Perturbative QCD and NLO
Monte Carlo Simulations

TASI 2011

John Campbell
MS 106, PO Box 500
Fermilab, Batavia IL 60510

June 10, 2011

Abstract

This is a set of notes for lectures which I am giving at TASI2011, provided for the
benefit of the students there. The notes provide an introduction to next-to-leading order
calculations in QCD and develop the spinor calculus necessary for the computation of
both tree-level and one-loop amplitudes (using analytic unitarity techniques). A large
part of these notes is based on lectures previously delivered by Keith Ellis in Würzburg
and Calcutta. If the notes are prepared for wider dissemination then the incomplete
referencing to the literature will be rectified.

Lecture 1
1 Outline

The outline of the lectures is as follows.

- Motivation and discussion of NLO.
- Modern tools for QCD calculations: introduction to the spinor formalism and useful identities.
- Simple tree level calculations in the spinor formalism.
- A one-loop calculation using analytic unitarity.

2 Motivation

In some situations a leading order pQCD estimate of a cross section is good enough. Canonical examples are the search for a narrow resonance (little theory input required) and searches involving significant changes in the shapes of distributions between backgrounds and signal (LO is probably good enough). However, even in these cases, an accurate knowledge of cross sections is still eventually required. If a narrow resonance is discovered, extraction of parameters such as its mass and width are most accurately performed with a precision calculation of the signal. In the case of shape changes, certain distributions may be sensitive to higher order corrections to backgrounds that can fake a signal. Moreover, in the case of very small signals hiding behind large or steeply-falling backgrounds, an accurate knowledge of both normalizations and backgrounds is required to extract a signal.

In this section we discuss broad features of NLO calculations and describe how they arise in actual relevant hadron collider processes.

2.1 Anatomy of a calculation

Consider one of the simplest possible hadron-collider processes, \(W\) production (Drell-Yan process). At NLO in the strong coupling, we must compute the additional diagrams shown in Fig. 2.1. Diagrams (a) represent real radiation of a gluon and are “squared-up” and thus proportional to \(g_s^2\). The 1-loop diagrams in (b) are themselves of order \(g_s^2\) and are interfered with the LO (tree) diagrams. Each of these contributions is separately divergent although, for physically observable quantities, the sum is not (KLN theorem).

In order to compute sensibly we just need a regularization scheme, usually some variant of dimensional regularization, \(d \rightarrow 4 - 2\epsilon\).

This can be illustrated rather simply with a toy model \[1\]. In the toy model we have a one-dimensional integral that corresponds to the extra phase space that must be integrated out in the real emission diagrams. In particular the integration variable corresponds to the energy of the gluon, or the angle at which it is emitted with respect to one of the quarks. In both those limits the real radiation matrix elements are divergent. Hence we begin with the integral,

\[
\mathcal{I} = \int_0^1 \frac{dx}{x} x^{-\epsilon} \mathcal{M}(x) .
\]  

(2.1)
The factor of $x^{-\epsilon}$ ($\epsilon < 0$) regularizes the divergence as $x \to 0$. $\mathcal{M}(x)$ corresponds to the real radiation matrix elements, with $\mathcal{M}(0)$ the tree level matrix elements. This factorization, into a tree-level matrix element multiplied by a singular function, is universal and essential for this procedure.

A particular method for isolating the singularities (phase-space slicing) proceeds as follows. First, introduce a parameter $\delta$ to parametrize the extent of the singular region. Within this region we use the factorized form of the real matrix elements, $\mathcal{M}(x) \approx \mathcal{M}(0)$ for $x < \delta \ll 1$. Hence we can write,

$$I = \mathcal{M}(0) \int_0^\delta \frac{dx}{x} x^{-\epsilon} + \int_\delta^1 \frac{dx}{x} x^{-\epsilon} \mathcal{M}(x)$$

$$= -\frac{1}{\epsilon} \delta^{-\epsilon} \mathcal{M}(0) + \int_\delta^1 \frac{dx}{x} \mathcal{M}(x) , \quad (2.2)$$

where in the last line we have performed the first integral and set $\epsilon = 0$ in the second since it no longer requires regularization. Expanding for small $\epsilon$ we can write the result as,

$$I = \left(-\frac{1}{\epsilon} + \log \delta\right) \mathcal{M}(0) + \int_\delta^1 \frac{dx}{x} \mathcal{M}(x) , \quad (2.3)$$

This procedure has isolated the singularity as a pole in $\epsilon$ multiplied by the tree-level matrix elements. This will cancel against the same term appearing in the 1-loop calculation (with the opposite sign). The remaining terms are finite and can be simply integrated numerically. We observe that this procedure has an additional feature – our answer should be independent of $\delta$ but there is a tension between retaining a good singular approximation (very small $\delta$) and avoiding large logarithmic cancellations (prefer larger $\delta$). An illustration of this check is shown in Fig. 2.2, taken from a calculation of $Wbb$ production at NLO [2].
The calculation of the real corrections is by now a largely automatic affair. Tree-level matrix elements can be generated using one of the many readily-available codes for doing so and the structure of the singularities is well known. Codes for setting up appropriate phase space slicing [3] or dipole subtraction [4] are also available.

Computation of the virtual corrections is more subtle but methods for automating this step are also being developed. Impressive results for $2 \rightarrow 6$ processes have recently been obtained [5, 6] while more general approaches to the problem are just beginning to appear [7]. These methods are based on numerical techniques that we will not attempt to describe here. However we will return to the issue of virtual corrections later in these lectures.

2.2 Scale dependence

An oft-cited reason for computing observables to higher order is the need for improved accuracy. This improvement does not simply rely on a successively-better approximation at each higher order in perturbation theory. Instead the improvement stems from the reduced dependence on the artificial renormalization and factorization scales that we must introduce in order to work at a finite order of the theory.

We will follow a simple formulation of the issue given in ref. [8]. Consider the single-jet inclusive distribution at the Tevatron, calculated from diagrams such as the ones shown in Fig. 2.3. At high transverse momentum (and therefore large parton $x$) it is dominated by these diagrams, from the quark-antiquark initial state.

For this process, we can write the lowest order prediction for the single jet inclusive
Figure 2.3: The leading order diagrams representing inclusive jet production from a quark-antiquark initial state.

dσ
dE_T = α_s^2(μ_R)σ_0 ⊗ f_q(μ_F) ⊗ f_\bar{q}(μ_F), \quad (2.4)

where σ_0 represents the lowest order partonic cross section calculated from the diagrams of Figure 2.3 and f_i(μ_F) is the parton distribution function for a parton i evaluated at the factorization scale μ_F. Similarly, after including the next-to-leading order corrections, the prediction can be written as,

dσ
dE_T = \left[ α_s^2(μ_R)σ_0 + α_s^3(μ_R) \left( σ_1 + 2b_0 \log(μ_R/E_T)σ_0 - 2P_{qq} \log(μ_F/E_T)σ_0 \right) \right] ⊗ f_q(μ_F) ⊗ f_\bar{q}(μ_F). \quad (2.5)

In this expression the logarithms that explicitly involve the renormalization (μ_R) and factorization scales have been exposed. The remainder of the O(α_s^3) corrections lie in the function σ_1.

From this expression, the sensitivity of the distribution to the renormalization scale is easily calculated using,

\frac{∂α_S(μ_R)}{∂(log μ_R)} = -b_0α_s^2(μ_R) - b_1α_s^3(μ_R) + O(α_s^4), \quad (2.6)

where the two leading coefficients in the beta-function, b_0 and b_1, are given by b_0 = (33 - 2n_f)/6π, b_1 = (102 - 38n_f/3)/8π^2. The contributions from the first and third terms in Eq. (2.5) cancel and the result vanishes, up to O(α_s^4).

In a similar fashion, the factorization scale dependence can be calculated using the non-singlet DGLAP equation,

\frac{∂f_i(μ_F)}{∂(log μ_F)} = α_S(μ_F)P_{qq} ⊗ f_i(μ_F). \quad (2.7)

This time, the partial derivative of each parton distribution function, multiplied by the first term in Eq. (2.5), cancels with the final term. Thus, once again, the only remaining terms are of order α_s^4. This is a generic feature of a next-to-leading order calculation. An observable that is predicted to order α_s^2 is independent of the choice of either renormalization or factorization scale, up to the next higher order in α_s.
Figure 2.4: The single jet inclusive distribution at $E_T = 100$ GeV, appropriate for Run I of the Tevatron. Theoretical predictions are shown at LO (dotted magenta), NLO (dashed blue) and NNLO (red). Since the full NNLO calculation is not complete, three plausible possibilities are shown.

Of course this is only a formal statement and the numerical importance of the higher order terms may be large. In order to make the discussion more concrete we will insert numerical values into the formulae indicated above. For simplicity, we will consider only the renormalization scale dependence, with the factorization scale held fixed at $\mu_F = E_T$.

In this case it is simple to extend Eq. (2.5) one order higher in $\alpha_S$,

$$\frac{d\sigma}{dE_T} = \left[ \alpha_S^2(\mu_R) \sigma_0 + \alpha_S^3(\mu_R) \left( \sigma_1 + 2b_0 L \sigma_0 \right) \right. \\
+ \left. \alpha_S^4(\mu_R) \left( \sigma_2 + 3b_0 L \sigma_1 + (3b_0^2 L^2 + 2b_1 L) \sigma_0 \right) \right] \otimes f_q(\mu_F) \otimes f_\bar{q}(\mu_F),$$

(2.8)

where the logarithm is abbreviated as $L \equiv \log(\mu_R/E_T)$. For a realistic example at the Tevatron Run I, $\sigma_0 = 24.4$ and $\sigma_1 = 101.5$. With these values the LO and NLO scale dependence can be calculated; the result is shown in Figure 2.4 adapted from [8]. At the moment the value of $\sigma_2$ is unknown but a range of predictions based on plausible values that it could take are also shown in the figure, $\sigma_2 = 0$ (solid) and $\sigma_2 = \pm \sigma_1^2/\sigma_0$ (dashed). It is clear that the renormalization scale dependence is reduced when going from LO and NLO and will become smaller still at NNLO.

Although Figure 2.4 is representative of the situation found at NLO, the exact details depend upon the kinematics of the process under study and on choices such as the running of $\alpha_S$ and the pdfs used. Of particular interest are the positions on the NLO curve which correspond to often-used scale choices. Due to the structure of Eq. (2.5) there will normally be a peak in the NLO curve, around which the scale dependence is minimized. The scale at which this peak occurs is often favoured as a choice. For example, for
inclusive jet production at the Tevatron, a scale of $E_{\text{jet}}^T/2$ is usually chosen. This is near the peak of the NLO cross section for many kinematic regions. It is also usually near the scale at which the LO and NLO curves cross, i.e. when the NLO corrections do not change the LO cross section considerably. Finally, a rather different motivation comes from the consideration of a “physical” scale for the process. For instance, $E_{\text{jet}}^T$ for the inclusive jet process or the $W$ mass in the case of $W$ production. Clearly, these three typical methods for choosing the scale at which cross sections should be calculated do not in general agree. If they do, one may view it as a sign that the perturbative expansion is well-behaved. If they do not agree then the range of predictions provided by the different choices can be ascribed to the “theoretical error” on the calculation.

A word of caution is due at this point. Although the improved scale dependence sketched out here is typical, it is by no means guaranteed. This is particularly true for cross sections at the LHC that proceed at LO through quark-antiquark initial states. Since the LHC produces an abundance of gluons, real radiation diagrams containing a gluon in the initial state can give rise to very large NLO corrections. Since they enter for the first time at NLO, they are strictly only accurate to LO and therefore give rise to a sizeable scale dependence. A well-known example is shown in Fig. 2.5 for the case of $Wb\bar{b}$ production at the 14 TeV LHC. In the absence of gluons the NLO calculation has the canonical behaviour; in their presence the rate is not well-controlled due to diagrams such as the one shown.

### 2.3 Other considerations

Another caveat arises when considering differential distributions at NLO. It is often the case that, either due to the limitations of the calculation or because of specific cuts that are applied, some distributions have a kinematic limit at LO. Often these limits are softened, or extended, at NLO due to the different kinematics in the real corrections. An example of such a situation is shown in Fig. 2.6 which depicts the transverse momentum distribution of the Higgs boson at the 7 TeV LHC. This is computed from the LO process.
Figure 2.6: The transverse momentum of a 200 GeV Higgs boson produced at the 7 TeV LHC, computed at LO and NLO in QCD.

$pp \to H + 1 \text{ jet}$ with a cut on the jet $p_T$ at 25 GeV that at LO translates into a cut on the Higgs $p_T$. At NLO the Higgs $p_T$ can be balanced by two partons that have a combined transverse momentum below the cut, so the region below 25 GeV is populated. This has two consequences. Since the region below 25 GeV originates solely from real radiation events, it should only be trusted as much as a LO calculation. In addition, the region around the kinematic boundary at 25 GeV is not well-described. The values of the histogram bins there are simply artefacts of the calculation and are not reliable.

Smoothing out these sorts of problems, and providing hadron-level predictions that can be directly compared with experimental results, is the domain of the parton shower and is beyond the scope of these lectures. There has been much progress in this area recently, with regard to improving the accuracy of the shower both in hard regions and with respect to including higher-order corrections. The reader is referred to refs. [9, 10, 11, 12, 13] for details.

### 3 The spinor method

In this section we introduce the spinor method and derive some identities that will be very useful in performing actual amplitude calculations.

#### 3.1 The Dirac equation for massless fermions

We begin by considering solutions of the Dirac equation for massless fermions. We recall the usual definitions of the gamma matrices,

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^\mu\nu,$$

$$\gamma_5 = i\gamma^0 \gamma^1 \gamma^2 \gamma^3.$$
To proceed, we must choose an explicit representation for the gamma matrices. The Bjorken and Drell representation is,

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (3.3)$$

However, it is more useful for us to use the Weyl representation, that is more suitable at high energy,

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.4)$$

We see that for both these representations the gamma matrices satisfy the relations,

$$(\gamma^\mu)^\dagger = \gamma^0 \gamma^\mu \gamma^0,$$  \quad (3.5)

$$(\gamma^\mu)^\ast = \gamma^2 \gamma^\mu \gamma^2.$$  \quad (3.6)

In a high energy processes we can often take the fermions to be massless, so that the Dirac equation they must satisfy is simply $i \slashed{p} u(p) = 0$. In order to derive an explicit solution to this equation it is useful to write out the expression for $i \slashed{p} = \gamma^0 p^0 - \gamma^1 p^1 - \gamma^2 p^2 - \gamma^3 p^3$ in the Weyl representation,

$$i \slashed{p} = \begin{pmatrix} 0 & 0 & p^+ & p^1 - ip^2 \\ 0 & 0 & p^1 + ip^2 & p^- \\ -p^1 + ip^2 & 0 & 0 \\ -p^1 - ip^2 & 0 & 0 \end{pmatrix}, \quad (3.7)$$

where $p^\pm = p^0 \pm p^3$. By inspection the solution to the massless Dirac equation is then,

$$u(p) = \begin{bmatrix} \sqrt{p^+} e^{i\varphi_p} \\ \sqrt{p^-} e^{-i\varphi_p} \\ -\sqrt{p^+} \\ -\sqrt{p^-} \end{bmatrix},$$  \quad (3.8)

where the phase factor is given by,

$$e^{\pm i\varphi_p} = \frac{p^1 \pm ip^2}{\sqrt{(p^1)^2 + (p^2)^2}} = \frac{p^1 \pm ip^2}{\sqrt{p^+ p^-}}, \quad p^\pm = p^0 \pm p^3. \quad (3.9)$$

It is convenient to decompose this solution into different helicity states using the projection operators (in the Weyl representation),

$$\gamma_R = \frac{1}{2} (1 + \gamma_5) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \gamma_L = \frac{1}{2} (1 - \gamma_5) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \quad (3.10)$$

Our original solution can then be written as the sum of the two helicity solutions,

$$u_+(p) = \gamma_R u(p) = \begin{bmatrix} \sqrt{p^+} e^{i\varphi_p} \\ \sqrt{p^-} e^{-i\varphi_p} \\ 0 \\ 0 \end{bmatrix}, \quad u_-(p) = \gamma_L u(p) = \begin{bmatrix} 0 \\ 0 \\ \sqrt{p^-} e^{-i\varphi_p} \\ -\sqrt{p^+} \end{bmatrix}. \quad (3.11)$$
In this representation the Dirac conjugate spinors are,
\[ \overline{\psi}_+(p) \equiv u_+^\dagger(p)\gamma^0 = \left[ 0,0,\sqrt{p^+},\sqrt{-p^+} e^{-i\varphi_p} \right], \]
(3.12)
\[ \overline{\psi}_-(p) = \left[ \sqrt{-p^+} e^{i\varphi_p},0,0,\sqrt{p^+} \right]. \]
(3.13)
The normalization of the spinors is such that,
\[ u_+^\dagger u_+ = 2p^0. \]
(3.14)
The two spinor solutions are further related by,
\[ u_+(p) = \left( \begin{array}{cccc} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{array} \right) \left[ \begin{array}{c} 0 \\ \sqrt{p^-} e^{i\varphi_p} \\ -\sqrt{p^+} \\ 0 \end{array} \right] = i\gamma^2 u_+^\dagger(p) = i\gamma^2 \gamma^0 [\overline{\psi}_-(p)]^T, \]
(3.15)
\[ \overline{\psi}_+(p) = i [u_-(p)]^T \gamma^2 \gamma^0. \]
(3.16)
These relations express the charge conjugation properties of massless spinors.

### 3.2 Notation

Consider a set of (massless) momenta \( p_i, i = 1,2,\ldots,n \). We introduce a bra and ket notation with the spinor labelled by the index \( i \) corresponding to momentum \( p_i \),
\[ |i^\pm \rangle \equiv |p_i^\pm \rangle \equiv u_\pm(p_i) = v_\mp(p_i), \]
(3.17)
\[ \langle i^\pm | \equiv \langle p_i^\pm | \equiv \overline{u}_\pm(p_i) = \overline{v}_\mp(p_i). \]
(3.18)
Due to these relations between the spinors for antiparticles (\( v \)) and particles (\( u \)), we will not use \( v \) henceforth.

The basic spinor products are defined by,
\[ \langle ij \rangle \equiv \langle i^- | j^+ \rangle = \overline{u}_-(p_i) u_+(p_j), \quad [ij] \equiv \langle i^+ | j^- \rangle = \overline{u}_+(p_i) u_-(p_j). \]
(3.19)
The helicity projection implies that products like \( \langle i^+ | j^+ \rangle \) vanish, as can be seen explicitly from Eqs. (3.11) and (3.12). Thus we have,
\[ \langle i^+ | j^+ \rangle = \langle i^- | j^- \rangle = \langle ii \rangle = [ii] = 0. \]
(3.20)

We can use our forms for the spinors to get explicit formulae for the spinor products valid for the case when both energies are positive, \( p_i^0 > 0, p_j^0 > 0 \),
\[ \langle i,j \rangle = \sqrt{p_i^+ p_j^+} e^{i\varphi_{ij}} - \sqrt{p_i^- p_j^-} e^{-i\varphi_{ij}} = \sqrt{|s_{ij}|} e^{i\phi_{ij}}, \]
\[ [i,j] = \sqrt{p_i^+ p_j^-} e^{-i\varphi_{ij}} - \sqrt{p_i^- p_j^+} e^{i\varphi_{ij}} = -\sqrt{|s_{ij}|} e^{-i\phi_{ij}} \]
(3.21)
where \( s_{ij} = (p_i + p_j)^2 = 2p_i \cdot p_j \) and,
\[ \cos \phi_{ij} = \frac{p_i^+ p_j^+ - p_j^+ p_i^+}{\sqrt{|s_{ij}|} p_i^+ p_j^+}, \quad \sin \phi_{ij} = \frac{p_i^+ p_j^- - p_j^+ p_i^-}{\sqrt{|s_{ij}|} p_i^+ p_j^+}. \]
(3.22)
Hence the spinor products are, up to a phase, square roots of Lorentz products. We also see the relations,
\[ \langle ij \rangle = -\langle ji \rangle, \quad [ij] = -[ji] \] (3.23)

For real momenta (most cases we are usually interested in) the two types of spinor product are related by complex conjugation,
\[ \langle ij \rangle^* = [ji] \] (3.24)

Note, however, that for complex momenta this is no longer true.

We can make a slight extension of this notation as follows. The helicity labels on the spinors can always be suppressed in favour of angle or square brackets, as in the spinor products. So we have,
\[ \langle i^- \rangle \rightarrow \langle i \rangle, \quad \langle i^+ \rangle \rightarrow [i], \quad |i^\pm \rangle \rightarrow |i \rangle, \quad |i^- \rangle \rightarrow |i \rangle. \] (3.25)

This notation is more common in the recent literature and we shall adopt it here.

It is appropriate here to comment on an alternative notation, due to Weyl and van der Waerden [14, 15]. Since the left-handed and right-handed spinors occupy different subspaces, we can write them in terms of two-index spinors (also called holomorphic and anti-holomorphic spinors),
\[ |j^+ \rangle = (\lambda_j)_{\alpha}, \quad |j^- \rangle = (\tilde{\lambda}_j)_{\dot{\alpha}}. \] (3.26)

The conjugate spinors can then be written as,
\[ \langle j^- \rangle = \varepsilon^{\alpha\beta} (\lambda_j)_\beta, \quad \langle j^+ \rangle = -\varepsilon^{\dot{\alpha}\dot{\beta}} (\tilde{\lambda}_j)_{\dot{\beta}}. \] (3.27)

where \( \varepsilon^{\alpha\beta} \) is the antisymmetric tensor in two dimensions. We thus see that the spinor products are given by,
\[ \langle j \, k \rangle = \varepsilon^{\alpha\beta} (\lambda_j)_\alpha (\lambda_k)_\beta \]
\[ [j \, k] = \varepsilon^{\dot{\alpha}\dot{\beta}} (\tilde{\lambda}_j)_{\dot{\alpha}} (\tilde{\lambda}_k)_{\dot{\beta}} \] (3.28)

The two different types of indices, dotted and undotted, both run from 1 to 2. Dotted indices are only contracted with other dotted indices, and undotted indices are only contracted with other undotted indices. We will not use this notation any further here.

### 3.3 Useful identities

The collinear limits of massless gauge amplitudes have square-root singularities, factors that are easily expressed in terms of spinor products. Therefore their use can lead to very compact analytic representations of gauge amplitudes. Computations involving spinor products are greatly simplified by three essential identities that we will prove here.
3.3.1 Reversal

Consider the spinor string \([a|\gamma^\mu|b]\), a quantity that naturally arises as the current for emission of a vector boson from a (right-handed) massless fermion line. We can manipulate it as follows:

\[
[a|\gamma^\mu|b] = \langle b|\gamma^\mu|a \rangle
\]

(3.29)

Hence we have the reversal relation:

\[
[a|\gamma^\mu|b] = \langle b|\gamma^\mu|a \rangle
\]

(3.30)

Note that this derivation relies on the charge conjugation properties of massless spinors.

3.3.2 Schouten identity

By explicit construction it is easy to show that,

\[
|b\rangle\langle c| - |c\rangle\langle b| = \langle c b | \gamma_R .
\]

(3.31)

By multiplying by \(\langle a|\) and \(|d\rangle\) on the left and right respectively we obtain the relation,

\[
\langle a b |\langle c d| - \langle a c|\langle b d| = \langle c b |\langle a d| .
\]

(3.32)

Thus we have the Schouten identity:

\[
\langle a b |\langle c d| = \langle a c|\langle b d| + \langle a d|\langle c b| .
\]

(3.33)

The same identity applies with angle brackets replaced by square ones. These identities can provide enormous algebraic simplification.

3.3.3 Fierz identity

We begin by noting that any 4x4 complex matrix \(M\) can be expressed as a linear combination of the 16 basis matrices \(\Lambda^i\) defined by,

\[
\Lambda^i = \{1, \gamma^\mu, \gamma^\mu, \gamma^\mu \gamma^\mu, \gamma^\mu \gamma^5, \gamma^5 \},
\]

(3.34)

with \(\sigma^{\mu\nu} = i[\gamma^\mu, \gamma^\nu]/2\). There is then a unique dual basis given by,

\[
\bar{\Lambda}_i = \frac{1}{4} \{1, \gamma_\mu, \sigma_{\mu\nu}, \gamma_\mu \gamma^5, \gamma^5 \},
\]

(3.35)

that satisfies the property,

\[
\text{tr} \left( \Lambda^i \bar{\Lambda}_j \right) = \delta^i_j .
\]

(3.36)
If we expand the matrix $M$ as $M = m_i \Lambda^i$ then this relation determines the coefficients $m_i$ and we have,

$$M = \Lambda^i \text{tr} \left( M \bar{\Lambda}_i \right).$$  \hspace{1cm} (3.37)

Making the explicit choice $M = u_-(a) \bar{\pi}_-(b)$ and using cyclicity of the trace we find,

$$u_-(a) \bar{\pi}_-(b) = \Lambda^i \left[ \bar{\pi}_-(b) \Lambda_i u_-(a) \right].$$  \hspace{1cm} (3.38)

We now evaluate the rhs of this equation using our basis elements, where we notice that most terms vanish. The result is,

$$u_-(a) \bar{\pi}_-(b) = \frac{1}{4} (1 - \gamma^5) \gamma^\mu \left[ \bar{\pi}_-(b) \gamma_\mu u_-(a) \right].$$  \hspace{1cm} (3.39)

Similarly, for the choice $M = u_+(b) \pi_+(a)$ we find,

$$u_+(b) \pi_+(a) = \frac{1}{4} (1 + \gamma^5) \gamma^\mu \left[ \pi_+(a) \gamma_\mu u_+(b) \right].$$  \hspace{1cm} (3.40)

Using the reversal identity in Eq. 3.30 these two results can be combined to find the Fierz identity:

$$\gamma^\mu \left[ \pi_+(a) \gamma_\mu u_+(b) \right] = 2 \left( u_-(a) \bar{\pi}_-(b) \right).$$  \hspace{1cm} (3.41)

In our abbreviated notation, the Fierz identity is thus written as,

$$\gamma^\mu [a | \gamma_\mu | b] = 2 (|a \rangle \langle b| + |b \rangle \langle a|).$$  \hspace{1cm} (3.42)

A similar identity clearly applies for the contraction $\gamma^\mu |a \rangle \langle \gamma_\mu | b].$

3.4 Summary

These identities, together with obvious extensions, are summarized in Table 3.1.

3.5 Massless boson polarizations

Incorporating massless bosons, such as photons and gluons, into the spinor formalism is also straightforward. For the polarization vector of a gluon with momentum $k$ the two physical polarizations are expressed as,

$$\epsilon_\pm^\mu(k, b) = \pm \frac{\langle k^\pm | \gamma_\mu | b^\pm \rangle}{\sqrt{2} \langle b^\pm | k^\pm \rangle}$$  \hspace{1cm} (3.43)

where $b$ is the corresponding gauge vector. Explicitly we have that,

$$\epsilon_+^\mu(k, b) = \frac{[k | \gamma_\mu | b]}{\sqrt{2} \langle b | k \rangle}, \quad \epsilon_-^\mu(k, b) = \frac{\langle k | \gamma_\mu | b \rangle}{\sqrt{2} \langle k | b \rangle}.$$  \hspace{1cm} (3.44)

It is straightforward to check that these satisfy the usual relations,

$$\epsilon_\pm^\mu(k, b) = [\epsilon_\pm^\mu(k, b)]^*, \quad \epsilon_\pm^\mu(k, b) k^\mu = 0$$

$$\epsilon_\pm^\mu(k, b) [\epsilon_\pm^\mu(k, b)]^* = -1, \quad \epsilon_\pm^\mu(k, b) [\epsilon_\pm^\mu(k, b)]^* = 0.$$  \hspace{1cm} (3.45)
\[
\langle pq \rangle = \langle p^-|q^+ \rangle, \quad \langle pq \rangle = \langle p^+|q^- \rangle
\]
\[
\langle p^+|q^+ \rangle = \langle p^-|q^- \rangle = \langle pp \rangle = [pp] = 0
\]
\[
\langle p^\pm|\gamma_\mu|p^\pm \rangle = \langle p|\gamma^\mu|q \rangle = [p|\gamma^\mu|q] = 2p_\mu
\]
\[
\langle pq \rangle = -\langle qp \rangle, \quad \langle pq \rangle = -\langle qp \rangle
\]
\[
2|p^\pm\rangle\langle q^\pm| = \frac{1}{2}(1 \pm \gamma_5)\gamma^\mu\langle q^\pm|\gamma_\mu|p^\pm \rangle
\]
\[
\langle pq \rangle^* = -\text{sign}(p \cdot q)\langle pq \rangle = \text{sign}(p \cdot q)\langle qp \rangle
\]
\[
|\langle pq \rangle|^2 = \langle pq \rangle\langle pq \rangle^* = 2|p \cdot q| \equiv |s_{pq}|
\]
\[
\langle pq \rangle|qp\rangle = 2p \cdot q \equiv s_{pq} \quad \text{(note: opposite sign in some literature)}
\]
\[
\langle p^\pm|\gamma_{\mu_1} \ldots \gamma_{\mu_{2n+1}}|q^\mp \rangle = \langle q^\mp|\gamma_{\mu_{2n+1}} \ldots \gamma_{\mu_1}|p^\mp \rangle
\]
\[
\langle p^\pm|\gamma_{\mu_1} \ldots \gamma_{\mu_{2n}}|q^\mp \rangle = -\langle q^\pm|\gamma_{\mu_{2n}} \ldots \gamma_{\mu_1}|p^\mp \rangle
\]
\[
\langle AB\rangle\langle CD \rangle = \langle AD\rangle\langle CB \rangle + \langle AC\rangle\langle BD \rangle, \quad \text{(Schouten)}
\]
\[
\langle A^\mp|\gamma_\mu|B^\mp \rangle\langle C^-|\gamma^\mu|D^- \rangle = 2\langle AD\rangle\langle CB \rangle, \quad \text{(Fierz)}
\]
\[
\langle A^\mp|\gamma_\mu|B^\mp \rangle \gamma_\mu = 2 \left[ |A^\mp\rangle\langle B^\mp| + |B^\mp\rangle\langle A^\mp| \right], \quad \text{(Fierz + Charge conj.)}
\]

\textbf{Table 3.1:} Useful relations for massless spinors.
by using the spinor identities derived in the previous section. Using the Fierz identity, the result when contracted with a gamma matrix (an oft-appearing quantity in actual calculations) is,

\[
\gamma^\mu \varepsilon^\dagger_{\mu}(k, b) = \frac{\sqrt{2} \left[ |k\rangle\langle b| + |b\rangle\langle k| \right]}{\langle b k \rangle}, \quad (3.46)
\]

\[
\gamma^\mu \varepsilon_\mu(k, b) = \frac{\sqrt{2} \left[ |k\rangle |b| + |b\rangle \langle k| \right]}{[k b]}, \quad (3.47)
\]

References


