A Crash Course for Summer Research

Summer 2021

Xiaochao Zheng

University of Virginia

In this document I provide a crash course on the topic of special relativity, 4-vectors; electron scattering and its kinematics including both fixed-target and collider settings; scattering cross sections, asymmetries, structure functions and parton distribution functions; and typical particle detectors. The target audience are students who have completed a minimum of one year of introductory physics courses and are interested in conducting research in the field of nuclear and particle physics with a focus of using electron scattering to study the nucleon structure and the Standard Model. Those who have completed Modern Physics (special relativity) may skip section 1, and those who completed an upper level electrodynamic or even general relativity can also comfortably skip most of section 2. Short problems are provided for all sections beyond section 1. Given the wide range of the physics background of the target audience, this document is best used in a "flipped" method, *i.e.* provided to students as pre-"lecture" reading material and then spend two hour-long "lecture" time to review concepts and do exercises together.

Contents

1 Introduction to Special Relativity

In this chapter we will present the theory of special relativity using language of (3)-vectors, i.e. the typical language used in intro-level physics courses. Most textbooks introduce special relativity from Einstein's two postulates, published in 1905. I will introduce it from a more historical point of view. It can be beneficial if you have learned electricity and magnetism at an upper level, if not, no worries, just glance through and you might find it helpful when you take the core course (PHYS3430) in the future.

1.1 Galilean Relativity

From our observations of ordinary objects we know that different observers don't agree on the velocity of objects. To an observer on a moving train a rock in his hand is at rest, while to an observer on the ground the rock is moving with the velocity of the train. However, it is well known that neither observer is preferred over the other. Galileo formulated this notion before Newton, and the equivalence of all inertial frames is known as *Galilean relativity*. More specifically:

- All systems in which Newton's Law is valid are said to be inertial frames;
- All inertial frames are moving with constant velocity with respect to each other;
- None of the inertial frame is any more fundamental than the other.

Galilean relativity works well for classical objects. One important consequence of such viewpoint is that since there is no frame that is more special than the other, all physical laws should be valid in all inertial frames. To see if this works, the general procedure is that we write down the law in one inertial frame, apply transformation of coordinates (time and position) using a certain rule which transforms the law into a new form of equation, and we examine whether the new equation is consistent with the physical law in the new frame. If it is, we say the physical law is invariant and we are happy. If it is not, then the physical law we started with must not be true and we are in trouble.

As an example, we write Newton's second law for an object of velocity $\vec{v}$, as observed by an observer S:

$$\vec{F} = m\vec{a} = m\frac{d\vec{v}}{dt}. \tag{1}$$

If another observer (S') is in relative constant motion $\vec{u}$ [1] with respect to observer S, then what he/she will observe are

$$\vec{r'} = \vec{r} - \vec{u}\;t; \quad \vec{v'} = \frac{d}{dt}\left(\vec{r} - \vec{u}\;t\right) = \vec{v} - \vec{u}; \quad \vec{a'} = \frac{d}{dt}\left(\vec{v} - \vec{u}\right) = \vec{a} \tag{2}$$

The two observers disagree on position and velocity but they agree on acceleration. Therefore they will conclude that Newton's second law is invariant.

[1] To avoid confusion, I will use $\vec{v}$ as the velocity vector of the object and $\vec{u}$ as the velocity vector of the inertial frame throughout this document.

1.2 Maxwell's Electrodynamics

From Galilean relativity one important conclusion we can make is that the concept of absolute velocity is meaningless. However, if you encountered Maxwell's equations for electricity and magnetism:

$$\nabla \cdot \vec{E} = \rho/\epsilon_0 \tag{3}$$
$$\nabla \cdot \vec{B} = 0 \tag{4}$$
$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{5}$$
$$\nabla \times \vec{B} = \mu_0 \vec{J} + \frac{1}{c^2}\frac{\partial \vec{E}}{\partial t} \ , \tag{6}$$

you may have noticed that Maxwell's equations yield a velocity, c, as a fundamental quantity. The system in which light has the velocity of c would seem to be special and we call it "fundamental frame". All other coordinate systems, except those at rest with respect to the fundamental one, would not observe light traveling at c. Because Earth is rotating around its axis and is moving around the sun, and the solar system is moving within the galaxy, it is hard to believe that the earth system is at rest w.r.t. the fundamental one. If we call the fundamental one the frame of the ether, then we could detect our motion relative to the ether by measuring the velocity of light along different directions and study the differences. Because c is so large, it was not easy to detect the effect. Nevertheless, an experiment was done by Michelson and Morley in 1887. But what they discovered is that the speed of light is exactly the same in all directions.

The results of the Michelson-Morley experiment came out like a mystery. The frame of the ether was not found. A more importantly question is –

> *Why should Newton's Laws obey Galilean invariance, which seemed in accord with notions of the homogeneity of space, while the laws of electricity and magnetism are not?*

The lack of equivalence of inertial frames in Maxwell's equations can be demonstrated from the equations themselves. An extreme measure of the lack of invariance was provided by Einstein in the following form. Imagine that you are traveling along with a plane wave of light. You would observe stationary electric and magnetic fields and both have a wave-like shape:

$$\vec{E} = E\cos(kz)\hat{x} \ , \qquad \vec{B} = \frac{1}{c}\cos(kz)\hat{y}$$

These fields satisfy $\nabla \cdot \vec{E} = 0$ and $\nabla \cdot \vec{B} = 0$. But

$$\nabla \times \vec{E} = \frac{\partial E_x}{\partial z}\hat{y} = -kE\sin(kz)\hat{y} \ , \qquad \frac{\partial}{\partial t}\vec{B} = 0 \ .$$

so Faraday's Law is not valid in this frame.

One can prove that Faraday's law is not invariant in the more general case, but we will omit that proof here. In short, **Maxwell's equations are not Galilean invariant.** Something must be wrong here, but what could it be? The first thought was that Maxwell's equations, which were only 20 years old back then, are incomplete. People tried to add new terms to these equations so the new equations will remain unchanged under Galilean transformation. However all these new terms led to predictions of new electromagnetic phenomena that did not exist at all when tested experimentally. So this attempt had to be abandoned. Gradually, people started to think that trouble must be sought elsewhere. Could it be that Galilean transformation is incorrect?

Meantime, Lorentz noticed a remarkable and curious thing when he made the following substitutions in Maxwell's equations:

$$x' = \frac{x - ut}{\sqrt{1-u^2/c^2}}, \tag{7}$$

$$y' = y, \quad z' = z, \tag{8}$$

$$t' = \frac{t - ux/c^2}{\sqrt{1-u^2/c^2}}\ . \tag{9}$$

Namely, Maxwell's equations remain invariant when this transformation is applied to them! [2]

Following a suggestion originally made by Poincaré, Einstein (who believed that Maxwell's equations should also obey equivalence of relativity) proposed that *all the physical laws* should remain unchanged under a Lorentz transformation. In other words, we should change, not the laws of electrodynamics, but the laws of mechanics. And because Maxwell's equations are valid in all inertial frames, the velocity c must be a special number that do not change when going from one frame to another. All these led Einstein to propose two simple postulates.

1.3 Einstein's Two Postulates

In 1905 Einstein published his famous paper and proposed two postulates:

1. The laws of physics (including Maxwell's equations) apply in all inertial frames.

2. The speed of light in vacuum is the same for all inertial observers, regardless of the motion of the source or the observer.

Although the second postulate came from Maxwell's equations naturally, the universality of the speed of light seems to be the most striking concept back then. It implies that if an object travels with c in one inertial frame, then it also travels with c in another inertial frame. This contradicts Galilean transformation but can be proved using Lorentz transformation, which we omit again here.

[2] Lorentz did not invent these transformation out of the blue. When studying retarded potentials for a charge moving at constant velocity, it's possible to show that it leads to Lorentz transformation. I hope you have done this as HW in PHYS3430, which means you would have followed the footstep of Lorentz and watched the birth of Lorentz transformation out of your homework paper).

1.4 Examples

After introducing the two postulates, we should in principle discuss their consequences, namely time dilation, length contraction, simultaneity, and derive velocity addition rule. But these are typical textbook material. Therefore we will omit these and dive straight into 4-vectors in the next section. If interested, you can read relevant Modern Physics textbooks to learn more or for a review. I have reserved two examples related to particle and nuclear physics, below.

Example 1: Experimental verification of time dilation in Particle Physics:

Although it is difficult to test special relativity using ordinary objects, the phenomenon of time dilation has been tested an uncountable number of times in nuclear and particle physics. In the sub-atomic world there are a host of particles and systems that decay to a more stable system. These decays always occur randomly, but the probability that a system will decay in a certain time interval is simply proportional to the time interval. If one has a system of N such particles at time t, all at rest in the observer's reference frame the number that decay in an infinitesimal time Δt is

$$\Delta N = \beta N \Delta t \tag{10}$$

or moving to the limit of $\Delta t \to 0$,

$$\frac{dN}{dt} = -\beta N \tag{11}$$

where the minus sign acknowledges the fact that the number of objects decreases with time. The factor β is a property of the unstable particle, and is the same for all such particles. If we solve the equation,

$$N = N_0 e^{-\beta t} = N_0 e^{-\frac{t}{\tau}} \tag{12}$$

N_0 is the number of particles at time $t = 0$, and the constant $\tau = 1/\beta$ is called the life-time of the particle. It is the time for then number of particles to decrease by a factor of $1/e$. This life-time is the life time as seen by an observer at rest with respect to the particle. If on the other hand the particle is moving relative to the observer, the observer will find that time in the particle frame is passing more slowly. If the particle is moving at $0.9c$, which is not unusual for beams of pions or mu mesons at accelerators, the time dilation factor is

$$\gamma = \frac{1}{1 - 0.9^2} \approx 2.4 \tag{13}$$

and the particle's life time will appear to be 2.4 times greater than when it was at rest. It will travel a longer distance in the laboratory: $0.9c \times 2.4\tau$, than the $0.9\tau c$ one would naively expect if there were no time dilation. One can observe the rate of disappearance of particles from a beam, or even directly observe the decays with detectors, and the time dilation factor describes the observed life-time exactly. Moving clocks run slow, and time ceases on a light beam. This is important when calculating the counting rate of unstable particles produced by a scattering process, because for scattering experiments detectors are usually placed far away from the scattering point.

Example 2: Spheres in inertial frames:

Two inertial frames S and $\bar{S}$, $\bar{S}$ is moving at a constant velocity $\vec{u} = u\hat{x}$ *w.r.t.* S and at $t = \bar{t} = 0$ the two origins coincide. What is the shape of a spherical object with radius R, at rest in S, as seen by an observer in the $\bar{S}$ frame?

Sol.: For simplicity we can center the object at the origin of S. Its surface satisfy

$$x^2 + y^2 + z^2 = R^2.$$

We can use Lorentz transformation to find its shape in the $\bar{S}$ frame. Putting in $x = \gamma(\bar{x} + u\bar{t})$, $y = \bar{y}$, $z = \bar{z}$ we find:

$$\frac{(\bar{x} + u\bar{t})^2}{1 - u^2/c^2} + \bar{y}^2 + \bar{z}^2 = R^2.$$

This is an ellipsoid moving along the $-\bar{x}$ direction. For each moment $\bar{t}$ it is centered at $\bar{x} = u\bar{t}$ and $\bar{y} = \bar{z} = 0$. Its equator radius is R (along y and z) and the polar radius is R/γ (along x). You may also call it an oblate spheroid.

Later when we study electron scattering and the structure of the proton (or the nucleon in general), we often talk about a certain inertial frame – the infinite momentum frame – in which the proton is "boosted" to infinite momentum and appears like a "pancake". As seen in the example above, when $\gamma \to \infty$, the oblate spheroid will be completed squashed and thus, becomes a pancake.

2 Four Vectors

2.1 The Concept of Four Vectors

According to Einstein's postulates, the laws of physics apply in all inertial frames and they follow Lorentz transformations:

$$\begin{cases} \bar{t} = \frac{t - ux/c^2}{\sqrt{1-u^2/c^2}} = \gamma(t - ux/c^2) \\ \bar{x} = \frac{x - ut}{\sqrt{1-u^2/c^2}} = \gamma(x - ut) \\ \bar{y} = y \\ \bar{z} = z \end{cases} \tag{14}$$

where $\gamma \equiv 1/\sqrt{1 - u^2/c^2}$ and the $\bar{S}$ frame is moving with constant speed u along the x direction w.r.t. the S frame. This transformation can be written in the matrix form as:

$$\begin{pmatrix} c\bar{t} \\ \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{15}$$

It is therefore more convenient to define a four-vector, or a set of four quantities:

$$\begin{aligned} x^0 &= ct \\ x^1 &= x \\ x^2 &= y \\ x^3 &= z \end{aligned}$$

and the matrix on the RHS of Eq.(15) is called the Lorentz transformation matrix, Λ. Equation (15) can now be written in a more compact form as

$$\bar{x}^\mu = \sum_{\nu=0}^{3} \Lambda^\mu_\nu x^\nu \tag{16}$$

where the superscript in Λ^μ_ν is for rows and the subscript is for columns. The convenience here is that even if the velocity of $\bar{S}$ frame is not along the x-direction, the details of Λ will change, but Eq.(16) and later (17) will remain its simple form.

There are two conventions about the superscript used in four-vectors:

(1) Earlier (before starting relativity) we often use x^1, x^2, x^3 for the ordinary coordinates x, y, z. To distinguish 3-vectors from 4-vectors, or from 3D coordinates to the new space-time (4D) coordinates, we now use only Greek letters $\mu, \nu, \rho, \ldots$ for 4-vectors, which run from 0 to 3, and Latin letters $i, j, k \ldots$ for 3-vectors, which run from 1 to 3.

(2) Summation is implied whenever a Greek letter is repeated in a product, once in the superscript and once in the subscript. (This is called the Einstein summation convention). For example we can now write Eq. (16) as

$$\bar{x}^\mu = \Lambda^\mu_\nu x^\nu \ . \tag{17}$$

The general definition of four-vector is: **any set of four numbers (or quantities)** (a^0, a^1, a^2, a^3) **that follow Lorentz transformation between inertial frames are called four-vectors**.

2.2 Invariant Products

Given two four vectors $a = (a^0, a^1, a^2, a^3)$ and $b = (b^0, b^1, b^2, b^3)$, their dot product is defined as $a \cdot b = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3$, or

$$a \cdot b \equiv g_{\mu\nu} a^\mu b^\nu = a^\mu b_\mu, \tag{18}$$

where $g_{\mu\nu}$ is called the metric tensor. Until we deal with gravity, it's safe to assume our spacetime is "flat", i.e.

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{19}$$

which is also called the Minkowski space-time metric. Using the convention introduced on the previous page, the summation is implied for both μ and ν. The new symbol, $b_\mu \equiv g_{\mu\nu} b^\nu$ (with summation over ν implied), is called the covariant vector and the previous b^μ is called the contravariant vector – one appears like a "row" and the other a "column". Please be familiar with Eq. (20) and how the summation is done. This is often called **"contraction" of 4-vectors**. Technically, you can write it in a similar format as the matrix operations: [3]

$$\begin{aligned} a \cdot b &\equiv (a^0, a^1, a^2, a^3) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} b^0 \\ b^1 \\ b^2 \\ b^3 \end{pmatrix} \\ &= (a^0, a^1, a^2, a^3) \begin{pmatrix} b^0 \\ -b^1 \\ -b^2 \\ -b^3 \end{pmatrix} = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3. \end{aligned} \quad (20)$$

We will see similar operations over and over until it becomes natural to you. When it becomes natural, it will no longer be important to distinguish row from columns, you only need to track "upper" and "lower" indices to make sure they "contract" properly.

Why are these dot products important? This is because **all dot products of four-vectors are invariant**. (Again, this is provable using Lorentz transformation.) We often deal with *only* these invariant quantities when talking about a process so we do not need to worry about specifying which reference frames we are working with. We will give one such invariant quantity below.

Space-time intervals

From now on we shall always consider time with spatial coordinates together, and study our subjects in this 4D "space-time".

- When we talk about an event, we must specify all 4 components;
- When we say we observe an event, it means the observer must be at exactly the same time and the same location as the event.

We are often interested in the space-time difference between two events. Let's call them event A and B, that happen at (t_A, x_A, y_A, z_A) and (t_B, x_B, y_B, z_B), respectively. The difference between A and B is

$$\Delta x^\mu = x^\mu_A - x^\mu_B$$

which is a 4-vector itself and thus its square is invariant. We define it as the invariant

[3] Be careful with the upper and lower indices, terms like $a^\mu b^\mu$ represent two "column" vectors and cannot be contracted, only terms like $a^\mu b_\mu$ – a row vector followed by a column vector – can.

interval [4]:

$$-I \equiv -(\Delta x)_\mu (\Delta x)^\mu = -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 = -(c\Delta t)^2 + d^2$$

where d is the ordinary spatial distance between these two events.

In contrast to the distance we use in the ordinary 3D space, which is always greater than or equal to zero: $d^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \geq 0$, the space-time interval does not have to be non-negative.

1. If $I < 0$ the interval is called timelike, meaning there is always an inertial frame in which the two events happen at the same position ($d = 0$) and at different times. Two events that are cause and effect are called "causal", and their interval must be timelike. In the frame in which the cause and effect happen at the same position, the cause happens earlier than the effect in time.

2. If $I > 0$ the interval is called spacelike, for there is always an inertial frame in which the two events happen at the same time ($\Delta t = 0$) and different locations. It also mean it is impossible for the two events to be causal: event A cannot possibly effect B instantaneously because no information can transmit in spacetime faster than the speed of light in vacuum c.

3. If $I = 0$ the interval is called lightlike, for this is the interval between two events which are connected by a signal traveling at the speed of light. Such two events can be causal.

Now one can see that the fact that the interval is an invariant quantity guarantees that if two events are causal in one inertial frame, then they must be also causal when observed in another inertial frame. Otherwise, all hell breaks lose. [5]

2.3 The 4-Momentum Vector

In your introductory physics courses you have learned the concept of kinetic and potential energies, and your professor may have made you to solve problems using both Newton's law and energy conservation. You may have wondered: what's the point of using two different methods if you can solve the problem using only one? Well, the answer may have become or will become clear when you take classical mechanics, and then more in quantum mechanics or electrodynamics. You will learn that the language of force and acceleration cease to be meaningful nor useful, and the language of more fundamental physical laws use solely concepts such as energy, Hamiltonion, Lagrangians etc.

[4]Depending on which textbook or metric convention you follow, you may see some versions of the I differs by a minus sign. A theorist colleague used to tell me: "The sign does not matter. The important thing is to pick a convention and stick to it and make sure all your work is self-consistent". To which I shall add: "And make sure the world understands you."

[5]Imagine in one inertial frame, if you ace the final exam then you get an A in a course; while when observed by your roommate in another inertial frame, this is no longer true and they conclude you get A's no matter what and invite a party the night before your next final exam. Of course, we have observed in the past year that some people do not see causality presented by Science no matter which frame they observe the world from, but that's a different story.

In this section we will introduce the concept of relativistic energy and relativistic momentum, and the 4-momentum vector. We will not cover proper time, proper velocity, etc, which (again) you can find in modern physics textbooks.

Relativistic Energy and Momentum

In classical mechanics the linear momentum of an object moving with velocity $\vec{v}$ is $\vec{p} = m\vec{v}$. The generalization of momentum to four-dimensional space time is made by defining the spatial components as

$$\vec{p} = \gamma m\vec{v} \ \text{ or } \ p^i \equiv \gamma m v^i, \tag{21}$$

where $\gamma \equiv 1/\sqrt{1 - v^2/c^2}$ with $v = |\vec{v}|$ the speed of the particle [6], and $i = x, y, z$ for the component format. For ordinary objects, $v \ll c$, $\gamma \approx 1$ and we revert to the classical momentum vector. To form the 4-momentum vector we also need a "time component", which is

$$p^0 = \frac{mc}{\sqrt{1 - v^2/c^2}} = \gamma mc = \frac{E}{c}, \tag{22}$$

where $E \equiv \gamma mc^2$ is the relativistic energy with m the (good old) mass of the object. One can see that for an object at rest, $E = mc^2$ which is called the "rest energy" of the object. For an ordinary object with $v \ll c$, $E \approx mc^2 + \frac{1}{2}mv^2$, i.e. the usual kinetic energy you learned in intro physics plus the new term rest energy. Note, however, that we rarely use the classical "kinetic energy" concept in the field of nuclear and particle physics because mass can become energy and vice versa, and kinetic energy is no longer conserved for a closed system. This is replaced by the new "relativistic energy" concept which is still conserved. Putting things together, we have the energy-momentum 4-vector or simply the 4-momentum vector:

$$p^\mu = (E/c, \vec{p}) = (\gamma mc, \gamma m\vec{v}). \tag{23}$$

Two great achievements of classical physics was to demonstrate that momentum and energy are separately conserved in a closed system (closed = no outside forces or interactions). The same rule applies here: **both relativistic energy and relativistic momentum of a closed system are conserved.**

For a system of N particles, the total momentum and energy are given by $P^\mu = \sum_{i=1}^{N} p_i^\mu$. Each component of this equation is conserved in any closed system, even when there are interactions between the particles to change their individual momenta.

Invariant Mass

The invariant square of the 4-momentum vector is

$$p^2 = p^\mu p_\mu = (p^0)^2 - |\vec{p}|^2 = \frac{m^2c^2}{1 - u^2/c^2} - \frac{m^2v^2}{1 - u^2/c^2} = m^2c^2 \ , \tag{24}$$

[6] Note: If you apply Lorentz transformation to figure out how it moves as observed from a moving frame, there will be another γ factor involved which is defined by the frame's velocity $\vec{u}$. Be very careful to not confuse the two.

where m is called the invariant mass. For a stable particle or ordinary object, mass m never changes so is obviously invariant. [7] However, this concept is extremely useful when solving scattering problems in particle and nuclear physics. As an example, when we study or look for an unstable particle that is produced in a particle collision or scattering, we do not directly detect the unstable particle because it decays quickly. If it decays into two daughter particles, which are detected and we know their 4-momenta as p_1^μ and p_2^μ, we add them together $p^\mu = p_1^\mu + p_2^\mu$, then take the square $p^2 = p^\mu p_\mu$ and make a graph of the number of events plotted vs. $\sqrt{p^2}$. This is called a "missing mass spectrum". If we see a sharp peak located at a certain value, e.g. 135 MeV, then we know the two particles detected are from the decay of an unstable partile of mass 135 MeV (which is a π^0 meson in this short example).

For ultrarelativistic particles, $E \gg mc^2$ and $E = pc$, which we often use for the electrons at JLab. For high energy colliders, one often omit the proton mass (about 1 GeV) as well.

2.4 Invariant Quantities in Two-Body Scattering

We will introduce the kinematics and some invariant quantities that are used in scattering or collision experiments. We will focus on the general defintions and calculations, and leave the physics of electron scattering (*i.e.* why we do it, what do we learn from it) to the next section. Prior knowledge of the Standard Model of Particle Physics is assumed, which can be found in typical textbooks (modern physics level is sufficient).

In the language of the Standard Model, all interactions are mediated by force-carrying bosons. In electron scattering off a nucleon or nuclear target, or ep and eA collisions, both electromagnetic (exchange of a virtual photon) and weak (exchange of virtual Z and W bosons) are possible. We will focus on EM interactions here. The simplest picture of such scattering process is shown in Fig. 1, with the 4-momenta vectors k and k' for the incident and the scattered electrons, respectively. The 4-momentum of the initial proton

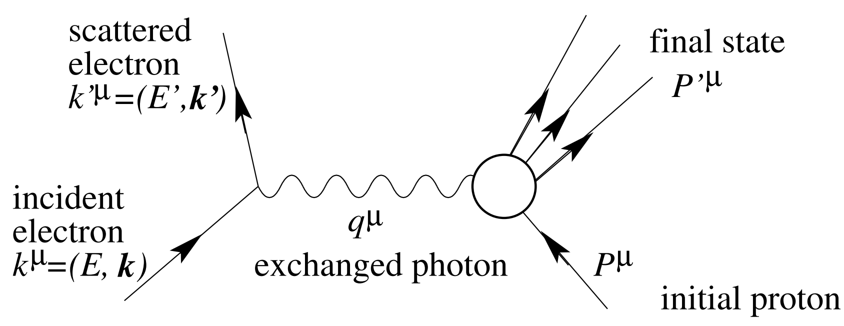

Figure 1: One photon exchange process of scattering between an electron and a proton (or a nucleus).

[7] You will soon realize that we never use kg to describe particle mass, we use GeV or MeV as particle mass units, such as "electron mass is 0.511 MeV" or "proton mass is 938.272 MeV", where $/c^2$ is implied. Similarly, momentum unit is MeV or GeV where $/c$ is implied. In fact, one often use "Natural Units" in which one sets $\hbar = c = 1$.

(or nucleus) is P. After scattering, the proton may stay intact or may be broken apart, and we denote P' as the total final state's 4-momentum. If the proton is broken apart, P' is the sum of all final state particles' 4-momenta. The exchanged photon carries 4-momentum q. In the so-called "inclusive" scattering studies, only the scattered electrons are detected, *i.e.*, k' is known, in addition to the two known initial states, k and P. The angle θ formed by the three-momentum vectors $\vec{k}$ and $\vec{k}'$ is called the scattering angle, and one also defines "pseudorapidity", $\eta = -\ln[\tan(\theta/2)]$, for collider settings.

Conservation of energy and momentum tell us:

$$q = k - k', \quad P' = q + p. \tag{25}$$

Note that when writing $a = b$ for 4-vectors, each equation implies four equations, one for time and three for the spatial components.

There is a set of commonly used invariant quantities that we use to describe the scattering kinematics:

- The center-of-mass energy $\sqrt{s}$ with $s \equiv (k + P)^2$;
- The four-momentum squared (in fact, the negative of it): $Q^2 \equiv -q^2$;
- The Bjorken scaling variable: $x \equiv \frac{Q^2}{2P\cdot q}$;
- The invariant mass of the $\gamma^* + p$ system: $W^2 \equiv P'^2$;
- The inelasticity: $y \equiv \frac{P\cdot q}{P\cdot k}$;
- The energy transfer from the e^- to the proton in the proton-rest frame: $\nu \equiv \frac{P\cdot q}{M_p}$;
- The Nachtmann variable: $\xi \equiv \frac{2x}{1+\sqrt{1+4x^2M_p^2/Q^2}}$. Note that $\xi \approx x$ for large Q^2.

Exercise 1 For electron scattering off a fixed proton target (of mass M_p), given initial electron's energy E, the scattered electron's energy E' and the scattering angle θ, express P in terms of M_p, and show that $Q^2 = 2EE'(1-\cos\theta)$, $\nu = E - E'$, $y = \nu/E$, $x = \frac{Q^2}{2M_p\nu}$, $W^2 = M_p^2 + 2M_p\nu - Q^2$, and $s = M_p^2 + 2M_pE$.

Exercise 2 Show that for the general case (collider or fixed-target scattering), $x = \frac{Q^2}{Q^2+W^2-M_p^2}$, $Q^2 = (s - M_p^2)xy$.

Exercise 3 For scattering of an 11 GeV beam off a fixed proton target (of mass $M_p = 0.938272$ GeV), at a scattering angle $\theta = 30°$, calculate the maximal x value one can reach under the constraint $W > 2$ GeV.

Exercise 4 Show that for the general case of *ep* scattering (collider or fixed-target), one has $W \geq M_p$, $0 \leq x \leq 1$, and $0 \leq y \leq 1$.

3 Introduction to Electron Scattering

In this section we introduce electron scattering as a tool to study the Standard Model of particle physics and the structure of the nucleon. Prior knowledge on the Standard Model is assumed, including 3 generations of quarks and leptons and the force mediator bosons. These can be found in typical modern physics textbooks.

3.1 The Subatomic World

To look into matter, you can buy a $100 microscope. A microscope utilizes optical light and thus its resolution is limited by the wavelength of the light, and so we can not use them to study atomic structure of matter (10^{-10} m). For smaller length scales, one uses energetic particles such as electrons. In the language of modern physics, we say that the resolution of such an electron beam is defined by its de Broglie wavelength $\lambda = \hbar/p$ where p is the (relativistic) momentum and $\hbar c = 197$ MeV·fm. Thus, a room-size electron microscope with 10^1 keV energy is good enough to study DNAs (recall Spider Man). To "look into the proton" (with a radius of roughly 0.8 fm), one needs an electron beam of at least several GeV energy such as that provided by the underground Continuous Electron Beam Accelerator Facility (CEBAF) at Jefferson Lab. A more strict definition of the resolution is based on the Q^2 and thus is process (kinematics) dependent.

You may have learned that protons and neutrons are made of three quarks: $p = uud$ and $n = udd$. In the simplest picture, they are. But a complete picture is that the nucleon is made of these three *valence quarks* that carry the quantum numbers of the nucleon (electric charge, weak charge...); a massive amount of *gluons*, the force mediator of the strong interaction; and a *sea of quark-antiquark pairs* that pop up from the gluons. One of the fundamental goals of nuclear and particle physics study is to learn what matter is made of, *i.e.* to study the structure of the nucleon in terms of its constituents and the interactions within. We will introduce concepts such as "form factors", "structure functions" and "parton distribution functions (PDF)". More recent developments on this subject include also "generalized PDFs" and "transverse momentum distributions (TMD)", to name a few.

Another concept we utilize is the spin $\vec{s}$. All elementary particles carry spin (except for the Higgs boson which carries $s = 0$). We also define helicity: $h \equiv \frac{\vec{s}\cdot\vec{p}}{|\vec{s}||\vec{p}|}$ where $\vec{p}$ is the 3-momentum vector. If $\vec{s} \uparrow\uparrow \vec{p}$ (parallel), $h = +1$ and we call such particles "right-handed". If $\vec{s} \uparrow\downarrow \vec{p}$ (antiparallel), $h = -1$ and we call them "left-handed". For both cases, if you align your thumb along $\vec{p}$, your four fingers would curl along $\vec{s}$.

Units:

- For size or length: fm or Fermi; 1 fm (1 Fermi)= 10^{-15} m;
- For mass or momentum or energy: GeV; $m_e = 0.511$ MeV, $m_p = 938.272$ GeV (recall $\hbar = c = 1$ in Natural Units);
- Cross section: barn or b: 1 barn= 10^{-28} m^2=10^{-24} cm^2; and pb and nb.
- Luminosity: cm^{-2}s^{-1} or cm^{-2} for integrated luminosity.

3.2 Inclusive Scattering Cross Sections

In electron scattering experiments, we scatter an electron beam off a fixed proton or nuclear target. In collider settings, we let an electron beam and a proton (or ion) beam collide head-on. In either case, we can focus on "inclusive" scattering which means only the scattered electrons are detected. The primary observable we measure in such experiments is called the "scattering cross section" σ, or a differential quantity, $\frac{d^2\sigma}{\mathrm{d}E'\mathrm{d}\Omega}$ or $\frac{d^2\sigma}{\mathrm{d}x\mathrm{d}Q^2}$ or $\frac{d^2\sigma}{\mathrm{d}x\mathrm{d}y}$ because the cross section depends on the kinematics. The scattering cross section has the dimension $[\mathrm{L}^2]$ and describes the probability for the scattering to happen [8]. The differential cross section thus describe the probability for the scattering to happen within a certain kinematics range $(E', E' + \mathrm{d}E')$ and $(\theta, \theta + \mathrm{d}\theta)$ [or $(x, x + \mathrm{d}x)$ and $(y, y + \mathrm{d}y)$].

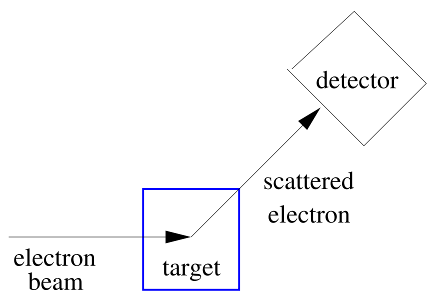

Figure 2: Detection of scattered electrons in the fixed-target setting (simplified).

In order to measure the angle and the momentum of the scattered electrons, we put a detector, often paired with a magnet (a spectrometer) to select the particle momentum, at a certain scattering angle θ, see Fig 2. The event rate (how many scattered electrons enter the detector per unit time) can be written, in the simplest case, as:

$$\frac{dN}{dt} = \frac{I_{beam}}{|e|}(\rho_N t) \times \Delta\Omega\Delta E' \times \frac{d^2\sigma}{\mathrm{d}E'\mathrm{d}\Omega} , \tag{26}$$

where I_{beam} is the electron beam current in Ampere (typically 1-100 μA at JLab), $|e|$ is the magnitude of the electron charge; ρ_N is the number density of the target and t the thickness; $\Delta\Omega\Delta E'$ is the product of the solid angle opening and the momentum acceptance of the spectrometer. Strictly speaking, Eq. (26) can be used only if $\Delta\Omega\Delta E'$ is small, or else one must replace the RHS by an integral $\int \frac{d^2\sigma}{\mathrm{d}E'\mathrm{d}\Omega}\mathrm{d}\Omega\mathrm{d}E'$.

The color on the RHS of Eq. (26) carries certain meanings: the term in blue is called the luminosity and is limited by the facility; the term in green is called the acceptance and is determined by the spectrometer (detector); the term in red is what we sought after, the physics. The precision of the measurement is determined by the statistical and the systematic uncertainties. The relative statistical uncertainty is purely determined by the event count – N, the product of the event rate and the experiment run time – as: $(\frac{\mathrm{d}\sigma}{\sigma})_{\mathrm{stat}} = 1/\sqrt{N}$. Since we have no control of the physics cross section (Nature controls

[8] Imagine you shoot a dart at a target. The bigger the target is, the more chance you hit it.

this term), for high precision measurements we must maximize the luminosity and the spectrometer acceptance. Figure 3 shows the luminosity of all existing lepton-proton scattering facilities and includes a few planned future facilities or upgrades.

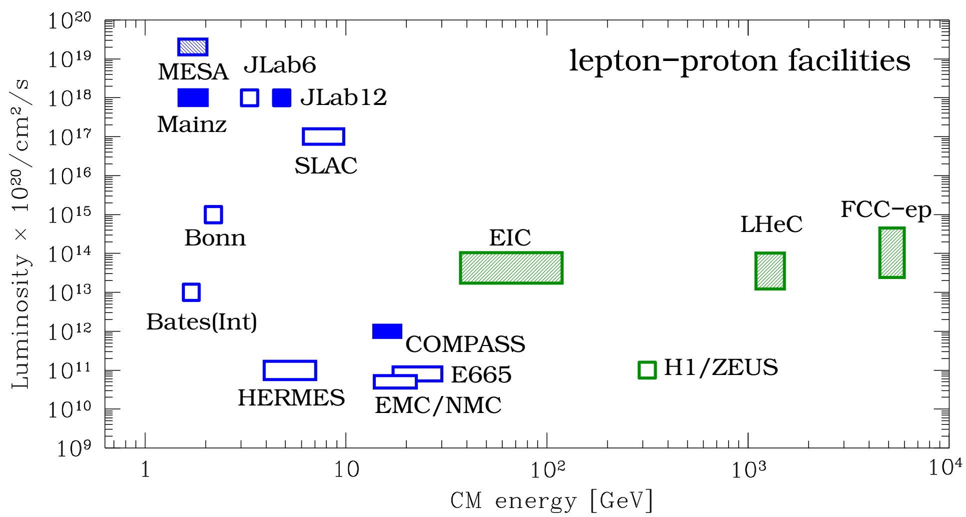

Figure 3: Luminosity vs. center-of-mass energy $\sqrt{s}$ for all fixed-target (blue) and collider (green) lepton-proton scattering facilities. Open rectangles indicate facilities that are retired (closed down) or superseded by upgrades, solid rectangles are facilities currently running, and shaded (hatched) rectangles are planned future facilities or upgrades.

Exercise 5 To convert between differential cross sections, one uses the Jacobian, for example: $\frac{d^2\sigma}{\mathrm{d}\Omega \mathrm{d}E'} = \mathcal{J}\frac{d^2\sigma}{\mathrm{d}x\mathrm{d}Q^2}$, where

$$\mathcal{J} = \det \begin{vmatrix} \frac{\mathrm{d}x}{\mathrm{d}E'} & \frac{\mathrm{d}x}{\mathrm{d}\Omega} \\ \frac{\mathrm{d}Q^2}{\mathrm{d}E'} & \frac{\mathrm{d}Q^2}{\mathrm{d}\Omega} \end{vmatrix}$$

Show that $\frac{d^2\sigma}{\mathrm{d}x\mathrm{d}Q^2} = \frac{\pi\nu}{xEE'}\frac{d^2\sigma}{dE'\mathrm{d}\Omega}$, $\frac{d^2\sigma}{dE'\mathrm{d}\Omega} = \frac{EE'}{\pi}\frac{d^2\sigma}{\mathrm{d}\nu\mathrm{d}Q^2}$, and $\frac{d^2\sigma}{\mathrm{d}\nu\mathrm{d}Q^2} = \frac{2Mx^2}{Q^2}\frac{d^2\sigma}{\mathrm{d}x\mathrm{d}Q^2}$. (You may use results from Exercise 1).

Exercise 6 Calculate the luminosity for a 50 μA 11 GeV electron beam incident on a 40-cm long liquid deuterium target. [9] Express your answer in cm^{-2}s^{-1}.

Exercise 7 The integrated luminosity of the 15 years running of HERA (the only $e^{\pm}p$ collider so far) is 0.5 fb^{-1}. Calculate how long it takes for a JLab beam of luminosity of Exercise 6 to reach 0.5 fb^{-1}. Express your answer using an appropriate unit of time. [10]

[9] You can find density and other commonly used physical properties of matter in Chap.6 of PDG.

[10] From this example one can see that collider and fixed-target experiments have very different strengths and weaknesses, and thus are complimentary to each other.

3.3 Structure Functions

3.3.1 The Three Kinematic Regime

The measured cross section for electron scattering off a nuclear target is illustrated in Fig. 4. It shows that the scattering can be divided into three types in which different types of physics are probed.

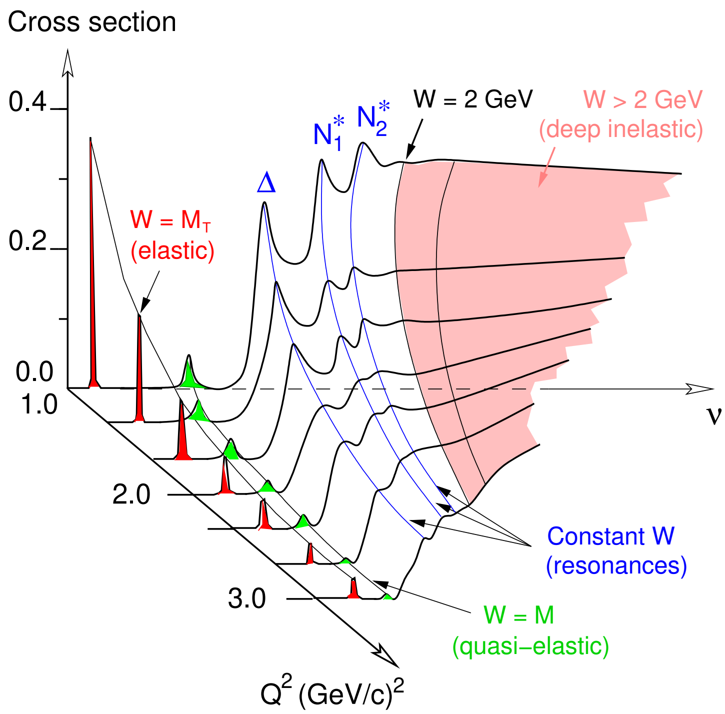

Figure 4: Scattering cross section vs. ν and Q^2 for nuclear targets.

Elastic and quasi-elastic scattering:

The red (narrow and sharp) peaks correspond to elastic scattering, where the nuclear target remains intact. Similarly, the electron can elastically scatter off individual protons and neutrons inside the nucleus. This is called quasi-elastic scattering (green peaks). For both cases the cross sections look like smeared delta-functions. [11] The elastic scattering cross section of relativistic electrons with initial energy E from a proton or a neutron with mass M is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\alpha^2\cos^2(\theta/2)}{4E^2\sin^4(\theta/2)}\frac{E'}{E}\left(\frac{G_E^2+\tau G_M^2}{1+\tau}+2\tau G_M^2\tan^2\frac{\theta}{2}\right), \tag{27}$$

where E' is the energy of the scattered electron for the given scattering angle (implied in $\mathrm{d}\Omega$) and $\tau \equiv 4Q^2/M^2$. The term on the RHS outside the parentheses corresponds to scattering off a point-like target, called the Mott cross section, and is calculated in

[11] Recall in classical two-body elastic scattering where one object is moving and collides with another object initially at rest. If the direction of the moving object after scattering is given, then its speed is fully determined, i.e. E' is a delta function for fixed values of θ.

quantum electrodynamics (QED). The terms within the parentheses then represent the non-point-like nature of the target, where $G^p_{E,M}$ (or $G^n_{E,M}$) are called the proton (neutron) electric and magnetic form factors.

One of the very first experimental evidence that the proton is not a point-like particle is the elastic scattering experiment carried out by Hofstadter *et al.* in the 1950's. Hofstadter was awarded the 1961 Nobel Prize in Physics "for his pioneering studies of electron scattering in atomic nuclei and for his thereby achieved discoveries concerning the structure of the nucleons".

Nucleon resonances:

The blue (wide) peaks are called nucleon resonances. In this case, the energy transferred to the nucleon start to excite the quark states (inside the nucleon), forming resonances. One identifies the resonances by calculating the invariant mass W. Nucleon resonances can be roughly divided into the $\Delta(1232)$, and the 2nd and the 3rd resonance regions.

Deep Inelastic Scattering:

The region with $W > 2$ GeV is called deep inelastic scattering (DIS), which has played and is still playing a centrol role in our understanding of the subatomic structure. The cross section considering electromagnetic interaction alone can be written as

$$\frac{d^2\sigma}{\mathrm{d}x\mathrm{d}Q^2} = \left(\frac{4\pi\alpha^2}{Q^4}\right)\left[y^2F_1(x,Q^2) + \left(\frac{1-y}{x} - \frac{My}{2E}\right)F_2(x,Q^2)\right] , \tag{28}$$

where $F_{1,2}$ are called **Structure Functions** (SF) of the target. We will focus primarily on DIS this summer. For a review and all type of SFs involved in electron or lepton scattering, including both electromagnetic and weak interactions, see Chap.18 of PDG. We will briefly review two significant facts learned from SF data next.

3.3.2 Scaling and the Parton Model

The first DIS data from SLAC in 1968 showed that approximately, $F_2 = 2xF_1$ and that the SF values are independent of Q^2, i.e. they are only functions of x. Shortly after scaling was observed, Feynman proposed the parton model:

- Protons and neutrons are made of point-like, spin 1/2 partons (these were later identified as quarks);
- In DIS, the electron (or the exchange photon) scatters elastically off individual quarks, and x represents the fraction of the nucleon's momentum carried by the struck quark in the infinite momentum frame;
- The quarks observed in DIS are only weakly interacting, *i.e.* they are "quasi-free".

The scaling phenomenon (called "Bjorken scaling") provided the first evidence of the existence of quarks. More specifically, the fact that $F_2 \approx 2xF_1$ proved quarks have spin quantum number $s = 1/2$, and scaling proved that quarks are point-like. [12] The 1990 Nobel Prize in Physics was awarded to Friedman, Kendall and Taylor "for their pioneering investigations concerning deep inelastic scattering of electrons on protons and bound neutrons, which have been of essential importance for the development of the quark model in particle physics". Existing world data on F_2 of the proton and the deuteron are shown in Fig. 5, where scaling is evident for the majority of the data coverage.

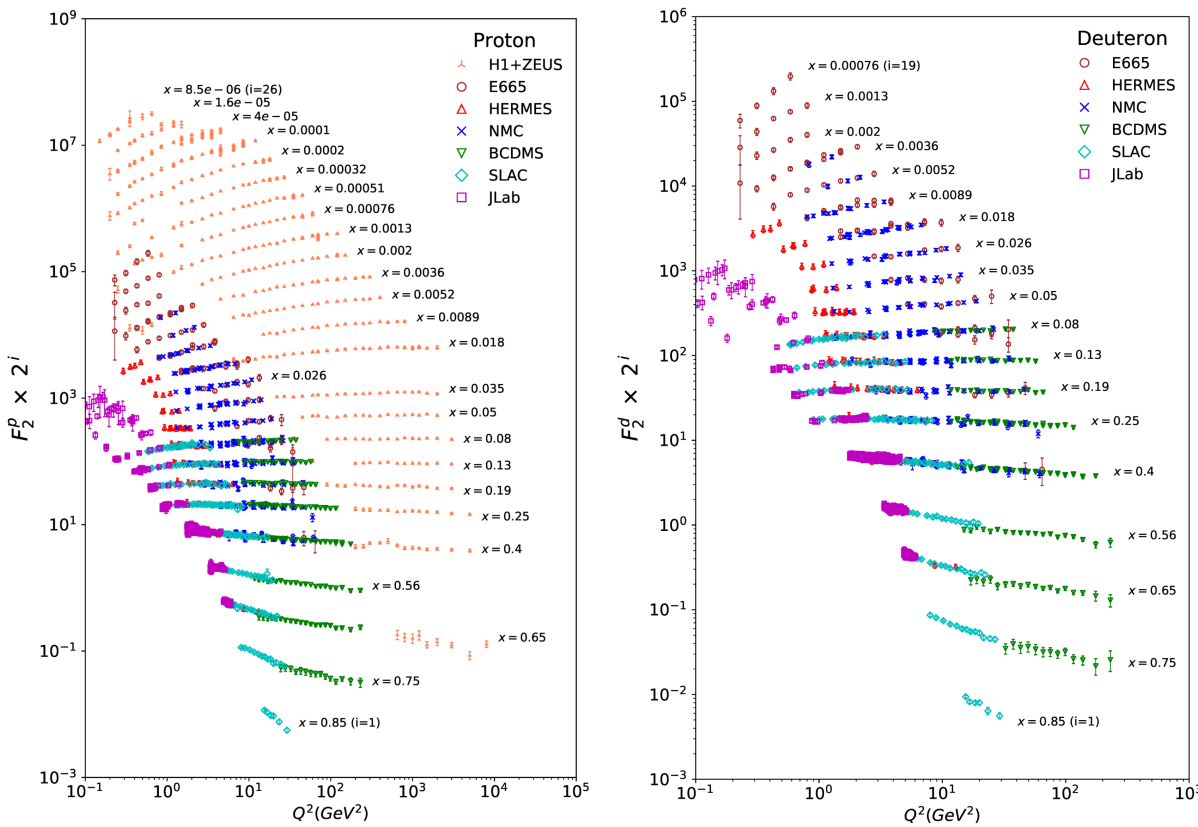

Figure 5: From PDG (2020 Figs. 18.8 and 18.9): Structure function data F_2 for the proton (left) and the deuteron (right). For clarity, data in each x bin are multiplied (offset vertically) by 2^i.

3.3.3 Asymptotic Freendom and Establishment of QCD

Immediately after the parton model was proposed, the Q^2 dependence of the strong interaction coupling constant α_S [13] was calculated in quantum chromodynamics (QCD)

[12] "Scaling" means what you see is regardless of the resolution of the probe used. Recall Q^2 represents the resolution of the electron probe, and thus a lack of Q^2 dependence in the SF data means scaling. When your observation of some objects is independent of the probe resolution, it implies you are seeing point-like objects.

[13] The analogy of α_S is the fine structure constant $\alpha \approx 1/137$, which represents the coupling strength of electromagnetic interaction.

and was found to be

$$\alpha_S = \frac{4\pi}{(11 - \frac{2n_f}{3})\ln(\frac{Q^2}{\Lambda^2})} , \tag{29}$$

where $n_f = 6$ is the number of quark flavors and Λ is the QCD mass scale. Thus at large Q^2 as probed in DIS experiments, α_S becomes very small and quarks can be considered as "quasi-free". This result led to one of the strongest evidence that QCD is the theory of strong interactions. The 1999 Nobel Prize in Physics was awarded to Gross, Politzer and Wilczek "for the discovery of asymptotic freedom in the theory of the strong interaction".

In addition to scaling, Fig. 5 shows that at very low and very high Q^2 values, the structure functions start to exhibit $\log(Q^2)$ dependence. This is because quarks do interact with each other through strong interaction and such interaction is reflected in a slight Q^2-dependence in the SFs. In QCD, such Q^2 evolution can be calculated precisely using the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equations. This means that given the SF at a certain Q^2 value, you can "evolve" the SF to a different Q^2 region. The perfect agreement between the calculated DGLAP evoluation and the observed Q^2 dependence in the SF data provide strong evidence that QCD works well, at least in the perturbative regime.

The Q^2 evolution applies not only to structure functions, but also to PDFs (next section), and interaction coupling constants. You have already seen one example of such Q^2 evolution of the strong coupling constant in Eq. (29). Many other quantities of the Standard Model, such as the fine structure constant α, the weak mixing angle θ_W, also evolves with Q^2. In fact, measuring the Q^2 dependence of $\sin^2\theta_W$ and comparing with the Standard Model (SM) prediction is being used as a tool to test the SM and to search for hints of beyond the standard model (BSM) physics, which some of you might encounter this summer.

3.4 Parton Distribution Functions

We now dive deeper and ask the question: Exactly what do these structure function data tell us about the structure of the nucleon? Recall that DIS is (the sum of) the electron elastically scattering off individual quarks, and elastic scattering off a point-like particle can be calculated precisely in QED. This means that from SF data we can deduce **the probability to find a certain quark inside the proton that carries x-fraction of the nucleon's momentum (in the infinite momentum frame)**. This probability is called **Parton Distribution Functions**, denoted as $q(x, Q^2)$ with $q = u, d, c, s, t, b$. In the parton model, one simply has

$$F_1(x, Q^2) = \frac{1}{2}\sum_q (e_q)^2 q(x, Q^2) , \tag{30}$$

where $e_q = \frac{2}{3}$ for u, c, t or $-\frac{1}{3}$ for d, s, b are the electric charge of the quark in unit of $|e|$, representing the strength of the electromagnetic interaction of that quark; and the summation is over quark flavors. For JLab energies, one only needs consider u, d, s, c, while for high energy colliders one must sum over more flavors.

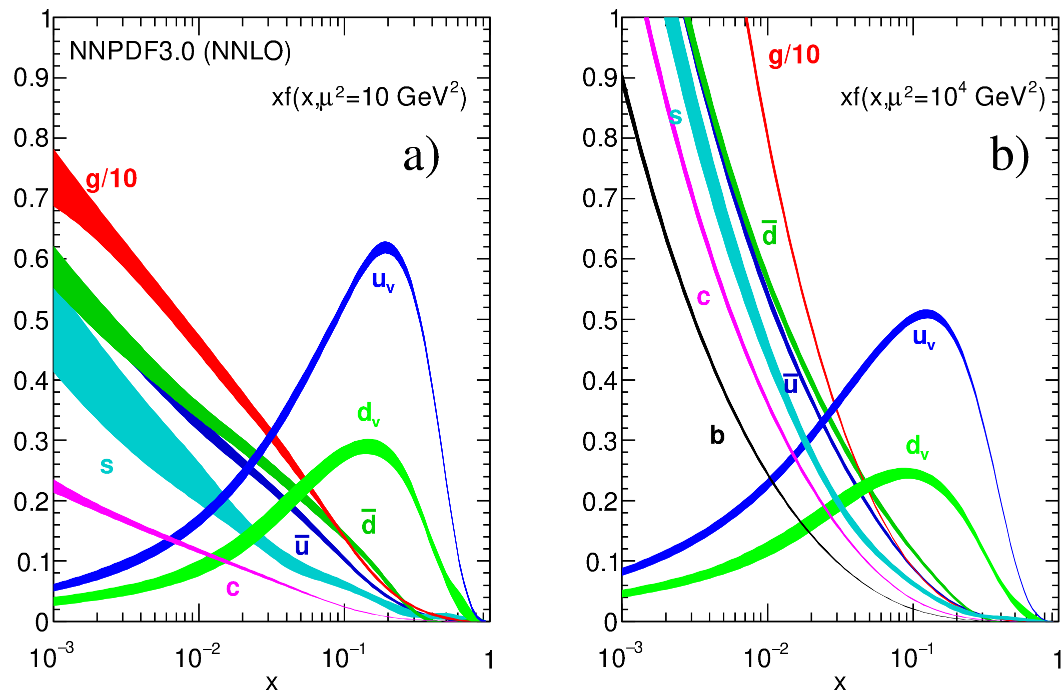

Figure 6: From PDG (2020 Fig. 18.4) The bands are x times the unpolarized PDF $f(x)$ (where $f = u_v, d_v, u, d$ and $s \approx \bar{s}$, $c = \bar{c}$) obtained in NNLO NNPDF3.0 global analysis at scales $Q^2 = 10$ GeV^2 (left) and $Q^2 = 10^4$ GeV^2 (right), with $\alpha(M_Z) = 0.118$.

From Eq. (30) one might assume that we can extract all PDFs if we collect enough data on DIS inclusive cross sections. The reality is more complicated because one has six quark flavors to separate but only two nucleons (protons and neutron) are provided by Nature. This is like solving for six unknowns using two equations. Therefore we must combine data from different processes: DIS using electrons and muons, DIS using neutrinos (with different quark couplings from electrons), Semi-inclusive DIS (where part of the hadronic final states are detected), and Drell-Yan processes ($q - \bar{q}$ annihilation in proton-proton collisions), to name a few. World data from different processes are collected and fit globally to achieve a specific parameterization of PDFs. Prominent groups in such global PDF fits are CT, MMHT, CJ (CTEQ-JLab), JAM (Jefferson Lab Angular Momentum), NNPDF (Neural Network), and more groups are leading the effort of spin-polarizaed PDFs which we will not cover in this short course. Fortunately, these parameterizations are accessible via the LHAPDF interface so one does not need to deal with 10 different codes (or programming languages). A representative plot of PDF is shown in Fig. 6, calculated using the NNPDF3.0 PDF grids. The sea quarks are denoted $\bar{u}, \bar{d}, \bar{s}, \bar{c} \ldots$, and the valence quark distributions are defined as $u_V \equiv u - \bar{u}$, $d_V \equiv d - \bar{d}$, where the subtraction is for the sea quark produced in the $q - \bar{q}$ pair (assuming to be equal to $\bar{q}$, i.e. assuming symmetric sea.)

A few words about what you need to study/know when using PDF grids:

- Data from what type of processes (electron DIS, muon DIS, neutrino DIS, Drell-

Yan, SIDIS) were included in the fit? What targets (proton, deuteron, heavy nuclei) were used? And what (W, Q^2) cuts were applied?

- What assumptions were used when performing the parameterization? For example, was it assumed that $s = \bar{s}$, $c = \bar{c}$, etc ("symmetric sea")? – unfortunately, even with the plethora of data from different processes, one cannot determine all PDFs for certain and assumptions such as symmetric sea and others are often used.

- Does the set provide both a set of central values and the uncertainties? How should one evaluate the uncertainty? (Different groups deal with the uncertainties differently, and this is not "unified" in LHAPDF. One must dig into the original publication to figure out how to calculate the uncertainties.)

Fitting what together?
As mentioned earlier, the F_1 given in Eq. (30) is only one of the many type of SFs that we can measure and use to extract PDFs. The interaction couplings on the RHS of Eq. (30) are the electric charge of quarks, which are well known. Other SFs accessible in electroweak interactions of electron scattering are products of PDFs and electroweak couplings of electron-quark interactions. If data are available for a wide kinematic range and to a high precision, then it is possible to *fit both the couplings and the PDFs together*. If the measured couplings deviate from their SM predictions, they may give hints on where to search for BSM physics, such as particles or interactions that are not currently in the SM.

3.5 Cross Section Asymmetries

While scattering cross sections are the primary observable, in some experiments we instead aim at measuring differences in cross sections, *i.e. cross section asymmetries*. The statistical uncertainty of the asymmetry measurement is

$$(\Delta A)_{\rm stat} \;=\; 1/\sqrt{N}, \tag{31}$$

where N is the total event count if it's approximately evenly split between the two scattering conditions. Two most commonly measured types of asymmetries are explained briefly in this section.

3.5.1 Double Spin Asymmetries

In scattering experiments using both a polarized beam and a polarized target, one can measure double-spin asymmetries by flipping the beam spins back and forth:

$$A_{\parallel} \;=\; \frac{\sigma^{\uparrow\Uparrow} - \sigma^{\downarrow\Uparrow}}{\sigma^{\uparrow\Uparrow} + \sigma^{\downarrow\Uparrow}}\;, \tag{32}$$

where $\uparrow$ ($\downarrow$) represents the beam spin parallel (anti-parallel) to the beam direction, and $\Uparrow$ represents the target spin always pointing parallel to the beam direction. Double-spin asymmetries are used to extract polarized SFs $g_1(x, Q^2)$ and $g_2(x, Q^2)$, and furthermore polarized PDFs $\Delta q(x, Q^2)$ and the spin structure of the nucleon.

3.5.2 Parity-Violating Asymmetries

If the target is unpolarized but we flip the beam spin back and forth, there is a cross section difference that arises from the interference between photon and Z-boson exchanges, *i.e.* due to the parity-violation (PV) nature of weak interactions. In such PV electron-scattering (PVES) experiments, one measures:

$$A_{PV} = \frac{\sigma^R - \sigma^L}{\sigma^R + \sigma^L} , \tag{33}$$

where $R(L)$ refers to right-handed (left-handed) incident electrons. PVES has been used to access the weak charge of the proton (measured via elastic scattering off a proton target), and the vector-axial coupling between electrons and quarks (via PV in DIS).

Exercise 8 Download and install LHAPDF or use a pre-built library [14], calculate PDF values for all quark flavors u, d, c, s, b and valence quark PDFs u_V, d_V, and their uncertainties (if applicable) for $x = 0.3$, $Q^2 = 5.0$ GeV^2 using the CT18 PDF grid or a PDF grid of your choice.

Exercise 9 Download and install LHAPDF or use a pre-built library, make a plot that is similar to Fig. 6 but using a different PDF set, such as CT18 or MMHT2014.

Exercise 10 Given $x = (0.3, 0.31)$ and $Q^2 = (5.0, 5.1)$ GeV^2 ranges and Eqs. (28) and (30), calculate the integrated cross section $\int \frac{d^2\sigma}{\mathrm{d}x\mathrm{d}Q^2}\mathrm{d}x\mathrm{d}Q^2$. You may use results from Exercise 7 and you may assume the cross section is approximately flat (does not depend on x or Q^2) within the given range.

Exercise 11 Using your result from Exercise 9 and the running condition (beam energy and luminosity) from Exercise 6, estimate how many days of running are needed to reach an uncertainty of 1% on the cross section within the given $x = (0.3, 0.31)$ and $Q^2 = (5.0, 5.1)$ GeV^2 bin? If the expected PV asymmetry is 400 ppm (1 ppm $= 10^{-6}$), how many days of running are required to reach a relative uncertainty of 1% on the asymmetry? You may assume the spectrometer acceptance is 1 for the given kinematic range, *i.e.* all scattered electrons within this given bin are detected successfully.

4 Particle Detectors

4.1 Particle Interaction with Matter

When particles pass through matter, they interact with electrons and nuclei and lose energy. How they lose energy depends nearly solely on two factors: the particle mass (whether it's electron's mass or something heavier) and their electric charges. Many details of particle passage through matter can be found in PDG Chap. 34. We introduce some concepts briefly below [15]:

[14] If you have a JLab computing account, try copy the short code from /work/eic/users/xiaochao/summer/

[15] Particles of lower energy (MeV or keV) lose energy differently from what is described here. Those interested in medical applications of radiation (gamma, protons, neutrons, electrons, α) should refer to appropriate textbooks of application of nuclear physics.

- Electrons ionize atoms and also radiate photons under the nuclear Coulomb force. For electrons of GeV or higher energy, they lose energy primarily due to the second effect, also called Bremsstrahlung radiation. Bremsstrahlung photons create e^+e^- pairs which are of high energy themselves and trigger more Bremsstrahlung. The process repeats and forms an electromagnetic shower until the e^+e^- produced are below a "critical energy" value and they lose energy more through ionization than Bremsstrahlung and eventually all particles stop. By detecting and stopping all particles in such EM shower, we can detect the energy of the incident electrons in experiments; The material thickness for incident electrons to lose $1/e$ of their initial energy is called "radiation length" X_0. Total absorption calorimeters are typically of thickness $16X_0$ or longer.

- Charged hadrons (lightest are the $\pi^\pm$ of mass 0.135 GeV/c^2) and heavy charged leptons (μ^-, τ^-) lose energy primarily by ionizing (kicking off atomic electrons) and the incident particles continue moving forward [16]. Hadrons of high energy also trigger hadronic showers by producing pions through nuclear interaction, among which the π^0 would decay into photons and can be detected. Hadronic showers exhibit a similar multiplicative behavior as EM showers, though they develop a lot slower in matter and thus hadrons are much harder to "stop". The material thickness to characterize the energy loss of hadrons is called "nuclear interaction length" λ_I. λ_I is much longer than X_0 for all medium to high Z material. Hadronic calorimeters used in high energy physics experiments often cannot be "total absorption" because of the limited space available for detectors for collider settings.

- Photons do not ionize, but they can produce $e^\pm$ pairs under nuclear Coulomb field and these subsequent $e^\pm$ can trigger EM shower. Thus one can use EM calorimeter to measure the energy of photons. In common applications, the best way to shield high energy photons is to use high Z material such as lead.

- Neutrons are the most difficult to lose energy when passing through matter because they are charge neutral (no ionization, no interaction with the nuclear Coulomb field). Neutrons trigger nuclear reaction, and/or they scatter elastically from protons and other atomic nuclei in the matter and lose energy. This is why water is a commonly used neutron moderator material (to slow down neutrons) in fission reactors. Similarly, the best way to shield neutrons is to use water or plastic (high hydrogen content to allow maximum energy transfer per collision). Neutrons can cause more radiation damage than gamma or charged particles of the same flux and are a concern for electron scattering experiments using nuclear targets (anything heavier than hydrogen). [17]

Besides X_0 and λ_I, the energy loss of particles passing through matter is described by the "stopping power" $\mathrm{d}E/\mathrm{d}x$. You can find graphs in PDG for quick look up of, for

[16] Recall in classical mechanics: if a moving object of mass m_1 collides head-on with another object of equal mass m_2 and if $m_1 \gg m_2$ then the motion of m_1 is nearly unchanged while m_2 is kicked to a high speed.

[17] The Iron Man movie obviously utilized some nuclear physics knowledge, portraying his arc-energy-heart as a neutron radiator.

example, how much energy cosmic muons lose when passing through certain thickness of scintillators. See Exercise below.

Exercise 12 Use PDG Fig. 34.1, estimate the energy loss of 1-10 GeV muons in 2 g/cm^2 thick of material.

4.2 Magnetic Spectrometers

Almost all particle detection requires some sort of magnetic field such that particles of interest are separated from charge-neutral and low energy background. Such magnetic fields select particles of certain (relativistic) momenta: Particles with desired momenta are bent by design and enter particle detectors that are behind the magnet. The bending angle of the particle is determined by the particle's momentum, its electric charge, and the field strength $\int \vec{B} \cdot d\vec{l}$ (called "B-d-l"). One selects the desired particle momentum by changing the field strength, often provided by superconducting magnets, and you may want to keep your fingers crossed that these magnets don't "quench" when you click the button!

There are many different types of magnetic fields used by particle spectrometers. At JLab's Halls A and C, each spectrometer utilizes a combination of dipole (for bending) and quadrupole magnets (for focusing). Charged particles traverse their fields like light rays bent by optical lenses, and one often talks about the "optics" of spectrometers. The CLAS(12) spectrometer in Hall B has a toroidal field, and the newly designed SoLID spectrometer in Hall A has a solenoid field (like an MRI magnet).

Exercise 13 Design a magnet that bends charged particles of relativistic momentum 5 GeV/c upward (above horizontal) by 45 degrees. Estimate how strong the field needs to be in order to keep the size of the magnet manageable? (That is, for the magnet to be no more than 1 or 2 stories tall).

4.3 Scintillators

The first step to detect particles is to form a trigger signal to inform the data acquisition (DAQ) that there is a particle coming in. Trigger signals are often provided by thin scintillators. Charged particles (both electrons and heavier ones) excite the molecular states of scintillating material, and de-excitation of such states emits scintillating light that can be detected by photomultiplier tubes (PMT). Scintillators typically have very fast timing response (to allow precise timing of the trigger), and are very thin (to not affect the particle momentum significantly) unless high precision timing is required (need high scintillating photon statistics). If one combines several scintillator paddles oriented in orthogonal directions, one forms a "scintillator hodoscope" to provide the approximate position of the particle, though more precise positioning info is provided by tracking detectors, coming up next.

4.4 Tracking Detectors

The second step is to determine the trajectory of the particle. Commonly used tracking detectors are drift chambers (DC), multi-wire proportional chambers (MWPC), gas emission multipliers (GEM), to name a few. Recall those E&M problem where you solved the field of a single wire held at $+V_0$ and found $E(s)$ to be extremely large for small s (distance from the wire)? Typical wire chambers work exactly the same way: they consist of many wires held at high voltages and are constantly flushed by noble gas (to avoid discharge). When charged particles pass through the wire chamber, they ionize the gas inside. Ionized electrons float to the wire and trigger signals in a handful of closest wires. By analyzing these wire signals one can deduce the hit position of the particle on the wire chamber. By combining two or more such chambers, one can detect the particle trajectory (two position readings x, y and two angles). Tracking information is then combined with the spectrometer optics to fully determine the particle momentum, the scattering angle, and the position of where the scattering occurs within the target. Tracking information is also often combined with the hit position of subsequent detectors to separate useful particles (strong correlation in hit positions of multiple detectors) from background events (no correlation).

The optics of the spectrometer can be measured using dedicated runs taken with carbon foil targets (to provide known reaction position along the beam) and with a "collimator" of many holes at the entrance of the spectrometer magnet (to provide known angle of the scattered particle). Optics data are analyzed to determine an optical "matrix" which transform scattering kinematics at the target to variables seen by the detectors.

Exercise 14 Exercise 14: One typical drawback of tracking detector is the large amount of wire signals to readout per event. Reading out such signals cause a "deadtime" effect, which means the system cannot react to another particle for a certain time window after the first particle hits. For detectors with a deadtime τ, the measured rate is $R_{\text{meas}} = R_{\text{true}}(1 - R_{\text{meas}}\tau)$ where R_{true} is the true event rate. Therefore a system with deadtime τ can detect only a maximum rate of $1/\tau$ regardless of how high the actual event rate is. Suppose you go take some experimental "shifts" in Hall A of JLab and you are told the DAQ can handle a maximum of 4 kHz event rate, estimate the deadtime of the system. (You can probably guess that the DAQ deadtime is limited by the vertical drift chambers. If your experiment does not need tracking information, then you can turn off the VDC and run at much higher rates.)

4.5 Time-of-Flight (TOF)

All detectors above tell us the trajectory, timing, and the momentum of the particles, but they do not tell us what they are (except for the sign of the electric charge they carry through magnetic field information). One typical method for particle identification (PID) is time-of-flight, see Exercise below.

Exercise 15 For two scintillator planes 1 meter apart, calculate the timing difference in the scintillator signals if a 1 GeV/c electron passes through and hit both along a direction normal to the planes. Repeat for π^- ($m = 0.135$ MeV/c^2). What timing "resolution" do you need to tell the two particles apart? (That is, the intrinsic width of the timing signal from the scintillator needs to be narrow enough for you to tell that there are two pulses, not one. You may assume the signal has the shape of a Gaussian peak, and its width in time is called the timing resolution of the detector signal.)

4.6 Gas Cherenkov Detectors

When charged particles pass through a material with speed $v > c/n$ with c the speed of light in vacuum and n the index of refraction of the material, they emit Cherenkov light. Because the spectrometer magnet selects particles based on their momentum, one can design a Cherenkov detector of specific n such that particles with lighter masses (such as electrons) emit Cherenkov light but those with heavier masses (such as pions) do not. Cherenkov detectors are thus a typical type of "threshold" PID detectors. The advantage over TOF is that one can tune the n value to meet experimental needs, while TOF is limited by the timing resolution and the momentum range of the particle. For very high energy, all particles fly fast and TOFs can no longer be used.

Exercise 16 For 4 GeV/c particles, calculate the range of n that allows electrons to emit light but not charged pions ($m = 0.135$ MeV/c^2). Similarly, calculate the range of n that allows charged pions to emit light but not heaver mesons such as kaons ($m = 0.497$ MeV/c^2). The former is often realized with "light gas" such as CO_2, and the latter can be realized using "heavy gas" or other more exotic materials.

4.7 Calorimeters

Calorimeters are used to detector the energy of particles. As mentioned above, calorimeters can be categorized as EM and hadron calorimeters. Because calorimeters stop the particles completely, they are the last detector along the particle path. Table 1 shows a simple comparison of the basic characteristics of the two types of calorimeters. For both cases, the energy resolution is determined by the fluctuation in the number of showers produced. Because hadronic showers are much slower to develop and they are not stochlastic in nature, the energy resolution of hadronic calorimeters are often at $50\%/\sqrt{E}$ level where E is the energy of the incident particle in GeV, and thus hadronic calorimeters are only used in high energy experiments.

EM calorimeters can also serve as a PID detector to identify electrons (lose all their energy in EM calorimeters and are stopped completed) from hadrons (lose a lot less energy than electrons in EM calorimeters). The amount of energy lost by hadrons when passing through EM calorimeters falls roughly within a minimum range of the $\mathrm{d}E/\mathrm{d}x$ curve (see PDF Fig. 34.1), called "minimum ionization peak" (MIP) energy, and the hadrons mostly punch through the EM calorimeter. In high energy experiments, one often builds additional detectors outside EM calorimeters to detect hadrons or muons that are not completely stopped.

Table 1: A simple comparison of EM vs. hadronic calorimeters.

Calorimeter	Electromagnetic	Hadronic
Shower nature	electrons produce Bremsstrahlung photons, photons produce e^+e^- pairs, the cycle repeats	nuclear interaction produces pions, $\pi^0 \to \gamma\gamma$, photons produce EM shower
Threshold energy for shower production	critical energy $E_c \approx 550$ MeV$/Z$	threshold energy $E_{\text{TH}} = 2m_\pi = 280$ MeV
Characteristic length for shower production	radiation length $X_0 = \frac{716\ \text{g/cm}^2 A}{Z(Z+1)\ln\left(\frac{287}{Z}\right)}$	nuclear interaction length $\lambda_I = 35A$ (g/cm^2)
Typical energy resolution	tunable, as good as $(3-5)\%/\sqrt{E}$	Cannot be better than $30\%/\sqrt{E}$

5 Useful Resources

As a start, I recommend the following textbooks and resources at the level of upper-undergraduate or beginning graduate studies:

- Quarks and Leptons, by Halzen & Martin
- Introduction to Elementary Particles, by Griffiths
- Modern Particle Physics, by Thomson
- Particle Data Group