Jet Physics Lecture 1

HCPSS 2020 Andrew Larkoski

Hello! I'm Andrew Larkoski, a professor at Reed College in Portland, Oregon, USA, and I'm excited to lecture at this year's Hadron Collider Physics Summer School. I have been tasked with discussing hadronic jets, which are collimated streams of particles created from the dynamics of QCD at high energies. In this first lecture, I'll introduce jets from a "bottom-up" perspective, forgoing any explicit discussion of the fundamental QCD Lagrangian from which they arise as an emergent phenomena. This will be similar in approach to the recommended reading of my lectures from CTEQ schools from a few years ago. I've also been asked to discuss machine learning for jet physics. Now, as a theoretical physicist with research interests in calculations, I won't be discussing neural network architecture, programming, computer science, etc., but will instead propose the following. Regardless of your individual thoughts of machine learning, the manner with which data is input and output of a machine suggests a novel way of thinking about fundamental problems in particle physics generally and jet physics specifically. I'll discuss this way of thinking and how it can be used to derive and learn about very general results in the next lecture.

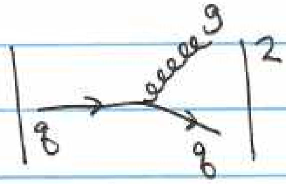
For today, though, let's first understand what QCD is, in the first place. My introduction will be very different (likely) than you've seen before. To define QCD, jets and their consequences, we will only make two assumptions or axioms from which everything in these lectures follows. They are:

1) At high energies, QCD is an approximately scale-invariant quantum field theory. This means that, to first approximation, the coupling of QCD, α_s, is constant, independent of energy.

2) Not only is α_s (approximately) constant, but further α_s is small; formally, we assume that $\alpha_s \ll 1$. This means that the degrees of freedom in the Lagrangian of QCD, quarks and gluons, are good quasi-particles for actually describing the physics of QCD.

These assumptions are sufficient to write down matrix elements for some simple processes. Actually, we'll just write down probability density functions, because "matrix element" (implicitly) assumes specific interactions, but our axioms say nothing explicit about interactions. By axiom 2, we can ask questions about the dynamics of quarks and gluons. So, we'll ask: what is the probability for a quark to emit a gluon? With a Lagrangian and Feynman rules, we would want to calculate the (squared) diagram:

$P_{qg \leftarrow q} =$



But we don't have Feynman rules, so we have to use our axioms.

We have a few things to establish before our axioms, however. The probability for gluon emission will, in general, depend on the four-momentum of the gluon. So, we need to identify the space on which the probability $P_{qg\leftarrow q}$ is defined. Then, given that space, we can ask what constraints scale-invariance imposes. That is, we need to identify the degrees-of-freedom of the emitted gluon.

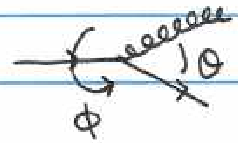
Let's start with the gluon's four-momenta written as:

$p_g = (E, p_x, p_y, p_z)$. The gluon is massless, so demanding

that this momentum be on-shell requires: $E = |\vec{p}|$, or that the momentum can be expressed in spherical coordinates as:

$$p_g = E\left(1, \sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta\right).$$

Here, θ is the quark-gluon opening angle and ϕ is the azimuthal angle about the quark.



Additionally, the total quark+gluon energy is fixed. Calling this E_{tot}, we can write $E = z E_{tot}$, for some fraction $z \in [0,1]$. Additionally, from the formulation of our assumptions and the problem at hand, there was nothing special about the azimuthal angle ϕ: the physics is independent of ϕ. Thus, we can fix ϕ to a convenient value, say, $\phi = 0$. (That is, we assume that the quark is unpolarized.) With these identifications, the gluon's momentum is:

$p_g = z E_{tot}(1, \sin\theta, 0, \cos\theta)$. We had assumed that

the final state quark's momentum is along the z-axis, so it is:

$$p_q = (1-z) E_{tot}(1, 0, 0, 1).$$

We have therefore identified the gluon's relevant degrees of freedom to be the energy fraction z and angle θ. The splitting probability is:

$P_{qg \leftarrow q} = p(z, \theta)\, dz\, d\theta$, for some other function $p(z,\theta)$.

There are a few things we can immediately say about the function $p(z,\theta)$. α_s is the QCD coupling; as such it controls the strength with which quarks and gluons interact with one another. Thus, this function, as it describes the probability of emission of a gluon off of a quark, is proportional to α_s:

$p(z,\theta) \propto \alpha_s$. Actually knowing QCD we can say more, so we might as well add it. The actual factors that come with α_s are:

$p(z,\theta) \propto \frac{2\alpha_s}{\pi} C_F$, where C_F is the fundamental quadratic Casimir of the SU(3) color symmetry of QCD. Again, our assumptions can't tell us about these factors, but they won't qualitatively change the picture we are developing. $C_F = 4/3$ in QCD, and so π, $C_F \sim \mathcal{O}(1)$, so truley α_s is what is controlling the coupling of quarks and gluons, by the assumption $\alpha_s \ll 1$. By the way, C_F is a measure of how quarks and gluons share the three colors of QCD.

Next, we'd like to determine the dependence of $p(z,\theta)$ on the energy fraction z and angle θ. To do this, we need to think about what our assumption of "scale invariance of QCD" means. QCD is a quantum field theory, and as such is Lorentz invariant. Thus, the only quantities that all observers agree on are those that are, well, Lorentz invariant. If we further state that QCD is scale-invariant, this means that a scaling of all Lorentz-invariant quantities produces the same physics. That is, probability distributions of Lorentz-invariant quantities should be further invariant to scale transformations.

Given our quark-gluon system, the only Lorentz invariant we can construct is the dot product of their momenta:

$$p_q \cdot p_g = z(1-z)E_{tot}^2(1-\cos\theta).$$

This is Lorentz invariant by construction and scale invariance means that the scaling $p_q \cdot p_g \to \lambda\, p_q \cdot p_g$, for any $\lambda > 0$, leads to identical physical phenomena. In general, this scaling is not simply implemented on the energy fraction or angle, but there is a limit in which it is simple. If the gluon has low energy or is soft so that $z \ll 1$ and is nearly collinear with the quark so that $\theta \ll 1$, note that

$z(1-z) \xrightarrow[z\ll 1]{} z$, $1-\cos\theta \xrightarrow[\theta \ll 1]{} \frac{\theta^2}{2}$. Then in this double limit, the dot product is:

$$p_q \cdot p_g \xrightarrow[z\ll 1,\, \theta \ll 1]{} z\theta^2 \frac{E_{tot}^2}{2}.$$

Thus, this soft and collinear limit corresponds to this dot product expressed as a power-law function in both z and θ. A scaling of $p_q \cdot p_g$ can therefore be accomplished by scaling of either z or θ (or both). Then, we identify the scale transformation under which QCD is invariant as:

$$z \to \lambda z \text{ or } \theta^2 \to \lambda \theta^2, \text{ for } \lambda > 0.$$

Before continuing, note that there is a coordinated scale transformation under which the product $z\theta^2$ is unchanged. If we scale

$$z \to \lambda z \text{ and } \theta^2 \to \frac{\theta^2}{\lambda}, \text{ then } z\theta^2 \to z\theta^2.$$

That is, such a coordinated scale transformation is actually a Lorentz transformation, a boost along the direction of the quark's momentum. Lorentz invariance states that if energies increase, angles must decrease to ensure that dot products are unchanged.

Now, if scale transformations are implemented by $z \to \lambda z$ and $\theta^2 \to \lambda \theta^2$ and the probability $P_{qg \leftarrow q}$ must be unchanged by this scaling, this uniquely fixes the z and θ dependendence of $p(z, \theta)$ to be:

$$P_{gq\leftarrow q} = p(z,\theta)\, dz\, d\theta = \frac{2\alpha_s}{\pi} C_F \frac{dz}{z} \frac{d\theta}{\theta}.$$

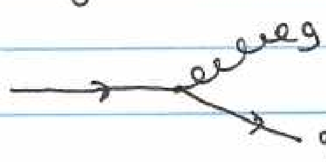
This is invariant to scalings and therefore satisfies our first assumption about QCD. Note that for this simple functional form, we had to work in the soft and collinear limit. While this may seem restrictive, we'll be able to get a lot of mileage out of it.

The first thing to note about this probability distribution is that, in the z or $\theta \to 0$ limits, it diverges. Actually, it's worse than that: not only does it diverge, but it is not even integrable as $z, \theta \to 0$! Thus, it isn't a "probability distribution" at all, because it cannot be normalized. So how can such a distribution describe some physical process which is non-singular? Note that the z or $\theta \to 0$ limits are those limits in which the gluon becomes unobservable. The $z \to 0$ limit is the limit in which the gluon has no energy. A detector, like the calorimeters at ATLAS or CMS, requires a finite energy of the particle to observe it: it never "sees" a 0 energy "hit". The $\theta \to 0$ limit is when the gluon is collinear with the quark. Because it is traveling in the exact same direction as the quark, there is no measurement you can perform to separate them out. The angular resolution of the cells of the calorimetry at ATLAS and CMS is finite:

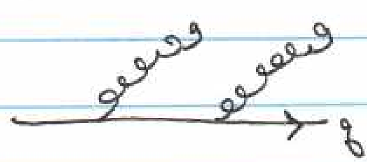
the particles must be a non-zero angle from one another to be distinguished.

That is, there is no measurement you can perform in the $z \to 0$ or $\theta \to 0$ limit to distinguish a quark emitting one gluon to the case in which the quark emits *no* gluons. That is, we call the $z \to 0$ or $\theta \to 0$ limits *degenerate*, as the physical configuration *degenerates* to a system with fewer gluons. But this also points to an extension. There is no measurement we can perform to distinguish a quark that emitted no gluons:

from a quark that emitted one soft and/or collinear gluon:



or two soft and/or collinear gluons:



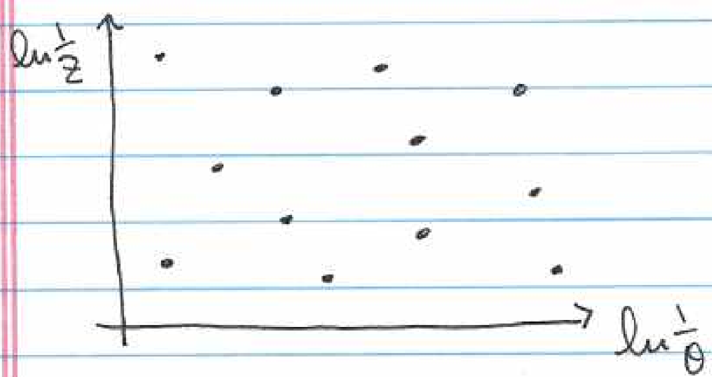
or *any* number of soft and/or collinear gluons! Every one of these configurations of soft and/or collinear gluons + quark have a divergent probability, and all are degenerate with each other. So how do we proceed?

Quantum mechanics saves us! Feynman diagram perturbation theory is a degenerate perturbation theory. And just like degenerate perturbation theory in quantum mechanics, we only get a finite result by summing over all degenerate configurations. In quantum field theory, this is called the KLN theorem (for Kinoshita, Lee and Nauenberg), and KLN ensures that summing over all the individual divergent degenerate configurations produces a finite result.

We'll end this lecture by seeing how this is done. Let's first go back to the probability distribution

$$P_{qg \leftarrow q} = \frac{2\alpha_s}{\pi} C_F \frac{dz}{z}\frac{d\theta}{\theta} = \frac{2\alpha_s}{\pi} C_F \, d\ln\frac{1}{z}\, d\ln\frac{1}{\theta}.$$

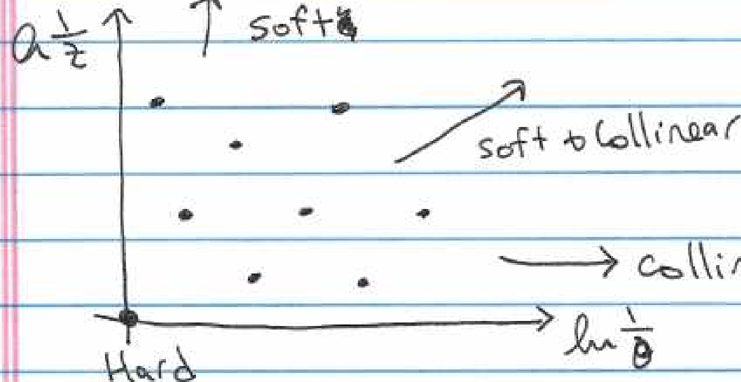
On the right, I just re-expressed the probability as flat logarithmically in z and θ. Thus, the phase space in the $(\ln\frac{1}{z}, \ln\frac{1}{\theta})$ coordinates is a semi-infinite region where $\ln\frac{1}{z}, \ln\frac{1}{\theta} > 0$. Further, the probability distribution is flat: that is, there is uniform probability for the emission to be anywhere on the plane. Thinking with degeneracy and KLN in mind, we can generalize this and say that the (arbitrary) gluon emissions off of a quark uniformly fill out the plane. Representing each emission by a dot, for emissions off of a quark, we might have:



Uniform in this logarithmic plane means exponentially far apart in "real" space, so gluon emissions off of a high energy quark are dominantly soft and/or collinear. That is precisely what a jet is.

This plane is called the Lund Plane, after researchers in Sweden who introduced it for studying jets. Also, this uniform emission distribution is the starting point of modern Monte Carlo event generators and parton showers. A modern Monte Carlo, like Pythia, Herwig or Sherpa, contains significant physics beyond this uniform assumption, such as: running α_s (emissions increase as z, θ decrease), fixed-order corrections (corrections to $\frac{1}{z}, \frac{1}{\theta}$ distributions), cut off by the scale of hadron masses, etc. Nevertheless, while simple, this uniform emission phase space has a lot of physics.

Before we do a calculation, I want to orient you in the Lund plane. Let's draw it again:



and I've also included regions to guide the eye. The origin, where $z, \theta \sim 1$, corresponds to high-energy, wide-angle gluon emission. The degenerate limits live off at ∞, with different physical origins for different ∞s. Vertical in the plane, $z \to 0$, is the soft limit, horizontal is the collinear limit $\theta \to 0$, and diagonal is a soft and collinear limit.

With this picture in place, let's now calculate the distribution of a particular observable called an angularity, τ_α. The angularity can be defined in the energy fraction z / angle coordinates as:

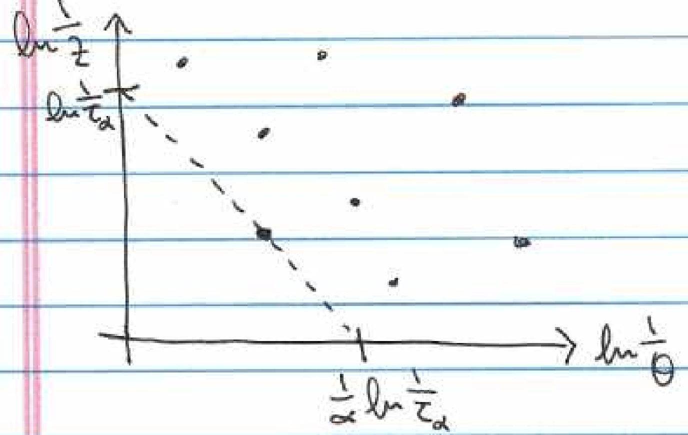
$\tau_\alpha = \sum_{i \in J} z_i \theta_i^\alpha$, where the sum runs over all particles i in the jet (collection of emissions),

and $\alpha > 0$ is an exponent that weights contributions from different angles. We require that $\alpha > 0$ to ensure that the observable is <u>infrared</u> and <u>collinear safe</u>: as we will see, this means that arbitrarily soft or collinear emissions cannot contribute to the observable (at least not dominantly so). One important thing to note is that there is no preferred ordering to emissions of a scale-invariant system. "Scale-invariant" means that <u>any</u> scale we impose on the system can and will exist in that system. We will see how measuring the angularities sets one particular ordering.

We had mentioned before that uniform logarithmically means exponentially far in real space. Because each term in the definition of the angularities is weighted by a product of energy and angle, there will be <u>one</u> emission in the jet that dominates the value of τ_α, and all others will be exponentially small. So, with one emission dominating, note that

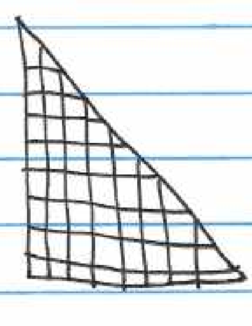
$$\tau_\alpha = z\theta^\alpha \quad \text{or} \quad \ln\frac{1}{\tau_\alpha} = \ln\frac{1}{z} + \alpha\ln\frac{1}{\theta},$$

and so a fixed value of τ_α corresponds to a straight line on the $(\ln 1/z, \ln 1/\theta)$ plane. We can draw this as:



We have drawn the x- and y-intercepts for the fixed value of τ_α represented by the dotted line. There is one emission on the dotted line, and all other emissions in the jet lie above the line. Indeed, emissions below the dotted line are forbidden, as they would act to increase the value of τ_α to be larger than what was measured.

So, for the measured value of τ_α, we must forbid all emissions at any point in the triangle below the dotted line. To calculate the probability that the measured value of τ_α is not larger than its value, we will do the following. We must forbid emissions everywhere in the triangle, so let's isolate it and break it up into many pieces:



There can be no emissions in any of the subregions. This is an "and" statement in probability, so we must multiply the probability for no emissions in all regions together.

Let's breake the triangle into N equal area regions. Then, the probability for an emission in one of the regions is uniform and equal to:

$$\text{Prob for emission} = \frac{2\alpha_s}{\pi} C_F \frac{\Delta}{N}, \quad \text{where } \Delta$$

is the area of the triangle: $\Delta = \frac{1}{2\alpha} \ln^2 \tau_\alpha$.
Then, the probability for <u>no</u> emission in one small region is one minus this:

$$\text{Prob for } \underline{\text{no}} \text{ emission} = 1 - \frac{2\alpha_s}{\pi} C_F \frac{\Delta}{N} = 1 - \frac{\alpha_s}{\pi} \frac{C_F}{\alpha} \frac{\ln^2 \tau_\alpha}{N}$$

Then, the probability for no emissions anywhere in the triangle is the product of probabilities of no emissions anywhere in each sub region:

$$P(\tau_\alpha \text{ less than measured value}) = \lim_{N \to \infty} \left(1 - \frac{\alpha_s}{\pi} \frac{C_F}{\alpha} \frac{\ln^2 \tau_\alpha}{N}\right)^N$$

$$= \exp\left[-\frac{\alpha_s}{\pi} \frac{C_F}{\alpha} \ln^2 \tau_\alpha\right].$$

The product transmogrifies into an exponential! This exponential factor is called the <u>Sudakov form factor</u>, and is simply a manifestation of the scale-invariant Poisson process of particle emission in high-energy QCD.

From a probability perspective, this Sudakov factor is the cumulative probability distribution or CDF of τ_α. To find the (differential) probability distribution function, we just differentiate:

$$p(\tau_\alpha) = \frac{d}{d\tau_\alpha} \exp\left[-\frac{\alpha_s}{\pi}\frac{C_F}{\alpha}\ln^2\tau_\alpha\right] = -\frac{2\alpha_s}{\pi}\frac{C_F}{\alpha}\frac{\ln\tau_\alpha}{\tau_\alpha}\, e^{-\frac{\alpha_s}{\pi}\frac{C_F}{\alpha}\ln^2\tau_\alpha}.$$

This probability is normalized on $\tau_\alpha \in [0,1]$. The issues with divergences with any fixed number of gluon emissions has been transformed into exponential suppression with the Sudakov form factor.

That's it for today - we'll use this intuition to understand aspects of machine learning next lecture. Below are a couple of exercises.

(1) Consider the measurement of two angularities, τ_α and τ_β, with, say, $\alpha > \beta > 0$. Calculate the Sudakov form factor for two angularities, the joint probability distribution $p(\tau_\alpha, \tau_\beta)$. Further, ensure that the joint probability marginalizes to the correct single probability distributions. That is,

$$\int_{\tau_0}^{\tau_1} d\tau_\beta\, p(\tau_\alpha, \tau_\beta) = p(\tau_\alpha), \text{ for particular bounds } \tau_0 < \tau_\beta < \tau_1.$$

For a hint to this problem, see arXiv: 1307.1699

(2) (This is an extension of exercise 9.3 in my particle physics textbook)

The ALEPH experiment at LEP measured the number of jets produced in $e^+e^- \to$ hadrons collisions. The experiment counted n jets, if, for every pair i,j of jets, the following inequality is satisfied:

$$2 \min[E_i^2, E_j^2](1-\cos\theta_{ij}) > y_{cut} E_{cm}^2$$

for $y_{cut} < 1$, E_i is the energy of jet i, θ_{ij} is the angle between jets i and j and E_{cm} is the center-of-mass energy. In the soft and collinear limits, determine the probability p_n for observing n jets, as a function of y_{cut}.

Note that the minimum number of jets is 2 ($e^+e^- \to q\bar{q}$) and gluons can be emitted from either the quark or anti-quark. Compare your result to Figure 7 of Eur. Phys. J. C 35 457-486 (2004), doi: 10.1140/epjc/s2004-01891-4.
What value of α_s fits the data the best? This is imperfect because we're omitting a lot of physics, but it will be qualitatively close.