraph at the bottom of page 42 and the top of page 43, which deals with fixing the index structure of (1.63). Our replacement for (1.63), equation (2), also fails to exhibit the same index structure on each side. As in the book, we conjure up a matrix that is numerically the identity matrix, but that has the correct index structure to fix (2). While the book could use matrices that had conveniently already been defined, i.e. $\sigma^0$ and $\bar{\sigma}^0$, here we simply define $\delta_{\dot{A}}^A$ to be numerically equal to the identity matrix. Then

$$\begin{align*}
\bar{\Psi}_A &= \delta_{\dot{A}}^A (\psi_A)^*, \\
\Psi^A &= \delta^A_{\dot{A}} (\psi^A)^*.
\end{align*} \quad (3)$$

Note that if we multiply (4) by $\delta_{\dot{C}}^A$ on the left, then we get

$$\delta_{\dot{A}}^C \Psi^A = \delta_{\dot{A}}^C (\psi^A)^* = (\psi^C)^*,$$

i.e.

$$\psi^A = \delta_{\dot{A}}^A \bar{\Psi}_A, \quad (5)$$

while from (3) we get

$$\psi_A = \delta_{\dot{A}}^A \Psi^A. \quad (6)$$

Equations (3)–(6) replace (1.90), (1.91) and the following two unnumbered equations in the book. Note also that this describes the transition from $F$ to $\bar{F}$, rather than to $\bar{F}^*$. Pages 45–47 then need to be changed. However, since this section needs substantial modification regardless of whether we change (1.63) or not, we postpone most of the discussion of these pages until section 3.1. For now, we merely note that changing (1.63) to (2) clearly necessitates changes to figure 1.2 on page 46.

Much that follows regarding spinor calculations can remain unchanged. On page 110, we must replace (1.190) with (4) and (5), while (1.200) becomes

$$\psi^A = \bar{\Psi}^\dot{A}.$$  

The rest of pages 110–111 require more significant modification, which we consider in section 4.
3.1 “Unconventional” definitions of $M^*$ and $M^{*-1\top}$

Pages 45–47 are confused, leading to some “unconventional” definitions of $M^*$ and $M^{*-1\top}$. These unconventional definitions are entirely unnecessary regardless of whether we fix (1.63), as described above, or not. They arise solely as a result of a confusion between $\hat{F}$ and $\hat{F}^*$. I will clear up the confusion here, using both the approaches discussed above.

Page 45 uses (1.63) to infer the transformation map $M^* : \hat{F} \rightarrow \hat{F}$ on the complex conjugate representation induced from the map $M : F \rightarrow F$ on the self-representation. This is easily adapted to the new version, (2), of (1.63). We get

$$\overline{\Psi}_A = \delta_A^A (\Psi_A)^*$$
$$= \delta_A^A (M_A B \Psi_B)^*$$
$$= \delta_A^A (M_A B)^* \Psi_B^*$$
$$= \delta_A^A (M_A B)^* \delta_B^B \overline{\Psi}_B.$$

We then simply define

$$\langle M^* \rangle_B^A := \delta_A^A (M_A B)^* \delta_B^B.$$

Note that, numerically, this just means that matrix $M^*$ is simply the complex conjugate of $M$.

We can continue to see how $\overline{\Psi}^A$ transforms. We get

$$\overline{\Psi}^A = \epsilon^{AB} \overline{\Psi}_B$$
$$= \epsilon^{AB} \delta_B^B (M_B^C)^* \delta_C^C \overline{\Psi}_C$$
$$= \epsilon^{AB} \delta_B^B (M_B^C)^* \delta_C^C \epsilon_{CD} \overline{\Psi}_D$$
$$= \epsilon^{AB} (M_B^C)^* \epsilon_{CD} \overline{\Psi}_D,$$

where we define

$$\epsilon^{AB} := \epsilon^{AB} \delta_B^B$$

and similarly for $\epsilon_{CD}$. Numerically, we find the same matrix as given at the bottom of page 45.

Performing the matrix multiplication, we find that

$$(\epsilon^{AB} (M_B^C)^* \epsilon_{CD}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} m_{11}^1 & m_{12}^2 \\ m_{21}^1 & m_{22}^2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} m_{22}^1 & -m_{21}^2 \\ -m_{12}^1 & m_{11}^2 \end{pmatrix},$$

which is $M^{*-1\top}$ (according to the standard definition), where we have used the fact that $M \in SL(2, \mathbb{C})$, so has determinant of 1.

If we were to decide to leave (1.63) (pretty much) as is, then we note that the complex conjugate representation would consist of spinors of the form $\overline{\Psi}^A$. On pages 45–47, this is the space referred to as $\hat{F}^*$. So surely the matrix $M^*$ should act on this space, i.e. on what pages 45–47 refer to as $\hat{F}^*$. If this is so, then $M^*$ is defined in the calculations on the first five lines of page 45, i.e.

$$(M^*)_B^A = (\overline{\Psi}^0)^{AA} (M_A B)^* (\sigma^0)_{BB}.$$

Note that, numerically, this just means that matrix $M^*$ is simply the complex conjugate of $M$.

We can continue to see how $\overline{\Psi}^A$ transforms. We get

$$\overline{\Psi}^A = \epsilon^{AB} \overline{\Psi}_B$$
$$= \epsilon^{AB} \delta_B^B (M_B^C)^* \delta_C^C \overline{\Psi}_C$$
$$= \epsilon^{AB} \delta_B^B (M_B^C)^* \delta_C^C \epsilon_{CD} \overline{\Psi}_D$$
$$= \epsilon^{AB} (M_B^C)^* \epsilon_{CD} \overline{\Psi}_D,$$

where we define

$$\epsilon^{AB} := \epsilon^{AB} \delta_B^B$$

and similarly for $\epsilon_{CD}$. Numerically, we find the same matrix as given at the bottom of page 45.

Performing the matrix multiplication, we find that

$$(\epsilon^{AB} (M_B^C)^* \epsilon_{CD}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} m_{11}^1 & m_{12}^2 \\ m_{21}^1 & m_{22}^2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} m_{22}^1 & -m_{21}^2 \\ -m_{12}^1 & m_{11}^2 \end{pmatrix},$$

which is $M^{*-1\top}$ (according to the standard definition), where we have used the fact that $M \in SL(2, \mathbb{C})$, so has determinant of 1.

If we were to decide to leave (1.63) (pretty much) as is, then we note that the complex conjugate representation would consist of spinors of the form $\overline{\Psi}^A$. On pages 45–47, this is the space referred to as $\hat{F}^*$. So surely the matrix $M^*$ should act on this space, i.e. on what pages 45–47 refer to as $\hat{F}^*$. If this is so, then $M^*$ is defined in the calculations on the first five lines of page 45, i.e.

$$(M^*)_B^A = (\overline{\Psi}^0)^{AA} (M_A B)^* (\sigma^0)_{BB}.$$

Note that, numerically, this just means that matrix $M^*$ is simply the complex conjugate of $M$.

We can continue to see how $\overline{\Psi}^A$ transforms. We get

$$\overline{\Psi}^A = \epsilon^{AB} \overline{\Psi}_B$$
$$= \epsilon^{AB} \delta_B^B (M_B^C)^* \delta_C^C \overline{\Psi}_C$$
$$= \epsilon^{AB} \delta_B^B (M_B^C)^* \delta_C^C \epsilon_{CD} \overline{\Psi}_D$$
$$= \epsilon^{AB} (M_B^C)^* \epsilon_{CD} \overline{\Psi}_D,$$

where we define

$$\epsilon^{AB} := \epsilon^{AB} \delta_B^B$$

and similarly for $\epsilon_{CD}$. Numerically, we find the same matrix as given at the bottom of page 45.

Performing the matrix multiplication, we find that

$$(\epsilon^{AB} (M_B^C)^* \epsilon_{CD}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} m_{11}^1 & m_{12}^2 \\ m_{21}^1 & m_{22}^2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} m_{22}^1 & -m_{21}^2 \\ -m_{12}^1 & m_{11}^2 \end{pmatrix},$$

which is $M^{*-1\top}$ (according to the standard definition), where we have used the fact that $M \in SL(2, \mathbb{C})$, so has determinant of 1.

If we were to decide to leave (1.63) (pretty much) as is, then we note that the complex conjugate representation would consist of spinors of the form $\overline{\Psi}^A$. On pages 45–47, this is the space referred to as $\hat{F}^*$. So surely the matrix $M^*$ should act on this space, i.e. on what pages 45–47 refer to as $\hat{F}^*$. If this is so, then $M^*$ is defined in the calculations on the first five lines of page 45, i.e.

$$(M^*)_B^A = (\overline{\Psi}^0)^{AA} (M_A B)^* (\sigma^0)_{BB}.$$

Note that, numerically, this just means that matrix $M^*$ is simply the complex conjugate of $M$.
Numerically, this is just the complex conjugate of $M_A^B$. Then the matrix found at the bottom of page 45, i.e.

$$\epsilon_{AC}(\tau^0)^C_A(M_A^B)^* (\sigma^0)_{BD}\epsilon^{DB},$$

should be numerically the same as $M^{*-1\top}$. This is easily confirmed, i.e.

$$(\epsilon_{AC}(\tau^0)^C_A(M_A^B)^* (\sigma^0)_{BD}\epsilon^{DB}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} m_1^{1*} & m_1^{2*} \\ m_2^{1*} & m_2^{2*} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} m_2^{2*} & -m_1^{2*} \\ -m_1^{2*} & m_1^{1*} \end{pmatrix},$$

as before.

### 4 Dirac conjugates, charge conjugation and index structure

The charge conjugation matrix is defined (numerically at least) as $C = i\gamma^2\gamma^0$ in equation (1.190) on page 102-20, while further details regarding its index structure are given on page 107. Later, however, on pages 110-111, we find some highly dubious mathematics. Upon fixing this, we find that, ignoring problems with index structure, it is more convenient to define $C$ as $-i\gamma^2\gamma^0 = i\gamma^0\gamma^2$. We will justify this change to $C$ in the following. Moreover, we will also revisit the index structure of $C$.

Before examining the details of the issues with the charge conjugation matrix, however, we will need to deal with issues around the calculation of the Dirac conjugate on page 110. In particular, we will find that if we use the version of (1.63) from the book, then we do not get the desired form of the Dirac conjugate. And yet, the book seems to get the desired result on line 15 and in (1.201a), that is

$$\Psi_W = (\psi^A \phi^*_A).\tag{9}$$

How has this been achieved? The error appears to be that, upon taking the transpose of $\Psi_W$, the indices have been raised/lowered.

This is clearly incorrect, since raising and lowering indices involves application of the $\epsilon$-matrices, which changes the numerical value of some of the elements — something taking the transpose should not do. Moreover, this error is repeated on obtaining (1.201b).

If we attempt to correct this calculation without changing (1.63) then (ignoring index structure for now) we get

$$\Psi_W = \psi^+_W\gamma_0^W$$

$$= (\phi^*_A \psi^{A*}) \begin{pmatrix} 0 & 1_{2\times2} \\ 1_{2\times2} & 0 \end{pmatrix}$$

$$= (\psi^{A*} \phi^*_A)$$

$$= (\psi^A \phi^*_A),$$

which is numerically different from (9). In other words, we would require a different matrix for $\gamma^0$ to obtain (9) — we would need to

---

\(^{22}\) It is actually only given in the Dirac representation, i.e.

$$C_D = i\gamma^2\gamma^0 = i\begin{pmatrix} 0 & -\sigma^2 \\ \sigma^2 & 0 \end{pmatrix}.$$

However, the $C = i\gamma^2\gamma^0$ part clearly does no change under similarity transformation.

\(^{21}\) I.e. the $\phi$ and $\psi$ are swapped, with the indices places such that we may consider $\Psi\Psi$ without having problems with the index structure.

\(^{22}\) As pointed out in sidenote -4, this is very bad. Appendix A details how the various matrix operators affect the index structure.
work with a different representation of the Dirac matrices. If, on the
other hand, we change (1.63) to (2), then we find that
\[
\Psi_W = \Psi_W^\dagger \Psi_W^0
\]
\[
= \left( \phi_A^* \quad \bar{\Psi}^\dagger_A \right) \left( \begin{array}{cc} 0 & \mathbb{I}_{2 \times 2} \\ \mathbb{I}_{2 \times 2} & 0 \end{array} \right) \\
= \left( \bar{\Psi}^\dagger_A \quad \phi_A^* \right) \\
= \left( \psi^A \quad \bar{\phi}_A \right),
\]
as desired. Furthermore, we obtain\(^{23}\)
\[
\Psi_W^T = \left( \psi^A \quad \bar{\phi}_A \right).
\]
We may now move on to page 111 and the charge conjugation
matrix. First note that the mathematics on the top of page 111 is
highly dubious. For we see that, since the index structure of the
Pauli matrices is given by \((\sigma^\mu)^{AA}\) and \((\bar{\Psi}^\mu)^{AA}\), we may calculate
\((i\sigma^2\sigma^0)^{AB}\) and \((i\bar{\sigma}^2\bar{\sigma}^0)^{AB}\) directly. We get
\[
(i\sigma^2\sigma^0)^{AB} = i \left( \begin{array}{cc} 0 & -i \\ i & 0 \end{array} \right) \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) = \left( \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right)
\]
which is certainly not \(\delta^B_A\) and
\[
(i\bar{\sigma}^2\bar{\sigma}^0)^{AB} = i \left( \begin{array}{cc} 0 & i \\ -i & 0 \end{array} \right) \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) = \left( \begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array} \right),
\]
which is certainly not \(\delta^A_B\). There is no need to calculate \((i\sigma^2\sigma^0)^{AB}\)
and even if there were, it would (most likely\(^{24}\)) be
\[
(i\sigma^2\sigma^0)^{AB} = \epsilon^{AC} (i\sigma^2\sigma^0)^{CB} = \mathbb{I}_{2 \times 2},
\]
and not as given in the book.
We therefore repeat the calculation correctly, starting with equa-
tion (10). We will perform this calculation numerically, by which I
mean that we will not be too concerned with index structure, which
will get somewhat squiffy. We get
\[
\Psi_W^c = (C_W)_{ab} \left( \Psi_W^a \Psi_W^b \right)
\]
\[
= \left( \begin{array}{cc} (i\sigma^2)^{AB} & 0 \\ 0 & (i\bar{\sigma}^2)^{AB} \end{array} \right) \left( \begin{array}{c} \psi^B \\ \bar{\Phi}_B \end{array} \right).
\]
Now, numerically, \((i\sigma^2)^{AB}\) is just \(-\epsilon_{AB}\), while \((i\bar{\sigma}^2)^{AB}\) is just \(-\epsilon^{AB}\).
We therefore get,
\[
\Psi_W^c = \left( \begin{array}{c} -\psi_A \\ -\bar{\phi}_A \end{array} \right)
\]
which is precisely minus what we would like. However, we now
change the definition of \(C\), (1.190) on page 102, to
\[
C_D = -i\gamma^2 \gamma_D^0 = i\gamma^0 \gamma^2 D = i \left( \begin{array}{cc} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{array} \right).
\]
so that

\[ C_W = i \begin{pmatrix} \sigma^0 \sigma^2 & 0 \\ 0 & \sigma^1 \sigma^2 \end{pmatrix}. \]

We find that \( C \) still satisfies the requirement of (1.189) and still satisfies the properties in (1.191). The only change is that we now get

\[ (\Psi^c_W)_{ab} = \left( \begin{array}{c} \psi_A \\ \Phi_A \end{array} \right), \]

as desired.

There remains the issue of the index structure of \( C \), but before that, we need to consider how the indices (should) work in (1.201a) on page 110, i.e. in the expression:

\[ \Psi = \Psi^\dagger \gamma^0 = \left( \phi_A^* \left( \bar{\Psi}^A \right)^* \right) \left( \begin{array}{cc} 0 & (\sigma^0)_{AB} \\ (\sigma^0)_{AB} & 0 \end{array} \right). \]

Clearly, the index structure is not working as desired here, as we are summing over two upper indices (\( A \)) and two lower undotted indices (\( A \)). We can fix the index structure either by inserting an ‘identity matrix in disguise’ between the \( \Psi^\dagger \) and the \( \gamma^0 \), or we can replace \( \gamma^0 \) with the numerically identical matrix:

\[ \hat{\gamma}^0 = \left( \begin{array}{cc} 0 & \delta^A_B \\ \delta^B_A & 0 \end{array} \right). \]

Then we get

\[ \left( \psi^A \bar{\Phi}_B \right) = \Psi = \Psi^\dagger \gamma^0 = \left( \phi_A^* \left( \bar{\Psi}^A \right)^* \right) \left( \begin{array}{cc} 0 & \delta^A_B \\ \delta^B_A & 0 \end{array} \right), \]

and the index structure works smoothly.

Now the index structure of \( C \) can be inferred from the \( \Psi^c = C \Psi^\dagger \), where, in the Weyl representation,

\[ \Psi^c = \left( \begin{array}{c} \psi_A \\ \Phi_A \end{array} \right) \quad \text{and} \quad \Psi^\dagger = \left( \begin{array}{c} \psi^A \\ \Phi^A \end{array} \right). \]

We get the index structure

\[ C = \left( \begin{array}{cc} ()_{AB} & ()_{A}^B \\ ()_{A}^B & ()_{AB} \end{array} \right), \]

contradicting page 107 — indeed, we find that

\[ C = \left( \begin{array}{cc} \epsilon_{AB} & 0 \\ 0 & \epsilon^{AB} \end{array} \right). \]

Alas, now we find that writing \( C \) in terms of the \( \gamma \)-matrices becomes tricky. The matrix \( i \gamma^0 \gamma^2 \) has the index structure of a \( \gamma \)-matrix, which does not match that of \( C \). Trying \( i \gamma^0 \gamma^2 \) produces a result that has the index structure of \( C^{-1} \), which provides a clue. It turns out that we can set

\[ C = i(\gamma^2)^{-1}(\gamma^0)^{-1} = -i \gamma^2(\gamma^0)^{-1} \]

to obtain the correct index structure.

\footnote{Feel free to ignore this section if the index structure does not worry you.}

\footnote{See appendix A for how complex conjugation and transposition affect index structure. Note that neither complex conjugation nor transposition raise or lower indices.}

\footnote{Something like}

\[ \left( \begin{array}{cc} (\sigma^0)_{AB} & 0 \\ 0 & (\sigma^0)_{AB} \end{array} \right). \]

\footnote{Or we can stop worrying so much about index structure!}

\footnote{Note that we cannot write}

\[ C = i(\gamma^0)^{-1} \gamma^2, \]

as the index structure of the two matrices does not permit us to take this product.

\footnote{I suspect that there is a deeper reason for the appearance of \((\gamma^0)^{-1}\) here.}
5 More epsilons and the Pauli matrices

After changing (1.63), we find that the $\epsilon$-matrices, i.e. $\epsilon^{AB}$, $\epsilon_{AB}$, $\epsilon^{AB}$ and $\epsilon_{AB}$ work well as given. However, there is considerable potential for sign errors with regards to $\epsilon^{\mu\nu\rho\sigma}$ and $\epsilon^{ijk}$. In particular, it is important to know how these objects change when an index is raised or lowered. It is also important to consider what happens to Pauli matrices, $\sigma^\mu$, when the index is lowered to produce $\sigma_\mu$.

Problems associated with these objects, in conjunction with other sign errors, appear to lead to further inconsistencies that need to be fixed.

Some decisions need to be made. We start with the Pauli matrices. On page 53, these are defined with the Lorentz index up, i.e.

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

However, this is immediately followed by a set of relations, (1.104), with the (latin) indices down. These, should, I believe, be replaced by

$$\{\sigma^i, \sigma^j\} = 2\delta^{ij}\mathbb{1}_{2\times2}, \quad i, j = 1, 2, 3,$$

$$[\sigma^i, \sigma^j] = 2i\epsilon^{ijk}\sigma^k, \quad i, j, k = 1, 2, 3,$$

$$\text{Tr}[\sigma^i] = 0, \quad i = 1, 2, 3,$$

$$\frac{1}{2} \text{Tr}[(\sigma^i)^2] = 1, \quad i = 1, 2, 3.$$  

While this fix is simple enough, it does raise two questions. What is $\sigma_\mu$ and what do we mean by $\epsilon^{ijk}$, i.e. how are indices on these objects raised and lowered? Staying with the Pauli matrices for now, we see an example of the index being lowered with $\eta_{\mu\nu}$ towards the bottom of page 55. In other words, $\sigma^\mu$ is simply treated like a vector in this regard. (Indeed, on page 66, we see that $(\phi\sigma^\mu\chi)$ transforms like a vector, so we would expect to use $\eta_{\mu\nu}$ to lower the index.) We therefore use this convention here.

Regarding $\epsilon^{ijk}$, it seems best to consider the index position to be unimportant, so that

$$\epsilon_{ijk} = \epsilon^{ijk} = \epsilon^{ij} = \epsilon_{ijk},$$  

so that it remains the object we are familiar with from basic vector calculus. The summation, in (1.104b), over repeated $k$ indices in the down position would suggest this. Further evidence appears on line 5 of page 69, where it is stated that

$$\sum_{j,k=1}^{3} \epsilon_{ijk} \epsilon^{jkl} = 2\delta_i^l. \quad (11)$$  

Since the book uses a $+--$ signature for the metric, if index position mattered and the metric were used to raise and lower
Hence the bottom line of page 69 should read

\[ \sum_{j,k=1}^{3} \epsilon_{jk3} \epsilon_{jk3} = -\sum_{j,k=1}^{3} \epsilon_{jk3} \epsilon_{jk3} = -\epsilon_{123} \epsilon_{123} - \epsilon_{213} \epsilon_{213} = -2, \]

contradicting (11).

We finally consider \( e^{\mu\nu\rho\sigma} \). This is actually clearly defined on page 26, where it is stated that \( \epsilon_{0123} = +1 \) and it can be inferred that index position does matter, so that \( \epsilon_{0123} = -1 \). However, we find that we then have difficulties on the bottom of page 69 and on page 70, where the calculations have been done, as given, turn out to be wrong. On the second last line we have \( e_{ijkl} \epsilon_{jkl} \). Now since \( \epsilon_{\mu\nu\rho\sigma} \) has been defined so that \( \epsilon_{0123} = +1 \), we must have \( e_{ijkl} = e_{ijk} \) and hence \( e_{ijkl} = -e_{ijk} \). Then

\[ e_{ijkl} \epsilon_{jkl} = -e_{ijk} \epsilon_{jkl} = -e_{jk1} \epsilon_{jkl} = -2\delta_i^1. \]

Hence the bottom line of page 69 should read

\[ = -\frac{1}{4i} 2\delta_i^1 \sigma^l = -\frac{1}{2i} \sigma^l = -e^{0i}. \]

Sign errors continue on page 70. Line 3 should read

\[ = -i \frac{1}{4} [\sigma^i, \sigma^j] = -i \frac{1}{4} (2i e_{ij}^k \sigma^k) = -i \frac{1}{4} (2i e_{ij}^k \sigma^k) = -\frac{1}{2} e^{ij} \sigma^k, \]

so we find, when we get to line 6, that \( e^{ij} \sigma^k \)

\[ = \frac{1}{2i} e^{ij} \sigma^k \sigma^l = \frac{1}{2} e^{ij} \sigma^k \sigma^l = \frac{1}{2} e^{ij} \sigma^k = -\sigma^{ij}. \]

The same two errors are repeated for part (ii), with sign errors introduced in going from line 13 to line 14 and on line 16.

As a result of this, with \( e^{\mu\nu\rho\sigma} \) as defined in the book, we get sign errors in (1.140), i.e. in the statement of proposition 1.33. Now, it turns out that the best way to fix this may be simply to change the sign of \( e^{\mu\nu\rho\sigma} \) on page 26 so that \( e^{\mu\nu\rho\sigma} = +1 \). Indeed, the only place in section 1.2 of the book where the sign seems to matter is on page 31, part (b), and, moreover, changing the sign of \( e^{\mu\nu\rho\sigma} \) appears to fix a sign error in (1.48). To see this sign error, note that if we stick with \( \epsilon_{0123} = +1 \), then we find that

\[ \frac{1}{2} e^{0ijk} M_{jk} = -\frac{1}{2} e_{0ijk} M_{jk} = -\frac{1}{2} e_{ijk} M_{jk} = -I_i = f^i, \]

so that

\[ \frac{1}{2} e^{0ijk} P_i M_{jk} = P_i f^i = -P J. \]

Changing the sign of \( e^{\mu\nu\rho\sigma} \) fixes this error too.

Alas, all is not plain sailing. In proposition 1.38 on page 76, we find that the calculations are correct for \( \epsilon_{0123} = +1 \). Changing \( e^{\mu\nu\rho\sigma} \) results in a change in sign of

\[ \frac{1}{2} \text{Tr} [e_{\mu\nu\rho\sigma} \sigma^\mu \sigma^\nu \sigma^\rho \sigma^\sigma], \]

which is now \(-24i\). As a result, \( c \) in proposition 1.38 becomes \( i \) and the sign before \( e^{\mu\nu\rho\sigma} \) in equations (1.144) to (1.146) must be changed.

\[ \text{These calculations were always going to be fiddly, as they mix the three dimensional \( \epsilon_{ijk} \) with the four dimensional \( e^{\mu\nu\rho\sigma} \). One therefore has to be particularly careful with the signs.} \]

\[ \text{Note that } e^{ij0} = -e^{0ij} = e^{ij}. \]

\[ \text{Another plus point for changing } e^{\mu\nu\rho\sigma} \text{ so that } \epsilon_{0123} = +1 \text{ is that this is similar to what we have for the } e \text{-matrices, i.e. } e^{12} = e^{12} = +1. \]
6 A dodgy proof

The proof on page 61 is suspect. The key lies in the footnote: "We introduce a unitary matrix \( U \) which diagonalizes \( M \)." But a matrix is only unitarily diagonalizable if it is a normal matrix
\[
\text{I.e. if } M^*M = MM^*. \text{ A Hermitian matrix is clearly normal.}
\]
\[
\text{A three step replacement proof is as follows.}
\]

1. \( MM^\dagger \) is Hermitian, since \((MM^\dagger)^\dagger = (M^\dagger)^\dagger M^\dagger = MM^\dagger\). Hence it is unitarily diagonalizable, i.e. we can find \( U \in U(2) \) such that \( UMM^\dagger U^\dagger = D \) is diagonal.

2. Diagonal elements of matrix \( AA^\dagger \) are positive real, since
\[
(AA^\dagger)_{ii} = \sum_j A_{ij}(A^\dagger)_{ji} = \sum_j A_{ij}A^*_{ij} = \sum_j |A_{ij}|^2.
\]

Hence \( D \) is positive real and the eigenvalues of \( MM^\dagger \) are positive real.

3. Since both the trace and the determinant are preserved by similarity transforms, we get
\[
\begin{align*}
\text{Tr}(MM^\dagger) &= \text{Tr} D = \lambda_1 + \lambda_2, \\
\text{det}(MM^\dagger) &= \text{det} D = \lambda_1 \lambda_2 = 1,
\end{align*}
\]
where \( \lambda_1 \) and \( \lambda_2 \) are the eigenvalues. From this we get
\[
\text{Tr}(MM^\dagger) = \lambda_1 + \frac{1}{\lambda_1},
\]
which is minimal when \( 1 - 1/\lambda_1^2 = 0 \), i.e. when \( \lambda_1 = 1 \). The minimal value is therefore \( 1 + 1/1 = 2 \). Hence
\[
\text{Tr}(MM^\dagger) \geq 2,
\]
which implies that \( \Lambda_0^0(M) \geq 1 \), as required.

A Index structure of complex conjugate, transposed and inverse matrices

I have already noted that expressions such as \((1.63)\) in the book, if used in calculations, lead to problems with the index structure of the expressions being manipulated, but that one can usually just continue calculating, provided one keeps track of which indices are summed over. Endeavouring to keep the index structure straight at all times requires the introduction of the ‘identity matrix in disguise’ and results in more complex and less memorable equations. However, if one does endeavour to keep the indices straight, and if one also wishes to use matrix notation, it is important to know how the index structure changes under complex conjugation, transposition and inversion of the matrices.
This is all straightforward, but it is easy to become confused, especially when combined with the addition complications of spinors. We start with complex conjugation. This is simply complex conjugation of all of the matrix elements, so this makes no change to the index structure whatsoever, i.e.

\[
(M^*)^{AB} = (M^{AB})^*,
\]
\[
(M^*)_{AB} = (M_{AB})^*,
\]
\[
(M^*)^{A}_B = (M^{A}_B)^*,
\]
\[
(M^*)^B_A = (M^B_A)^*,
\]
even.

The alert reader will note that this seems to contradict equations (7) and (8) above. Well, we should perhaps replace (7) and (8) with something like

\[
\overline{M}^B_A := \delta^A_A (M^B_A)^* \delta^B_B = \delta^A_A (M^*)^B_A \delta^B_B,
\]

and

\[
\overline{M}^A_B = (\sigma^0)^{AA} (M^B_A)^* (\sigma^0)^{BB} = (\sigma^0)^{AA} (M^*)^B_A (\sigma^0)^{BB},
\]

since to find the transformation matrix for the complex conjugate representation that corresponds to \(M^B_A\) in the self-representation, we must not only take the complex conjugate, but we must add dots to the indices to match the dots that we conventionally add to the indices of the right-handed spinors.

We can check that this works correctly by taking the complex conjugate of an expression such as

\[
\psi'_B = M^B_A \psi_B.
\]

We will use equation (12) to relate \(M\) and \(\overline{M}\), from which we obtain

\[
(M^*)^B_A = \delta^A_A \overline{M}^B_A \delta^B_B.
\]

So now

\[
\psi'^*_A = (M^B_A)^* \psi'^*_B
\]
\[
= (M^*)^B_A \psi'^*_B
\]
\[
= \delta^A_A \overline{M}^B_A \delta^B_B \overline{\psi}_C
\]
\[
= \delta^A_A \overline{M}^B_A \overline{\psi}_B
\]

We then multiply both sides on the left by \(\delta^C_A\) to get

\[
\overline{\psi}_A = \overline{M}^B_A \overline{\psi}_B,
\]
as desired, without any problems arising from the index notation.

Next we consider taking the transpose. This merely switches the order of the indices, without affecting their height. After all, raising or lowering indices requires the use of the \(\epsilon\)-matrices, which change
the values of some of the numbers — something transposition
certainly should not do. Hence

\[(M^T)^{AB} = M^{BA},\]
\[(M^T)_{AB} = M_{BA},\]
\[(M^T)^A_B = M_B^A,\]
\[(M^T)^{\dot{A}}_{\dot{B}} = M^{\dot{B}}_{\dot{A}} , \text{ etc.}\]

Finally, we consider the inverse matrix. Here, we merely note
that both \(MM^{-1}\) and \(M^{-1}M\) should work with the usual index
structure rules. Hence, if the second index of \(M\) is down and un-
dotted, then the first index of \(M^{-1}\) must be up and undotted for
\(MM^{-1}\) to work. Similarly, if the first index of \(M\) is up and dotted,
then the second index of \(M^{-1}\) must be down and dotted for \(M^{-1}M\)
to work. Hence

\[M^{AB} \rightarrow (M^{-1})_{AB},\]
\[M_{AB} \rightarrow (M^{-1})^{AB},\]
\[M^A_B \rightarrow (M^{-1})^A_B ,\]
\[M^{\dot{A}}_\dot{B} = (M^{-1})^{\dot{B}}_{\dot{A}} , \text{ etc.}\]

References