1 Introduction

In this series of 5 lectures, we will study the jets in heavy ion collisions. The systems we would like to study are inevitably quantum systems since the entities we deal with are microscopic subatomic particles. In this lecture, I will emphasize not only the quantitative aspects of the jet studies, but also the qualitative aspects. This is because without qualitative understanding of a complicated calculation or a phenomena, no real understanding of phyiscs takes place even though the derived formula can describe nature.

Why do hard probes

Relativistic Heavy Ion Collisions

- Why do it?
  - To study QGP
  - Most extreme environment ever created: $T \sim 1$ GeV.
    This existed only at around 1 microsecond after the Big Bang
- How do we understand it?
  - Theory: Many-body QCD
  - Experimental probes:
    - Soft
    - Hard

Hard Probes are useful

- Hard Probes $\sim$ Large momentum/energy phenomena
- pQCD applies – We know how to do this
- Produced *before* QGP is formed in the same way as in hadron-hadron collisions
- Difference between *pp*, *pA* and *AA* tells us about the medium.
- Caveat: How well do we know the *nuclear initial state?*

What do we want to learn?

- Medium properties
  - What is it made of? Quarks? Gluons? Hadrons?
  - Thermodynamic properties – Temperature, Equation of state, etc.
  - Transport properties – Mean-free-path, transport coefficients, etc.
- Tools
  - Jets
  - Hard Photons

Outline

1. pQCD
2. Jet Quenching
3. Hard Photons

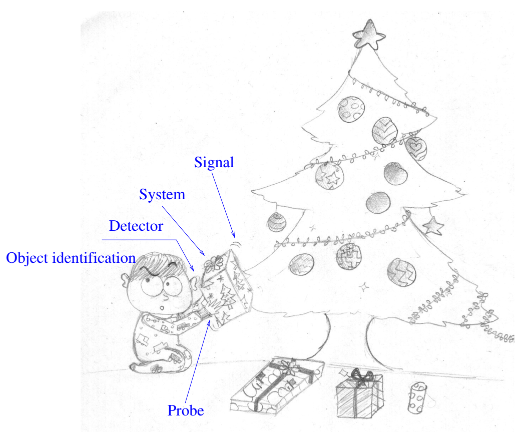
- My goal for these lectures: Qualitative understanding

What is a hard probe?

- Early hard probe experiments

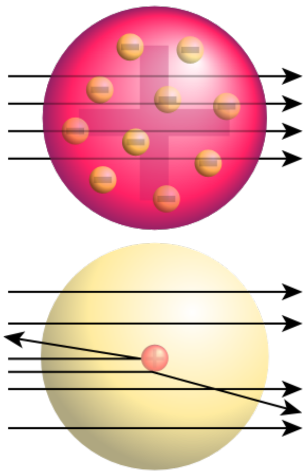
What is a hard probe?

- Early hard probe experiments



What is a hard probe?

- Early hard probe experiments

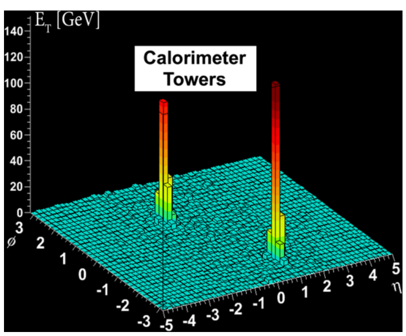


- Rutherford's α scattering experiment (1911)

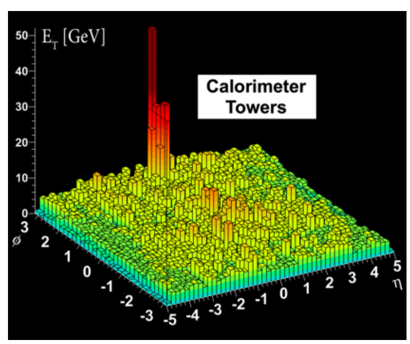
$$\frac{d\sigma}{d\cos\theta} = \frac{\pi}{2} Z^2 \alpha_{\rm EM}^2 \left(\frac{\hbar c}{E_{\rm kin}}\right)^2 \times \frac{1}{(1-\cos\theta)^2}$$

- Small angle scattering dominates $d\sigma/d\cos\theta \propto 1/\theta^4$
- But backscattering prob. is finite, favoring Rutherford's model over Thompson's (which causes no backscattering)

Fast-forward to the present



ATLAS: Intact dijets in Pb+Pb



ATLAS: One jet is fully quenched in Pb+Pb

- Simplest conclusion to draw: The medium is *opaque*.
- We want to know much more than that!

Hard Probe Requirements

- Must be known & calculable using pQCD.
- Must be created *before* QGP forms
- Both requirements satisfied if the energy scale is much large compared to $\Lambda_{\rm QCD} \approx 200\,\text{MeV}$ and the length (time) scale is much shorter than $\sim 1\,\text{fm}$.
- Example: Jets (high energy partons) with $E \gg 1\,\text{GeV}$ and Heavy quarks (c, b) with $M \gg 1\,\text{GeV}$

Hard Probes

Probes

- Propagation of hard partons or "Jets"
- Quarkonium suppression
- High p_T electromagnetic probes (real and virtual photons)

Goal

- To characterize *QGP*
- To characterize initial state (nPDF, CGC?)

Review of some basic concepts

Review of some basic concepts

- Basic unit:

$$\hbar c = 197.3\,\text{MeV}\cdot\text{fm} \approx 0.2\,\text{GeV}\cdot\text{fm}$$

- With $\hbar = c = 1$
- Units
  Mass: GeV/c^2
  Momentum: GeV/c
  Energy: GeV
  Length: $\hbar c/\text{GeV}$
- $200\,\text{MeV} \leftrightarrow 1/\text{fm}$
- $1\,\text{fm} \leftrightarrow 1/(200\,\text{MeV})$
- Thermal energy $k_B = 8.617 \times 10^{-5}\text{eVK}^{-1}$
  With $k_B = 1$,
  $1\,\text{eV} = 11,605\,\text{K}$ or $290\,\text{K} \approx \frac{1}{40}\,\text{eV}$

Jeon (McGill) | Hard Probes | Stony Brook 2013 | 14 / 33

The unit system we are going to use is what's often referred to as the "natural unit" where we set $\hbar = c = 1$. You must have heard about it and perhaps used it in your studies already. But what do these mean? That is, what does setting the presumably dimensionful quantity

$$\hbar = 1 \quad \text{and} \quad c = 1 \tag{1}$$

mean?

To figure that out, we need to see where $\hbar$ and c appears. The speed of light appears in Special Relativity as the ultimate speed limit. That is, no particle, nor information, can travel faster than c. Furthermore, it is *constant* across *all inertial frames*. That is, it does not matter where in the universe you measure it. As long as you are in an approximately inertial frame (The surface of the Earth is obvioulsy not an inertial frame, but close enough since the acceleration is relatively small), the value of c you measure is exactly the same. In the KMS unit system, it is *defined to be* exactly

$$c = 2.99792458 \times 10^8\,\text{m}/s \tag{2}$$

Now note the unit. It is measured in meter per second. Hence, knowing the speed of light, one can easily convert time to length and vice versa as in

$$\Delta z = c\Delta t \tag{3}$$

But what's so special about this relationship? I can always write down such a relationship with any constant speed v.

The difference of course is that due to the fact that nothing can move faster than the light, space and time actually mixes up

$$\begin{aligned} t' &= \gamma t - \gamma\beta z \\ z' &= \gamma z - \gamma\beta t \end{aligned} \tag{4}$$

where

$$\beta = v/c \tag{5}$$

is the speed of the moving frame and

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{6}$$

is the Lorentz γ factor. Why is this so special? Well, consider the usual space and the rotation that transforms (x, y) to (x', y') as

$$\begin{aligned} x' &= \cos\phi\, x + \sin\phi\, y \\ y' &= \cos\phi\, y - \sin\phi\, x \end{aligned} \tag{7}$$

Note that the new x' is a mixture of the old x and y. One can do this because both x and y are spatial distances. All you are doing here is to re-define which direction you would like to call "south-north" and which one "east-west". Physics should not change just because you've re-defined the axis. In particular, this transformation does not change the distance

$$x^2 + y^2 = x'^2 + y'^2 \tag{8}$$

Now take another look at this relationship

$$t' = \gamma t - \gamma\beta z \tag{9}$$

It says that the time measured in a moving frame is related not only to the time in the rest frame but also the *position* where it was measured.

This is very peculiar, but Einstein's bold conclusion was that under this transformation that mixes space and time, physics should be the same. In other words, space and time are not two independent quantities. They can mix up. Therefore they must represent coordinates of a *unified* single 4-dimensional space. If one defines

$$x^0 = ct \tag{10}$$

then all components of the 4-vector

$$x^\mu = (ct, x, y, z) \tag{11}$$

can be measured with the lengh unit. This, of course, is a prelude to the General Relativity where the geometry of this 4-D space reprereents gravity. In the subatomic world, the most convenient length unit is the femto meter or fermi (fm) which is roughly the size of a proton. Hence, setting $c = 1$ is tantamount to measuring time in fm. The Lorentz transformation of course leaves the proper time

$$\tau^2 = x_\mu x^\mu = x'_\mu x'^\mu \tag{12}$$

unchanged with $g_{\mu\nu} = \text{diag}(1, -1, -1, -1)$.

Similarly, Einstein's famous formula

$$E = mc^2 \tag{13}$$

relates energy and momentum. We of course know that mass times velocity squared represents kinetic energy. What Einstein is telling us, however, is that there is an intrinsic energy that is associated with the mass and one can actually convert energy into mass or mass into energy. That is, they are basically the same quantity and the fact that c is constant just reflects that fact.

What about the Planck constant $\hbar$? Does it have a similar meaning? Well, you probably saw the Planck constant first time in the relationship

$$p = \hbar k \tag{14}$$

or

$$E = \hbar\omega \tag{15}$$

where p and E are the mechanical momentum and the energy of a particle, and k and ω are the wavenumber and the frequency of a wave. You must have used these relationship to get through your quantum mechanics exams. But have you really appreciated them?

In classical physics, there are two ways of transmitting energy-momentum, equivalently, information. Suppose that the information you want to transmit is in the binary form. That is, 0 or 1. Then one way to convey the information to another person/detector is to use particles. For instance, you can make a particle with high mass represent 1 and a particle with low mass represent 0. Then shoot them in sequence, taking care that the velocity is the same. Another way to transmit the same information is to use wave, like your cell phone, where waveform can represent the binary information.

They are clearly different. In the case of the particles, particles physically move a long distance to carry the information. In the case of the wave (think of sound), each particle moves only within the mean free path of scatterings, yet the information can be carried to a long distance. What the Planck and Einstein relationships above assert is that in the subatomic world, they are one and the same. The wave number, which has the dimension of $1/[L]$, is the same as the momentum which has the dimension $[M][L]/[T]$. The frequency, which has the dimension of $1/[T]$, is the same as the energy which has the dimension $[M][L^2]/[T^2]$.

Hence, the Planck constant $\hbar$ relates the momentum to the length, and the energy to the frequency. I would like to emphasize how different the two quantities being related are. The wavenumber controls how many times the wave oscillates in a unit length. Why should that be the same as the momentum that can push things around? The frequency controls how many times the wave oscillates in a given unit time. Why should that be the same as energy which can heat things up, for instance? Yet, quantum mechanics says they are because ultimately we are all made of de Broglie wave.

Now recall that c relates the time and the length. Hence, in that unit, both momentum and the energy carry the same unit – mass. What $\hbar$ then does it for us to be able to measure mass in terms of the length unit, e.g. as in $p = \hbar k \rightarrow [M] = \hbar/[L]$.

Now at the end of the calculation, one must revert to the usual units. For this, you can memorize

$$\hbar c = 0.1973\,\text{GeV}\,\text{fm} \approx 0.2\,\text{GeV}\,\text{fm} \tag{16}$$

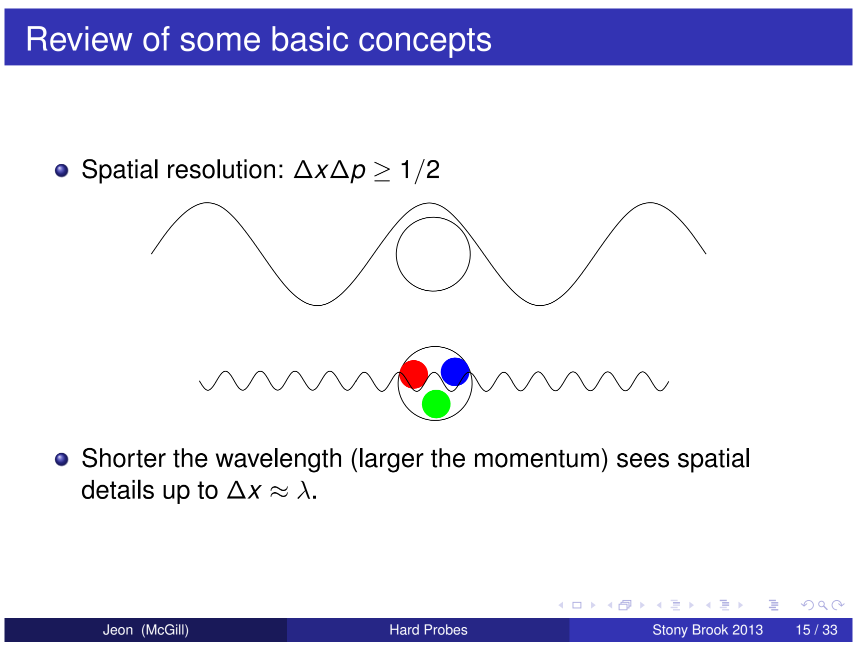
The approximation $\hbar c \approx 0.2\,\text{GeV}\,\text{fm}$ is accurate up to $1.5\,\%$ and can be used for estimation puroposes. This is basically the uncertainty principle

$$\Delta x \Delta p \geq \frac{\hbar}{2} \tag{17}$$

Hence, all dimensionful quantities, length, time, mass, momentum, energy can be measured with one unit, either fm or GeV, since, naturally, the length unit 1 fm is related to the energy unit 0.2 GeV as

$$0.2\,\mathrm{GeV} \approx \frac{1}{1\,\mathrm{fm}} \tag{18}$$

Consequences of the uncertainty principle



Consider the uncertainty principle

$$\Delta x \Delta p \geq \frac{1}{2} \tag{19}$$

and let's look at what it means.

Suppose you want to investigate a system which has the size of L. How do you study such a system? Well, how do you study any physical system? Answer: You need a probe.

Think of the x-ray pictures. The system you want to study is the human body. The way to study it is to bombard it with the high frequency light: x-ray. You then study how each part of the human body *response* to x-ray. Soft tissue will let it pass easily but bones will stop it. So by detecting the response of the system to the probe by making a photographic image of it, you can now study the human body. Now in this case, the x-ray is like a bullet. One does not really need to know that the x-ray is a wave. It could very well be a collection of small particles. This is because the wavelength of the x-ray is from

$$E = \omega = k = \frac{2\pi}{\lambda} \tag{20}$$

is much smaller than the human scale

$$\begin{aligned} \lambda &\sim \frac{2\pi}{\mathrm{keV}} \\ &= \frac{2\pi}{\mathrm{keV}} 2 \times 10^5\,\mathrm{keVfm} \\ &\sim 10^6\,\mathrm{fm} = 10^{-9}\,\mathrm{m} \end{aligned} \tag{21}$$

The ultra-sound scan, however, is a different matter. The usual ultra-sound scan uses the vibration frequency of about 5 MHz. The speed of sound in water is about 1,500 m/s. Hence, the wavelength λ is

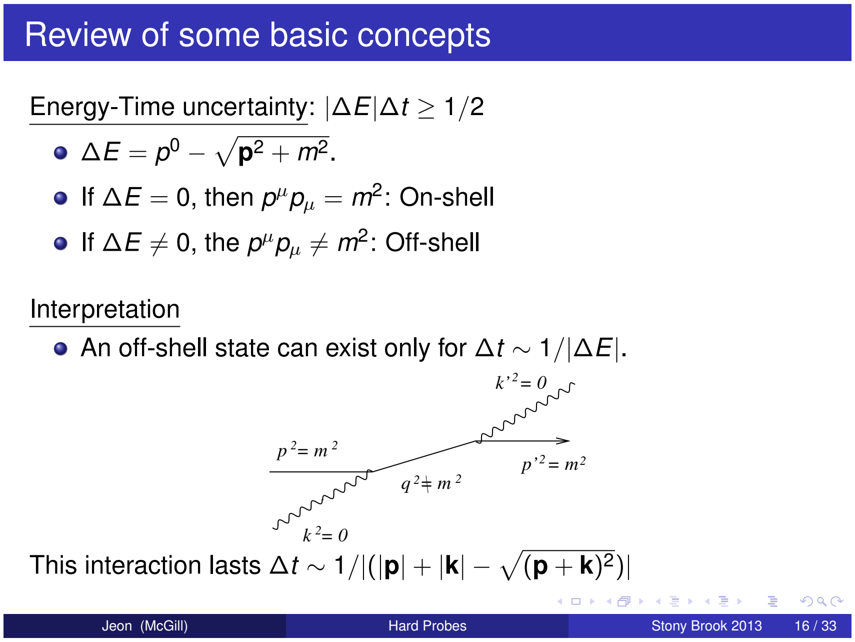
$$c_s = \frac{\lambda}{T} = \lambda f \tag{22}$$

or

$$\lambda \approx 0.3\,\mathrm{mm} \tag{23}$$

and you can't really see any detalis smaller than that scale. Why is that?

This is because the sound wave cannot really distinguish an obstacle that is of the size of the wavelength or smaller. Recall that when scattered, a wave undergoes diffraction. If the object is smaller than the wavelength, however, there is no diffraction, meaning that the object is invisible to the wave. The uncertainty principle is exactly that statement: One cannot obtain information on the position of an object that is smaller than the wavelength of the wave. Hence, to see smaller and smaller details of an object, one must use higher and higher wavenumber. As long as the speed of the wave is roughly constant (in the case of the sound wave) or constant (in the case of light), this means higher frequency and that means higher energy.



Now consider the statement

$$E = \hbar\omega \tag{24}$$

Writing $\omega = 2\pi f = 2\pi/T$, where f is the frequency and T is the period, we get

$$ET = h \tag{25}$$

which can be similarly interpreted that the time resolution depends on the energy scale.

How can this be used? Let's write it this way:

$$\Delta E \Delta t \geq \frac{1}{2} \tag{26}$$

One can interprete this in many ways. One way to interprete is the accuracy of the energy (frequency) measurement. If you just mesaure the amplitude of wave between one zero amplitude and the next zero amplitude, how accurately can you know the actual frequency of the wave? Well, since you have observed the behaviour only once, you can't really be sure and the uncertainty is large. If you haven't even observed one full oscillation, then you are only looking at a small part of the curve and your uncertainly accordingly goes higher. If you have observed a very small amount of time, you cant be sure at all of its frequency. On the other hand, if you have observed the wave for many ups and downs, you can measure the frequency pretty accurately although you can never shrink the uncertainty to zero (you never know exactly what the next time interval will bring!).

Another way to interprete this relationship, which is more useful to us, is the relationship between the "off-shell-ness" of a particle and its life time. If a particle is on-shell that means it satisfies

$$p^\mu p_\mu = E^2 - \mathbf{p}^2 = m^2 \tag{27}$$

where m is the particle mass. (We use the mostly negative metric convention $g_{\mu\nu} = \text{diag}(1, -1, -1, -1)$). In fact, there aren't that many elementary particles that actually satisfies this relationship with a single mass parameter m. For this to be true, the particle must be truly stable. For instance, a proton or an electron. Any particle that decays into other more stable particles in a finite life-time cannot be truly on-shell. The reason is that if a particle is truly on-shell, then we know the energy exactly. That is $\Delta E = 0$. Hence, the time uncertainty Δt diverges. It leaves for an indefinite amount o time.

On the other hand, suppose a particle decays. Then the wavefunction of that particle behaves like (in the non-relativisitc limit)

$$\psi(t) \sim e^{-iEt - \Gamma t/2} \tag{28}$$

where $\Gamma = 1/\tau_d$ is the decay rate and τ_d is the corresponding decay constant. Fourier transforming, we get

$$\psi(\omega) \sim \frac{\Gamma}{(E-\omega)^2 + \Gamma^2/4} \tag{29}$$

which indicates that the energy of the particle is uncertain by Γ. Since $E = \sqrt{m^2 + \mathbf{p}^2}$, that means the mass of the particle is not fixed, but distributed within $m \pm \Gamma$. One recorvers the Einstein relationship only in the limit $\Gamma \to 0$. One can say that the particle 4-momentum is off-shell by $\sim \Gamma$, then the lifetime of the particle is $\sim 1/\Gamma \sim 1/\tau_d$.

In relativistic quantum mechanics, a particle can not only decay into other particles but it can also absorb other particle. Consider an absorption of photon by an electron as in the Compton scattering. When an on-shell electron absorbs an on-shell photon, the total momentum becomes

$$p' = p + k \tag{30}$$

and the 4-momentum squared becomes

$$\begin{aligned}(p')^2 &= p^2 + k^2 + 2p \cdot k \\ &= m_e^2 + 2E_p|\mathbf{k}| - 2|\mathbf{p}||\mathbf{k}|\cos\theta\end{aligned} \tag{31}$$

where θ is the angle between the two momenta. This is clearly not equal to m_e^2. Even though the electron is completely stable, this state of the electron with $(p')^2 > m_e^2$ is not stable. It will eventually become an on-shell electron and an on-shell photon again. How long this unstable state can live depends on the on-shell-ness

$$\begin{aligned}\Delta Q^2 &= (p')^2 - m_e^2 \\ &= 2E_p|\mathbf{k}| - 2|\mathbf{p}||\mathbf{k}|\cos\theta\end{aligned} \tag{32}$$

For high energy $|\mathbf{p}| \gg m_e$, we have

$$E_p = |\mathbf{p}| + \frac{m_e^2}{2|\mathbf{p}|} + \cdots \tag{33}$$

and that gives

$$\Delta Q^2 = 2|\mathbf{p}||\mathbf{k}|(1 - \cos\theta) + m_e^2 \frac{|\mathbf{k}|}{|\mathbf{p}|} + \cdots \tag{34}$$

This can be small or large compared to m_e^2 depending on the angle θ and the ratio $|\mathbf{k}|/|\mathbf{p}|$.

The off-shell-ness can be small if $\theta \ll m_e/\sqrt{|\mathbf{p}||\mathbf{k}|}$ and $|\mathbf{k}|/|\mathbf{p}| \ll 1$. In this case, the electron is almost on-shell and it takes a long time until the photon is finally radiated. On the other hand, if $\Delta Q^2 \gg m_e^2$, that means $\theta \sim 1$ when $|\mathbf{p}| \gg m_e$ and $|\mathbf{k}| \gg m_e$ and this intermediate state is short lived.

Now ask it another way. Given ΔQ^2, what are the possible $\mathbf{p}$ and $\mathbf{k}$? We need to satisfy

$$\Delta Q^2 = 2|\mathbf{k}|(E_p - |\mathbf{p}|\cos\theta) \tag{35}$$

This is one equation but has 3 variables, $|\mathbf{k}|$, $|\mathbf{p}|$ and $\cos\theta$. The angle is given by

$$\cos\theta = \frac{\Delta Q^2}{2|\mathbf{k}||\mathbf{p}|} - \frac{E_p}{|\mathbf{p}|} \tag{36}$$

So when the energies are high, it is more likely that θ will be a large angle than a small angle.

What is a jet?

First, we should know what a jet is. A jet is a collimated shower of hadrons that occur in hadron-hadron collisions and electron-positron collisions.

Before we talk about jets, let's think about what happens when a two hadrons collide. To talk about that, we first need to talk about what a hadron is.

A hadron is defined as elementary particles which participate in the strong nuclear interaction. Good examples are protons and neutrons that make up the atomic nuclei and pions, kaons, ρ-mesons, Δ-baryons, Λ-baryons, etc that appear only in the interaction of the protons and neutrons (collectively called the nucleons).

What is a hadron?

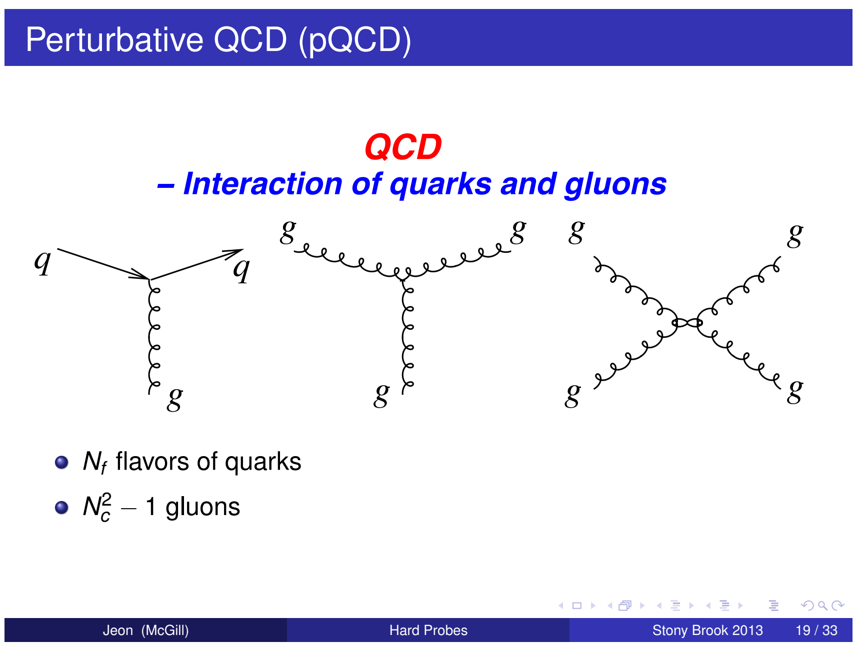
As mentioned above, a hadron is a particle that participate in strong nuclear interaction. One can say that the history of hadrons start with Hideki Yukawa. In 1935, Hideki Yukawa first published his idea on why the nuclear force between the nucleons is strong but short-ranged. His idea was that just like the electro-magnetic force is mediated by virtual photons, the nuclear force is also mediated by virtual "mesons". And he figured out that the range of the force is inversely proportional to the mass of the meson and the fact that the range of the nuclear force is about 1 fm indicates that the mass of the meson is about 0.2 GeV. And if there is a virtual particle in this interaction, there must be the corresponding real particle. The pions are subsequently discovered in Cosmic Ray experiments (1940's equivalent of particle accelerator experiments) and Yukawa got his Nobel prize in 1949.

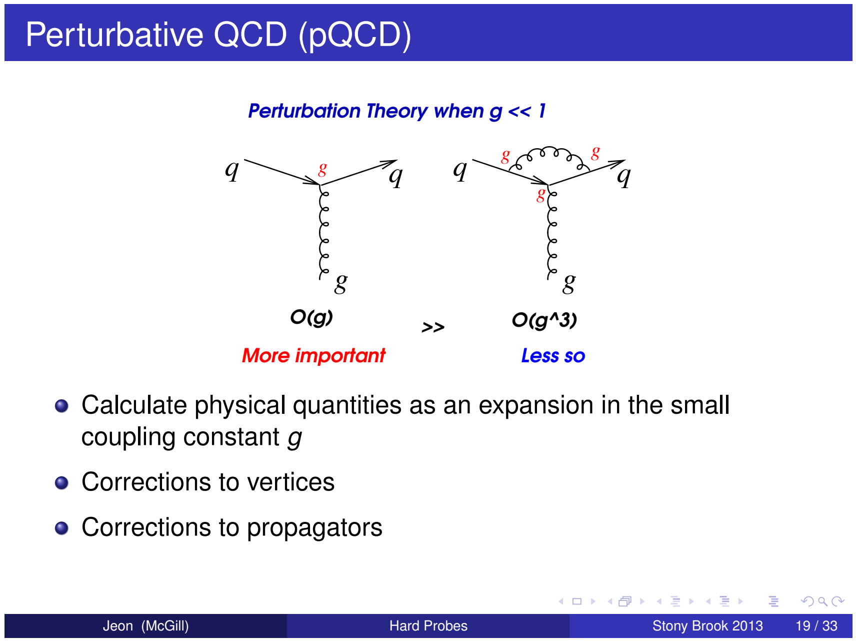
After the initial discovery, there followed a plethora of hadrons discovered in the Cosmic Ray experiments as well as accelerator experiments. The question then became: Why so many? Where did they all come from? Is there any organizing principle in analogy with the atomic periodic table?

This was answered partially when Gell-Mann and Zweig introduced the concept of quarks and their flavors, that is, u, d, s and fully answered by Han, Naumbu, Fritzsch and Gell-Mann when they wrote down the theory of quark interactions as a modern SU(3) gauge theory of colours called the Quantum Chromodynamics (QCD in short).

We now know that the integer spin mesons are made of one quark and one anti-quark and half-integer spin baryons are made of three quarks.

Quantum Chromodynamics





Quantum Chromodynamics (QCD) is the theory of strong nuclear interaction. In this theory, the interaction Lagrangina is given by

$$\mathcal{L} = \bar{q}_{f,a} i \not{D} q_{f,a} - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a \tag{37}$$

where a is the color index, $q_{f,a}$ is the spinor corresponding to the quark of the flavor f,

$$G_{\mu\nu} = \frac{1}{ig}[D_\mu, D_\nu] \tag{38}$$

is the field strength tensor of gluons (QCD equivalent of photons) and $D_\mu = \partial_\mu + igA_\mu$ is the covarient derivative with $A_\mu = A^a_\mu \tau_a$ being the gluon field. The matrices τ_a are the 8 generators of the SU(3) group.

$$\not{D} = \gamma^\mu \left(\partial_\mu + igA^a_\mu \tau_a \right) \tag{39}$$

is the covariant derivative contracted with the Dirac gamma matrices. The quarks have finite masses, but for the purpose of this lecture, we can set them to zero.

For quark, the color index runs from 1 to 3 or RGB since it is in the fundamental representation. For gluons, the color index runs from 1 to 8. It is in the color-anti-color combination except that there is no gluon with the color neutral combination $R\bar{R} + G\bar{G} + B\bar{B}$.

The QCD Lagrangian above looks similar to the QED Lagrangian, but the big difference is in the way gluons interact. In QED, the field strength tensor is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{40}$$

and hence, the photon Lagrangian $-F_{\mu\nu}F^{\mu\nu}/4$ does not contain any interaction. Photons interact directly with only the charged matter particles. Any photon-photon interaction must be mediated via a fermion loop. Not so with the gluons. The non-Abelian field strength is

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g f_{abc} A^b_\mu A^c_\nu \tag{41}$$

and the gluon Lagrangian $\mathcal{L}_G = -G^a_{\mu\nu}G^{\mu\nu}_a/4$ does contain the cubic and the quartic interactions among gluons.

QCD is a non-Abelian gauge theory because the color symmetry group is SU(3). Let's talk about that for a moment.

2 Gauge field

How do particles interact? The free-particle Hamiltonian operator for a many-body system is

$$\hat{H}_0 = \sum_{i=1}^{N} \frac{\hat{\mathbf{p}}_i^2}{2m_i} \tag{42}$$

There is no reason why the world cannot be like that. But then it would be boring. Nothing will ever happen in this world. Even the "position" of a particle isn't a well defined concept because there is no way to measure it. The Hamiltonian above is invariant under any amount of translation in space, $\mathbf{x}_i \to \mathbf{x}_i + \mathbf{a}_i$. Hence, pin-pointing the position of any particle is meaningless. So the question is, how can one make particles interact?

An obvious answer to this question is just to add the interaction potential:

$$\hat{H}_V = \sum_{i=1}^{N} \frac{\hat{\mathbf{p}}_i^2}{2m_i} + \sum_{i>j} V(\mathbf{x}_i - \mathbf{x}_j) \tag{43}$$

But why? Why is this the "obvious" way for particles to interact? This may seem a bit of a silly question to ask. Classically, this is tantamount to asking why the force in the Newton's law $\mathbf{F} = m\mathbf{a}$ is given by $\mathbf{F} = -\nabla V$. In classical mechanics, they are actually just definitions. The force is defined to be what causes acceleration. The potential energy is defined to be a quantity that gives $\mathbf{F} = -\nabla V$. There is no sense asking why they are so.

What we are asking here is a bit different. It is almost philosophical. We are asking to justify the existence of the potential energy in quantum mechanics given the kinetic energy exists.

Let's start with the free Hamiltonian again

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m} \tag{44}$$

That is, this world entirely consists of a single particle. One may ask: What do we mean by time in this case? What do we mean by position and momentum? What do we mean by the phase of a wavefunction? Well, they don't mean a thing in this case. For these concepts to be meaningful, it must be in relation to other particles. OK. Let's add another particle.

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}_1^2}{2m_1} + \frac{\hat{\mathbf{p}}_2^2}{2m_2} \tag{45}$$

Now the concept of total and relative momenta may be formulated in a meaningful way, but the question of "position" is still a meaningless one. In the wavefunction

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = Ae^{i\mathbf{p}_1\cdot(\mathbf{x}_1-\mathbf{x}_0)+i\mathbf{p}_2\cdot(\mathbf{x}_2-\mathbf{x}_0')} \tag{46}$$

the positions $\mathbf{x}_1$ and $\mathbf{x}_2$ can be measured against two arbitrary points $\mathbf{x}_0$ and $\mathbf{x}_0'$ and it would still make perfect sense. Note that the combined factor $e^{i\mathbf{p}_1\cdot\mathbf{x}_0+i\mathbf{p}_2\cdot\mathbf{x}_0'}$ is a constant, pure phase. Therefore, another way of stating the same thing is that the wavefunction is defined only modulo an arbitrary phase.

This comes about because the Schrödinger equation is a linear equation and the probability density of the particle is given by

$$P(t, \mathbf{x}) = \psi^*(t, \mathbf{x})\psi(t, \mathbf{x}) = |\psi(t, \mathbf{x})|^2 \tag{47}$$

If one extracts a pure phase from ψ

$$\psi'(t, \mathbf{x}) = e^{i\alpha}\psi(t, \mathbf{x}) \tag{48}$$

one still has

$$|\psi'(t, \mathbf{x})| = |\psi(t, \mathbf{x})| \tag{49}$$

Hence, the *physical* probability density distribution does not change even if the phase changes by an arbtrary amount.

Now if we are to add meaningful interaction between the two particles, the concept of relative position should be made meaningful although the concept of absolute position is still ambigous. But that's OK. So how does one go about adding an interaction? That is, how can the two particle "know about each other"? For that, we borrow some concepts from geometry. The idea here is that the presence of a particle changes the geometry of some manifold which then forms the background of the motion of the second particle. For the General Relativity, it is actually the geometry of space-time that changes in the presence of any form of energy. For other interactions, it's a bit more subtle. By the way, this was originally Weyl's idea. He extended GR to include a scale factor and could argue the necessity of a vector poential which looked very much like the 4-vector potential of E & M. However, it turned out that in this way, E & M and GR mixes up in such a way that the behavior of a atomic clock can depend on the trajectory of the clock. Experimentally, this is not the case. So we need to find an alternative way.

If the geometry of the space-time is out of the picture, what other geometry is available? In classical mechanics, there is no other. In quantum mechanics, however, there is another possibility. It's the phase of the wave-function. It may be somewhat odd to think of a phase having a geometry. It can be made a little bit more familiar if you think of a complex number as a 2-D vector.

The value of a complex number $\psi(x^\mu) = f_1(x^\mu) + i f_2(x^\mu)$ at a fixed space-time position $x^\mu = (t, \mathbf{x})$ can be represented by a 2-D vector $\boldsymbol{\psi} = (f_1(x^\mu), f_2(x^\mu))$. Using the basis vectors, we have

$$\boldsymbol{\psi} = f_1(x^\mu)\mathbf{e}_1(x^\mu) + f_2(x^\mu)\mathbf{e}_2(x^\mu) \tag{50}$$

where $\mathbf{e}_1$ is the real direction and $\mathbf{e}_2$ is the imaginary direction. Now suppose we make a transformation of the basis

$$\mathbf{e}_1' = \cos\alpha\mathbf{e}_1 - \sin\alpha\mathbf{e}_2 \tag{51}$$

$$\mathbf{e}_2' = \cos\alpha\mathbf{e}_2 + \sin\alpha\mathbf{e}_1 \tag{52}$$

Inversely,

$$\mathbf{e}_1 = \cos\alpha\mathbf{e}'_1 + \sin\alpha\mathbf{e}'_2 \tag{53}$$

$$\mathbf{e}_2 = \cos\alpha\mathbf{e}'_2 - \sin\alpha\mathbf{e}'_1 \tag{54}$$

Hence

$$\begin{aligned}\boldsymbol{\psi} &= f_1(\cos\alpha\mathbf{e}'_1 + \sin\alpha\mathbf{e}'_2) + f_2(\cos\alpha\mathbf{e}'_2 - \sin\alpha\mathbf{e}'_1) \\ &= (f_1\cos\alpha - f_2\sin\alpha)\mathbf{e}'_1 + (f_2\cos\alpha + f_1\sin\alpha)\mathbf{e}'_2\end{aligned} \tag{55}$$

That is,

$$f'_1 = \cos\alpha f_1 - \sin\alpha f_2 \tag{56}$$

$$f'_2 = \cos\alpha f_2 + \sin\alpha f_1 \tag{57}$$

and

$$f_1 = \cos\alpha f'_1 + \sin\alpha f'_2 \tag{58}$$

$$f_2 = \cos\alpha f'_2 - \sin\alpha f'_1 \tag{59}$$

And if we write

$$\boldsymbol{\psi}' = f'_1\mathbf{e}'_1 + f'_2\mathbf{e}'_2 \tag{60}$$

then obviously

$$\boldsymbol{\psi} = \boldsymbol{\psi}' \tag{61}$$

Now in the complex number space, the above transformation is equivalent to

$$\begin{aligned}\psi' &= f'_1 + if'_2 \\ &= (\cos\alpha + i\sin\alpha)(f_1 + if_2) \\ &= e^{i\alpha}\psi\end{aligned} \tag{62}$$

and it looks like that they represent two different numbers. This would be the case if we *rotated* the vector instead of *using different basis vectors*. Without the concept of the basis vectors, this is the only way to interprete. However, if one applies the concept of the tangent space (or more technically correct way, the fibre bundle), then ψ and ψ' obviously represent the same quantity but expressed in different basis. Quantum mechanically, we know that ψ and ψ' represent the same physical probability distribution. Hence, the interetation that $\boldsymbol{\psi}$ and $\boldsymbol{\psi}'$ reprsent the same quantity is more natural than regarding ψ and ψ' as equivalent modulo a phase.

These "basis vectors" at each x^μ are an additional concept on top of the concept of the complex number. It represents the possibility that what we call the "real" function space and the "imaginary" function space do not have to be the same at different space-time points. Is there a way to use this to represent interactions between two particles? If there are only free particles, there is no reason why the basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ should depend on the space-time point. But if they interact, that is, if they "know about each other", then one can express the phase of one wavefunction with respect to the other. In other words, $\boldsymbol{\psi}_1$ can be expressed in the basis defined by the wavefunction $\boldsymbol{\psi}_2$ and vice versa.

If the two particles are in a relative motion, then the basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ seen by, say particle 2, does not need to remain constant. They will actually be a function of time and space. For particle 2, the wavefunction can be represented by

$$\boldsymbol{\psi}(x^\mu) = f_1(x^\mu)\mathbf{e}_1(x^\mu) + f_2(x^\mu)\mathbf{e}_2(x^\mu) \tag{63}$$

where the basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ depends on what the particle 1 is doing. The change in the wavefunction in x_ν direction is

$$D_\nu \boldsymbol{\psi} \equiv \lim_{dx^\nu \to 0} \frac{\boldsymbol{\psi}(x + dx^\nu) - \boldsymbol{\psi}(x)}{dx^\nu} \tag{64}$$

This derivative now includes the effect of the changing axis. We have now

$$\begin{aligned} D_\nu \boldsymbol{\psi} &= \frac{\boldsymbol{\psi}(x + dx^\nu) - \boldsymbol{\psi}(x)}{dx^\nu} \\ &= \mathbf{e}_1 \partial_\nu f_1 + \mathbf{e}_2 \partial_\nu f_2 + f_1 \partial_\nu \mathbf{e}_1 + f_2 \partial_\nu \mathbf{e}_2 \\ &= \mathbf{e}_1 \partial_\nu f_1 + \mathbf{e}_2 \partial_\nu f_2 \\ &\quad + f_1 \mathbf{e}_1(\mathbf{e}_1^T \partial_\nu \mathbf{e}_1) + f_1 \mathbf{e}_2(\mathbf{e}_2^T \partial_\nu \mathbf{e}_1) + f_2 \mathbf{e}_1(\mathbf{e}_1^T \partial_\nu \mathbf{e}_2) + f_2 \mathbf{e}_2(\mathbf{e}_2^T \partial_\nu \mathbf{e}_2) \end{aligned} \tag{65}$$

Since $\mathbf{e}_1^T \mathbf{e}_1 = \mathbf{e}_2^T \mathbf{e}_2 = 1$, $\mathbf{e}_1^T \partial \mathbf{e}_1 = \mathbf{e}_2^T \partial \mathbf{e}_2 = 0$ and

$$\mathbf{e}_T^1 \partial \mathbf{e}_2 = -(\partial \mathbf{e}_1^T)\mathbf{e}_2 = -\mathbf{e}_2^T(\partial \mathbf{e}_1) \tag{66}$$

Hence

$$D_\nu \boldsymbol{\psi} = \mathbf{e}_1 \left(\partial_\nu f_1 + f_2(\mathbf{e}_1^T \partial_\nu \mathbf{e}_2) \right) + \mathbf{e}_2 \left(\partial_\nu f_2 - f_1(\mathbf{e}_1^T \partial_\nu \mathbf{e}_2) \right) \tag{67}$$

Let

$$A_\nu^{ij} = \mathbf{e}_i^T \partial_\nu \mathbf{e}_j = \mathbf{e}_i \cdot \partial_\nu \mathbf{e}_j \tag{68}$$

then

$$A_\nu^{ij} = J_{ij} A_\nu \tag{69}$$

where $A_\nu = \mathbf{e}_2 \cdot \partial_\nu \mathbf{e}_1$ and

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{70}$$

is the matrix version of i since

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \tag{71}$$

and

$$\mathbf{J} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} -f_2 \\ f_1 \end{pmatrix} \tag{72}$$

Then

$$\mathbf{e}_i \cdot D_\nu \Psi = \partial_\nu f_i + A_\nu J_{ij} f_j \tag{73}$$

Hence as a complex number

$$\begin{aligned} D_\nu \psi &= \partial_\nu (f_1 + i f_2) + i A_\nu (f_1 + i f_2) \\ &= (\partial_\nu + i A_\nu) \psi \end{aligned} \tag{74}$$

Therefore the presence of another particle manifests itself as a non-zero 4-vector A_ν. So far, however, we have not talked about *how* A_ν depends on the other particle. We'll do that later.

In summary, the Hamiltonian

$$\hat{H} = \frac{\hat{\mathbf{p}}_1^2}{2m_1} + \frac{\hat{\mathbf{p}}_2^2}{2m_2} \psi \tag{75}$$

in the presence of interaction becomes

$$\hat{H} = \frac{(\hat{\mathbf{p}}_1 - \mathbf{A}_1)^2}{2m_1} + \frac{(\hat{\mathbf{p}}_2 - \mathbf{A}_2)^2}{2m_2} + A_0 \tag{76}$$

where $\mathbf{A}_1$ is the vector potential caused by particle 2 and $\mathbf{A}_2$ is the vector potential caused by particle 1. A_0 is the potential energy between the two particles coming from $i\partial_t \to i\partial_t - A_0$.

3 Gauge Invariance

Consider a situation where the second particle is much heavier than the first. Then this particle is not going to be influenced by the presence of the second one. Hence A_μ for the second particle can be regarded as fixed, or external.

The time-dependent Schrödinger equation in this case

$$i(\partial_t + iA_0)\psi = \frac{(\hat{\mathbf{p}} - \mathbf{A})^2}{2m}\psi \tag{77}$$

posseses an obvious symmetry. If ψ is a solution of this equation, then $\psi' = e^{i\alpha}\psi$ is also a solution provided that α is constant.

But what if $\alpha(t, \mathbf{x})$ is a function of space-time? Certainly, $|\psi'(t, \mathbf{x})| = |\psi(t, \mathbf{x})|$ is still valid. Hence, the *classical physics* described by both functions are the same. Another way of saying it is, if one get the corresponding classical trajectory from the two functions following the maximum, they will be exactly the same. It may sound then reasonable to demand that the two function satisfies the same Schrödinger's equation. But that's not possible. The function $\psi'(t, \mathbf{x})$ defined in Eq.(48) does not satisfiy the same Schr̈dinger equation since

$$\begin{aligned}\partial_\mu\psi &= \partial_\mu(e^{-i\alpha}\psi') \\ &= e^{-i\alpha}\left(\partial_\mu - i\partial_\mu\alpha\right)\psi'\end{aligned} \tag{78}$$

Hence, ψ' satisfies

$$i\left(\partial_t + iA_0 - i\partial_t\alpha\right)\psi' = \frac{1}{2m}g^{ij}(\partial_i + iA_i - i\partial_i\alpha)(\partial_j + iA_j - i\partial_j\alpha)\psi' \tag{79}$$

Our metric is $g^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$. Clearly, ψ' depends on α. Hence, we have a situation where infinitely many different Schrödinger's equations describe exactly the same *classical physics*. This also means that the potential energy is not unique. Huh? How can that be? What happened to the energy consevation?

What did we do wrong? Algebraically, nothing. Geometrically, yes, we've done something wrong: We forgot to rotate the basis vectors as well as the components. Recall that a complex number $\psi = f_1 + if_2$ can be regarded as a 2-dimensional vector $\boldsymbol{\psi} = (f_1, f_2)$ defined in a 2-D plane anchored at $(t, \mathbf{x})$. We've seen that if we transform

$$\mathbf{e}_1' = \cos\alpha\mathbf{e}_1 - \sin\alpha\mathbf{e}_2 \tag{80}$$

$$\mathbf{e}_2' = \cos\alpha\mathbf{e}_2 + \sin\alpha\mathbf{e}_1 \tag{81}$$

and at the same time

$$f_1' = \cos\alpha f_1 - \sin\alpha f_2 \tag{82}$$
$$f_2' = \cos\alpha f_2 + \sin\alpha f_1 \tag{83}$$

then

$$\boldsymbol{\psi}' = \boldsymbol{\psi} \tag{84}$$

So in the system where $\mathbf{e}_{1,2}'$ are used, we get,

$$(D_\nu\psi)' = (\partial_\nu + iA_\nu')\psi' \tag{85}$$

not $(\partial_\nu + iA_\nu)\psi'$. Note that the 4-vector potential also changed from A_ν to A_ν'. Of course, here $\psi' = f_1' + if_2'$ and $A_\nu' = -\mathbf{e}_1'^T\partial_\nu\mathbf{e}_2'$. But this should be exactly the same as $D_\nu\psi$ since $\boldsymbol{\psi}'$ and $\boldsymbol{\psi}$ are the same, just expressed in different basis. Therefore, there must be a consistent relationship between A_ν and A_ν'.

To find the relationship between A_ν and A_ν', we start with $D_\nu\boldsymbol{\psi}$ and transform the basis from $\mathbf{e}_i$ to $\mathbf{e}_i'$ and then back to $\mathbf{e}_i$ as follows

$$\mathbf{e}_i' = R_{ij}\mathbf{e}_j \tag{86}$$

$$\mathbf{e}_i'f_i' = R_{ij}\mathbf{e}_jf_i' = (f_i'R_{ij})\mathbf{e}_j \tag{87}$$

or

$$f_j = R_{ji}^Tf_i' \tag{88}$$

$$\begin{aligned}&\partial_\mu\mathbf{e}_if_i\\ &\quad= \mathbf{e}_i\partial_\mu f_i + (\partial_\mu\mathbf{e}_i)\cdot\mathbf{e}_j\mathbf{e}_jf_i\\ &\quad= \mathbf{e}_i\partial_\mu f_i + (\partial_\mu\mathbf{e}_j)\cdot\mathbf{e}_i\mathbf{e}_if_j\\ &\quad= \mathbf{e}_i\left(\partial_\mu f_i + (\partial_\mu\mathbf{e}_j)\cdot\mathbf{e}_if_j\right)\end{aligned} \tag{89}$$

Let

$$A_\mu^{ij} = -\mathbf{e}_i\cdot(\partial_\mu\mathbf{e}_j) = -A_\mu^{ji} \tag{90}$$

Then

$$D_\mu f_i = \partial_\mu f_i + A^{ij}_\mu f_j \tag{91}$$

In 2D

$$A^\mu_{ij} = J A^\mu \tag{92}$$

where $A^\mu = \mathbf{e}_2 \cdot (\partial_\mu \mathbf{e}_1)$ is the common function.

Now changing basis to the primed one,

$$\begin{aligned}
\partial_\mu(\mathbf{e}_i f_i) &= \partial_\mu(R^T_{ij}\mathbf{e}'_j f_i) \\
&= (\partial_\mu R^T_{ij})\mathbf{e}'_j f_i + R^T_{ij}(\partial_\mu \mathbf{e}'_j) f_i + R^T_{ij}\mathbf{e}'_j(\partial_\mu f_i) \\
&= (\partial_\mu R^T_{ij})\mathbf{e}'_j f_i + \mathbf{e}'_k R^T_{ij}\left[(\partial_\mu \mathbf{e}'_j)\cdot \mathbf{e}'_k\right] f_i + R^T_{ij}\mathbf{e}'_j(\partial_\mu f_i)
\end{aligned} \tag{93}$$

To project it back to the unprimed basis,

$$\begin{aligned}
\partial_\mu(\mathbf{e}_i f_i) &= (\partial_\mu R^T_{ij})R_{jl}\mathbf{e}_l f_i + R_{kl}\mathbf{e}_l R^T_{ij}\left[(\partial_\mu \mathbf{e}'_j)\cdot \mathbf{e}'_k\right] f_i + R^T_{ij}R_{jl}\mathbf{e}_l(\partial_\mu f_i) \\
&= (\partial_\mu R^T_{ij})R_{jl}\mathbf{e}_l f_i + R_{kl}\mathbf{e}_l R^T_{ij}\left[(\partial_\mu \mathbf{e}'_j)\cdot \mathbf{e}'_k\right] f_i + \mathbf{e}_l(\partial_\mu f_l)
\end{aligned} \tag{94}$$

which gives

$$D_\mu f_l = \partial_\mu f_l + R^T_{lk} {A'}^{kj}_\mu R_{ji} f_i + (R^T_{lj}\partial_\mu R_{ji}) f_i \tag{95}$$

and

$$A^{li}_\mu = R^T_{lk} {A'}^{kj}_\mu R_{ji} + R^T_{lj}\partial_\mu R_{ji} \tag{96}$$

In 2D,

$$\begin{aligned}
\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}
\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}
\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}
&= \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}
\begin{pmatrix} -\sin\alpha & \cos\alpha \\ -\cos\alpha & -\sin\alpha \end{pmatrix} \\
&= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}
\end{aligned} \tag{97}$$

is invariant. Hence

$$R^T_{lk} {A'}^{kj}_\mu R_{ji} = J A'_\mu \tag{98}$$

The additional matrix

$$\left(\begin{array}{cc} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{array}\right)\partial_\mu\left(\begin{array}{cc} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{array}\right) = \left(\begin{array}{cc} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{array}\right)(\partial_\mu\alpha)\left(\begin{array}{cc} -\sin\alpha & \cos\alpha \\ -\cos\alpha & -\sin\alpha \end{array}\right)$$
$$= (\partial_\mu\alpha)\left(\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right) = -(\partial_\mu\alpha)J \qquad (99)$$

so that

$$A_\mu = A'_\mu - \partial_\mu\alpha \qquad (100)$$

Therefore,

$$A_\nu = A'_\nu - \partial_\nu\alpha \qquad (101)$$

From our derivation, it is clear that these two connections, A_ν and $A'_\nu = A_\mu - \partial_\nu\alpha$, must be equivalent since α is just an arbitrary reparametrization of underlying geometry. This is called the "gauge freedom". Physical quantities must be then "gauge invariant" meaning that they should not depend on the choice of α.

If one can find a α in such a way that

$$A_\nu + \partial_\nu\alpha = 0 \qquad (102)$$

then there exists a frame where $A'_\nu = 0$. Since what we did was a mere re-definition of the frame, geometry (and hence physics) should not depend on our choice α. So if one can really set $A'_\nu = 0$, then the space must be a free-space.

Since A_ν is not unique due to this gauge freedom, in geometry, the physical quantity that uniquely represents the non-flat nature of the geometry is the curvature tensor

$$F_{\mu\nu} = i[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu \qquad (103)$$

Since $[\partial_\mu, \partial_\nu] = 0$, it is trivial that $\partial_\nu\alpha$ term disappears in $F_{\mu\nu}$. Classical physics does not directly depend on the 4-vector potential A_μ, it only depends on $F_{\mu\nu}$.

We've argued that the presence of non-trivial A_μ or a non-zero "curvature" $F_{\mu\nu}$ necessarily implies the presence of other particles. Question to ask here is how? This is a question of dynamics. As such, one needs a Lagrangian to answer that.

The Lagrangian-density for the Schrödinger's equation is

$$\mathcal{L} = i\psi^*\partial_t\psi - \frac{\nabla\psi^* \cdot \nabla\psi}{2m} - V(\mathbf{x})\psi^*\psi \tag{104}$$

One can easily see that changing $\psi \to e^{i\alpha}\psi$ leaves $\mathcal{L}$ intact. That is, it does not depend on α. What is the consequence of this invariance? To find out, we start with the Lagrangian. A change in ψ leads to

$$\begin{aligned}
\delta\mathcal{L} &= \mathcal{L}(\psi + \delta\psi, \partial_\mu\psi + \partial_\mu\delta\psi) - \mathcal{L}(\psi, \partial_\mu\psi) \\
&= \delta\psi\frac{\partial\mathcal{L}}{\partial\psi} + (\partial_\mu\delta\psi)\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi} \\
&\quad + \delta\psi^*\frac{\partial\mathcal{L}}{\partial\psi^*} + (\partial_\mu\delta\psi^*)\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi^*} \\
&= \delta\psi\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi} + (\partial_\mu\delta\psi)\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi} \\
&\quad + \delta\psi^*\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi^*} + (\partial_\mu\delta\psi^*)\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi^*}
\end{aligned} \tag{105}$$

where we used the Euler-Lagrange equations

$$\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi} = \frac{\partial\mathcal{L}}{\partial\psi} \quad \text{and} \quad \partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi^*} = \frac{\partial\mathcal{L}}{\partial\psi^*} \tag{106}$$

The above then becomes

$$\delta\mathcal{L} = \partial_\mu\left(\delta\psi\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi}\right) + \partial_\mu\left(\delta\psi^*\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi^*}\right) \tag{107}$$

Suppose

$$\delta\psi = e^{i\alpha}\psi - \psi \tag{108}$$

Now we know that

$$\mathcal{L}(\psi + \delta\psi, \partial_\mu\psi + \partial_\mu\delta\psi) = \mathcal{L}(\psi, \partial_\mu\psi) \tag{109}$$

since $\mathcal{L}$ does not depend on a constant α. Hence, using $\delta\psi \approx i\alpha\psi$ and $\delta\psi^* \approx -i\alpha\psi^*$, one can easily see that

$$\partial_\mu\left(\psi\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi} - \psi^*\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi^*}\right) = 0 \tag{110}$$

That is, there is a conserved current. This is, of course, Noether's theorem (Emmy Noether, 1918). We define

$$\begin{aligned} J^\mu &= \frac{i}{2}\left(\psi\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi} - \psi^*\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi^*}\right) \\ &= \left(\psi^*\psi, -\frac{i}{2m}(\psi^*\nabla\psi - \psi\nabla\psi^*)\right) \end{aligned} \tag{111}$$

which you should recoginze as the probability current in quantum mechanics.

Now suppose α is a function of space-time. Then,

$$\begin{aligned} \partial_\mu\psi' &= \partial_\mu(e^{i\alpha}\psi) \\ &= e^{i\alpha}\left(\partial_\mu + i\partial_\mu\alpha\right)\psi \end{aligned} \tag{112}$$

Therefore

$$\mathcal{L}(\psi', \partial_\mu\psi') \neq \mathcal{L}(\psi, \partial_\mu\psi) \tag{113}$$

and

$$\begin{aligned} \delta\mathcal{L} &= \partial_\mu\left(i\alpha\psi\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi}\right) - \partial_\mu\left(i\alpha\psi^*\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi^*}\right) \\ &= 2\partial_\mu(\alpha J^\mu) \\ &= 2(\partial_\mu\alpha)J^\mu + 2\alpha\partial_\mu J^\mu \\ &\neq 0 \end{aligned} \tag{114}$$

and hence the conserved current does not result. But this is not good. Charge is conserved and probability is conserved. So there must be a conserved current. So what is one supposed to do?

We start with our Lagrangian that includes the 4-vector potential. Then it is easy to see that the Langrangian is invariant under gauge transformations

$$\begin{aligned} \mathcal{L}' &= i\psi'^*(\partial_t + iV')\psi' \\ &\quad + \frac{g^{ij}}{2m}\left[(\partial_i + iA'_i)\psi'\right]^*\left[(\partial_j + iA'_j)\psi'\right] \\ &= i\psi^*(\partial_t + iA_0)\psi \\ &\quad + \frac{g^{ij}}{2m}\left[(\partial_i + iA_i)\psi\right]^*\left[(\partial_j + iA_j)\psi\right] \\ &= \mathcal{L} \end{aligned} \tag{115}$$

where we have identified V with the time component of a 4-vector A_μ which transforms $A'_\mu = A_\mu - \partial_\mu \alpha$.

With A_μ added, the Euler-Lagrange equations are

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} = \frac{\partial \mathcal{L}}{\partial \psi} \tag{116}$$

and

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi^*} = \frac{\partial \mathcal{L}}{\partial \psi^*} \tag{117}$$

Check:

$$\partial_t(-i\psi) + \partial_i \left[(\partial^i + iA^i)\psi \right] /2m = -iA^j(\partial_j + iA_j)\psi/2m - A_0\psi \tag{118}$$

or

$$i(\partial_t + iA_0)\psi \;\;=\;\; -\frac{1}{2m}(\nabla + i\mathbf{A})\cdot(\nabla + i\mathbf{A})\psi \tag{119}$$

Now under a gauge transformation,

$$\begin{aligned} \delta\mathcal{L} &= \partial_\mu \left(\delta\psi \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} \right) + \partial_\mu \left(\delta\psi^* \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi^*} \right) + \delta A_\mu \frac{\partial \mathcal{L}}{\partial A_\mu} \\ &= \partial_\mu \left(i\alpha\psi \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} \right) - \partial_\mu \left(i\alpha\psi^* \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi^*} \right) - (\partial_\mu \alpha) \frac{\partial \mathcal{L}}{\partial A_\mu} \\ &= 0 \end{aligned} \tag{120}$$

With an arbitrary α, this gives 2 equations. The coefficient of α must vanish:

$$\partial_\mu \left(\psi \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} - \psi^* \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi^*} \right) = 0 \tag{121}$$

The coefficient of $\partial_\mu \alpha$ must vanish:

$$i \left(\psi \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} - \psi^* \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi^*} \right) = \frac{\partial \mathcal{L}}{\partial A_\mu} \tag{122}$$

From

$$\mathcal{L} \;\;=\;\; i\psi^*(\partial_t + iA_0)\psi + \frac{g^{ij}}{2m} \left[(\partial_i + iA_i)\psi \right]^* \left[(\partial_j + iA_j)\psi \right] \tag{123}$$

We have

$$i\left(\psi\frac{\partial\mathcal{L}}{\partial\partial_t\psi} - \psi^*\frac{\partial\mathcal{L}}{\partial\partial_t\psi^*}\right) = i\left(\psi i\psi^* + \psi^* i\psi\right) = -2\psi^*\psi \tag{124}$$

$$i\left(\psi\frac{\partial\mathcal{L}}{\partial\partial_i\psi} - \psi^*\frac{\partial\mathcal{L}}{\partial\partial_i\psi^*}\right) = i\left(\psi[(\partial^i + iA^i)\psi]^*/2m - \psi^*[(\partial^i + iA^i)\psi]/2m\right) \tag{125}$$

$$\begin{aligned}
\partial_t(-2\psi^*\psi) &= -2(\partial_t\psi^*)\psi - 2\psi^*(\partial_t\psi) \\
&= -2i[(-D_i^\dagger D_i^\dagger)\psi/2m + A_0\psi]^*\psi + 2i\psi^*[(-D_iD_i)\psi/2m + A_0\psi] \\
&= \frac{i}{m}\Big[\psi(\partial_i - iA_i)(\partial_i - iA_i)\psi^* - \psi^*(\partial_i + iA_i)(\partial_i + iA_i)\psi\Big] \\
&= \frac{i}{m}\Big[\psi(\nabla^2 - i(\nabla\cdot\mathbf{A}) - 2i\mathbf{A}\cdot\nabla)\psi^* - \psi^*(\nabla^2 + i(\nabla\cdot\mathbf{A}) + 2i\mathbf{A}\cdot\nabla)\psi\Big] \\
&= \frac{i}{m}\partial_i\Big[\psi(\partial_i - iA_i)\psi^* - \psi^*(\partial_i + iA_i)\psi\Big]
\end{aligned} \tag{126}$$

Hence we recover a conserved current. Charge conservation is safe. The requirement that the coefficient of $\partial_\mu\alpha$ vanishes just provides the consistency check for the form of J^μ.

The fact that demanding gauge invariance led to a conserved current does not tell us anything about the *dynamics* of the gauge field A_μ. We need additional piece of physics to get that.

Note that the classical physics is embedded in Schrödinger equation as the 0-th order equation for S if one writes $\psi = e^{iS/\hbar}$. Let's try that. The Schrödinger equation in the presence of A_μ is

$$\left(i\partial_t - A_0\right)\psi = -g^{ij}(-i\partial_i + A_i)(-i\partial_j + A_j)\psi/2m \tag{127}$$

Letting $\psi = e^{iS}$, we get for the LHS,

$$\left(i\partial_t - A_0\right)e^{iS} = e^{iS}\left(-\partial_t S - A_0\right) \tag{128}$$

For the RHS,

$$\begin{aligned}
(\text{RHS})2m &= -g^{ij}(-i\partial_i + A_i)(-i\partial_j + A_j)\psi \\
&= (-i\partial_i + A_i)(i\partial^i - A^i)\psi \\
&= \left(-\nabla^2 + i(\partial_i A^i) + iA^i\partial_i + iA_i\partial^i - A_iA^i\right)\psi \\
&= \left(-\nabla^2 + i(\nabla\cdot\mathbf{A}) + 2i\mathbf{A}\cdot\nabla + \mathbf{A}^2\right)\psi \\
&= \left(-\hbar^2\nabla^2 + i(\hbar\nabla\cdot\mathbf{A}) + 2i\hbar\mathbf{A}\cdot\nabla + \mathbf{A}^2\right)e^{iS/\hbar} \\
&= e^{iS/\hbar}\left((\nabla S)\cdot(\nabla S) - i\hbar(\nabla^2 S) + i\hbar(\nabla\cdot\mathbf{A}) - 2\mathbf{A}\cdot\nabla S + \mathbf{A}^2\right) \\
&\approx e^{iS/\hbar}\left((\nabla S)\cdot(\nabla S) - 2\mathbf{A}\cdot\nabla S + \mathbf{A}^2\right) \qquad (129)
\end{aligned}$$

Upon identifying ∇S with the classical momentum $\mathbf{p}$, one can then see that this leads to the classical Hamilotonian

$$H_{\text{cl}} = \frac{(\mathbf{p} - \mathbf{A})^2}{2m} + A_0 \qquad (130)$$

Therefore, it is natural to identify A^μ with the classical electromagnetic 4-vector potential. The classical 4-vector potential of course satisfies the Maxwell equations.

OK. Let's recap. What did we just see? We started with the non-relativistic Schrödinger equation and the corresponding Lagrangian. Then we saw that the phase of the free-particle wavefunction can be freely changed without changing any observables. Then we argue that the presence of electromagnetic force changes the functional space in such a way that one must introduce a "connection" which relates the basis vectors at one point to the basis vectors at another point. And then demanding that the physics of both the connection and the Schr̈dinger equation remains the same, we ended up with the full set of Maxwell equation. This is the simplest prototype of gauge theories.

Now suppose we have two species of particles. They are identical except for the eigenvalue of one Hermitian operator. The wavefunction of the two species can be represented by the vector (Note: 2D complex space – *different* 2D space)

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \qquad (131)$$

Suppose that the physics of these two species are the same. That is, as long as the transformation is unitary, the linear transformation

$$\Psi' = U\Psi \qquad (132)$$

does not change physics. Now the 2×2 complex matrix U must be unitary since we don't want to mess with the norm of the wavefunction. Hence demanding

$$(\Psi')^\dagger \Psi' = \Psi^\dagger \Psi \tag{133}$$

gives

$$U^\dagger U = U U^\dagger = 1 \tag{134}$$

That is, the matrices U form the U(2) group. If we now specify only the matrices satisfying $\mathrm{Det}\, U = 1$, we have the special unitary group SU(2).

If one uses the matrix identity

$$\log \mathrm{Det}\, A = \mathrm{Tr} \log A \tag{135}$$

one can say that

$$\mathrm{Tr} \log U = 0 \tag{136}$$

Hence, we should be able to represent U in the exponential form

$$U = \exp\left(iG\right) \tag{137}$$

where G is a traceless 2×2 matrix. The factor i in front of G is customary. The fact that $UU^\dagger = U^\dagger U = 1$ then demands that

$$G = G^\dagger \tag{138}$$

That is, G is a traceless and Hermitian 2×2 matrix. It has the general form

$$G = \begin{pmatrix} c & a - ib \\ a + ib & -c \end{pmatrix} = a \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + c \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{139}$$

which you should recognize as the Pauli matrices. So the most general form of the element of the SU(2) group is

$$U(\boldsymbol{\theta}) = \exp\left(i\boldsymbol{\theta} \cdot \boldsymbol{\tau}\right) \tag{140}$$

where

$$\boldsymbol{\tau} = \frac{1}{2}\boldsymbol{\sigma} \tag{141}$$

are called the *generators* of the SU(2) group or the Lie algebra of SU(2).

Now we repeat the E&M gauge theory consideration. The Schrödinger equation

$$i\partial_t \Psi = -\frac{\nabla^2}{2m}\Psi \tag{142}$$

is obviously invariant under the transformation

$$\Psi = U^\dagger \Psi' \tag{143}$$

when $U^\dagger$ is constant. We are using $U^\dagger$ here to conform with the most textbook definition of the gauge tranformation. It does not matter whether you use U or $U^\dagger$. But it is convenient to agree on the convention.

What about the case when U is not constant but varies from a space-time point to another? In that case,

$$\begin{aligned}\partial_\mu \Psi &= \partial_\mu (U^\dagger \Psi') \\ &= (\partial_\mu U^\dagger)\Psi' + U^\dagger \partial_\mu \Psi' \\ &= U^\dagger \left(\partial_\mu + U\partial_\mu U^\dagger\right)\Psi' \end{aligned} \tag{144}$$

Now since $U^\dagger U = 1$,

$$U\partial_\mu U^\dagger = -(\partial_\mu U)U^\dagger \tag{145}$$

Hence $U\partial_\mu U^\dagger$ is anti-Hermitian. Furthermore, in the definition of the derivative

$$\begin{aligned}U\partial_\mu U^\dagger &= \lim_{h\to 0} U(x)\frac{U^\dagger(x) - U^\dagger(x-h)}{h} \\ &= \lim_{h\to 0} \frac{1 - U(x)U^\dagger(x-h)}{h}\end{aligned} \tag{146}$$

the matrix $U(x)U^\dagger(x-h)$ is an element of the SU(2) group. Hence, it must have a representation

$$U(x)U^\dagger(x-h) = \exp\left(i\boldsymbol{\theta}\cdot\boldsymbol{\tau}\right) \tag{147}$$

Furthermore, we must have $\boldsymbol{\theta} \to 0$ as $h \to 0$. Therefore, expanding the exponential we get

$$U\partial_\mu U^\dagger = \lim_{h\to 0}\frac{i\boldsymbol{\theta}}{h} \tag{148}$$

and it is an element of the SU(2) algebra.

Again, for free particle wavefunctions, these transformations cannot matter. One can always redefine the phase of the wavefunction to absorb the "pure gauge" vector potential of the form $U\partial_\mu U^\dagger$. Now consider an analogue of electromagnetic force acting on this system. We can again divide the real and the imaginary part of ψ_1 and ψ_2 and their changes in space and time. However, that is overly complicated and does not really give one any additional insight. We'll simply assert that all that can be done and the result is "covariant derivative"

$$\partial_\mu \Psi \to D_\mu \Psi = \left(\partial_\mu + iA^a_\mu \tau_a\right)\Psi \tag{149}$$

The gauge transformation then yields

$$\begin{aligned} D_\mu \Psi &= D_\mu U^\dagger \Psi' \\ &= \left(\partial_\mu + iA^a_\mu \tau_a\right) U^\dagger \Psi' \\ &= U^\dagger \left(\partial_\mu + iUA^a_\mu \tau_a U^\dagger + U\partial_\mu U^\dagger\right)\Psi' \end{aligned} \tag{150}$$

Or with $A_\mu = A^a_\mu \tau_a$,

$$\begin{aligned} A'_\mu &= UA_\mu U^\dagger - iU\partial_\mu U^\dagger \\ &= UA_\mu U^\dagger + i(\partial_\mu U)U^\dagger \end{aligned} \tag{151}$$

The next thing to do is to demand that the physics of Ψ as well as the physics of A_μ should not depend on U. This would be the case if again the physical force is represented by the curvature tensor

$$G_{\mu\nu} = -i[D_\mu, D_\nu] \tag{152}$$

where the communtator now includes the matrix part as well. Because $\Psi = U^\dagger \Psi'$,

$$D_\mu \Psi = U^\dagger D'_\mu \Psi' \tag{153}$$

or

$$UD_\mu U^\dagger \Psi' = D'_\mu \Psi' \tag{154}$$

Hence

$$D'_\mu = UD_\mu U^\dagger \tag{155}$$

and that means

$$G'_{\mu\nu} = -i[D'_\mu, D'_\nu] = -iU[D_\mu, D_\nu]U^\dagger = UG_{\mu\nu}U^\dagger \tag{156}$$

Hence, as long as the dyanmics depends only on the trace of the products of D_μ the dynamics is independent of U. The QCD Lagrangian

$$\mathcal{L} = \bar{q}_{f,a} i \not{D} q_{f,a} - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a \tag{157}$$

is exactly in that form.

Now there isn't much that is really specific to SU(2) in the above analysis. The only thing that is specific to SU(2) is the appearance of the Pauli matrices. If you had N different species of particles which are identical but in one aspect, then we will have the SU(N) gauge theory with $N^2 - 1$ different A_μ. For the strong nuclear force, the number turns out to be $N = 3$. This was discovered because of the presence of the Δ^{++} baryon. It contains 3 u quarks and has the 3/2 spin. One can then show that the wavefunction is symmetric. This violates the Pauli Exclusion principle. The only way out is that there are actually 3 different kinds of u quark. A red u quark, a green u quark and a blue u quark. They are identical in all other aspects, but they do differ in the "color charge". The the wavefunction can be properly anti-symmetrized by anti-symmetrizing the color labels. In SU(3), there are 8 generators

$$\tau_a = \frac{\lambda_a}{2} \tag{158}$$

where λ_a are the Gell-Mann matrices.

So, what is the "color" then? Of course, it is not really the literal red, green and blue colors. The name color (or Chromo) was attached to it because it has 3 components just like the 3 fundamental colors and also because it turned out that the observed hadrons are all in the color neutral (white) state. The "color" simply denotes the 3 different states which are distinguished by the eigenvalues of the 2 diagonal Gell-Mann matrices

$$\lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{159}$$

All the terms in the Standard Model are generated this way.

Now one should ask: Can we see color? The analogous of color in the Abelian gauge theory is simply the electric charge. We can certainly see and feel the electric charge. But what about colors? Can we see them? To answer this question, one should answer this question first: What do we mean by "see"? Again, going back to the E&M the way we "see" the effect of the EM fields is to experience the electric and the magnetic forces. That is, we detect the presence of electric charges, moving or stationary by having a test charge and "see" it reacting.

In this regard, the important fact is that the gauge transformation cannot change the fact that the electron has the negative charge and the positron has the positive charge. This is because in this case, the number of different charges is really 1. The other one is the charge of the anti-particle or the anti-charge.

Now consider the color. The u quark for instance, should really be represented by the 3D vector

$$u = \begin{pmatrix} u_R \\ u_G \\ u_B \end{pmatrix} \tag{160}$$

In this case, the RGB are all charges. The anti-u-quark has the anti-color components $\bar{u} = (\bar{u}_{\bar{R}}, \bar{u}_{\bar{G}}, \bar{u}_{\bar{B}})$. In this case of 3 charges, a gauge transformation

$$u' = Uu \tag{161}$$

does mix up colors. The "red" component of u' is the linear combination of the old red, green and blue. Yet, the physics should remain the same. Hence, the label "color" has meaning only within the microscopic processes within a color neutral object and physical objects which should not depend on what I call "red" must be color neutral objects.

A similar comments goes to the field strength. The reason we can "feel" the electromagnetic force is that under the gauge transformation, they do not change at all. But in non-Abelian gauge theory, the field strength tensor changes:

$$G'_{\mu\nu} = UG_{\mu\nu}U^\dagger \tag{162}$$

The jargon is: They color-rotates. Hence, you cannot directly observe them. At the end, you can only detect the result of these forces acting on the color-neutral objects.

The color neutral objects that can be made out of quarks and anti-quarks are $\bar{q}q$ and $\epsilon_{abc}q^a q^b q^c$. The combination $\bar{q}q$ is colorless (gauge invariant) because of the unitarity, $U^\dagger U = 1$ and the second combination is colorless because of the matrix identity

$$\epsilon_{abc}U^{aa'}U^{bb'}U^{cc'} = \epsilon_{a'b'c'}\mathrm{Det}U \tag{163}$$

and the fact that $\mathrm{Det}U = 1$.

Asymptotic freedom

One of the most imporant aspect of non-Abelian gauge theory is how the strength of the gauge coupling behaves as the interaction energy scale changes. Consider a bare red quark propagating. Suppose a blue quark wants to interact with it. It can do so by emitting a gluon that carries $B\bar{R}$ color and becoming a red quark. The red quark can then absorb the gluon and become a blue quark. In this way, they have exchanged not only the energy and momentum but also the color. In this "classical" picture of color interaction, there is no reason why the strength of the interaction has to depend on the size of the exchanged momentum. Of course, gluon exchange does not have to exchange color as there are two gluons that are color neutral (corresponds to λ_3 and λ_8).

Quantum mechanically, however, the story is different. Now, before we talk about the color charges, let's talk about the electric charge first. Say an electron interacting with another electron. In this case, the interaction is mediated by a photon. Quantum mechanically, vacuum is not really empty. In fact, it is teeming with the briefly lived electron-positron pairs. Recall that the photon – electromagnetic wave – can only see the objects which are larger than the de Broglie wavelength. Hence, an exchange photon with the wavelength λ feels not only the electric charge of the original electron but also all the e^+e^- pairs bubbling in and out of the vacuum inside the wavelength. Because of the electric field of the original electron, these pairs prefer the orientation in which e^+ is towards the original electron and e^- is slightly outside. This is very analoguous to what happens in a polarizable medium when a charge is dropped into it. Now recall from E&M that in the polarizable medium with the dielectric constant $\epsilon > \epsilon_0$, the Coulomb field is given by

$$V = \frac{1}{4\pi\epsilon}\frac{q}{r} = \frac{1}{4\pi\epsilon_0}\frac{(\epsilon_0/\epsilon)q}{r} \tag{164}$$

which reduces the apparent charge. This is because in any size sphere containing the point charge, there is bound to be more opposite charge part of the dipole inside the sphere than outside. Hence, the net charge of any sphere is smaller than the original charge.

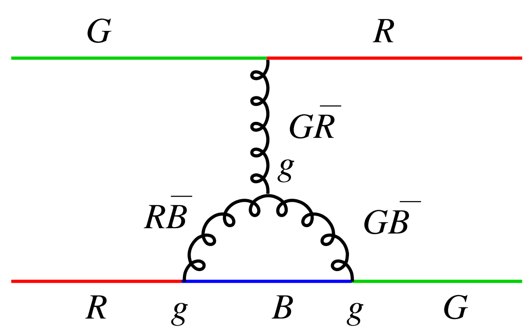
In the linear dielectric medium where

$$\mathbf{P} = \chi\epsilon_0\mathbf{E} \tag{165}$$

the net charge is always reduced by a factor (ϵ_0/ϵ). In the vacuum, the polarization effect due to the virtual e^+e^- pair "dipole" is a little more complicated. Nevertheless, it should be clear that as one approaches the original charge, one is bound to see more and more of the original charge strength. That is, the electromagnetic coupling constant α_{EM} becomes stronger as the distance scale decreases or the momentum scale increases.

When there are more than a single kind of charges, however, this picture completely changes. When an electron encounters an e^+e^- vacuum excitation, it may annihilate with the positron and briefly become a photon and an electron. The point here is that once you start with an electron your system always constains an electron no matter what happens.

Now consider a red quark propagating in vacuum. Suppose it encouters a $B\bar{B}$ vacuum fluctuation. It can then briefly become $R\bar{B}$ gluon and a blue quark. Now suppose a green quark is trying to interact with this red quark by emitting a $G\bar{R}$ gluon. If the wavelength of the gluon is large, then it will see the whole $q_B + g_{R\bar{B}}$ system and a net R charge. However, if the wavelength of the gluon is sufficiently small, then this $G\bar{R}$ gluon will either see the blue quark or the $R\bar{B}$ gluon. It can't interact with the blue quark any more. It can interact with the $R\bar{B}$ gluon which should then become the $G\bar{B}$ gluon and then it goes back to the original quark which now is green.



But this process now involves 3 interactions instead of just one. Hence, the effective interaction strength is *lower* than the case where the green quark could directly interact with the red charge as a whole. This is one way to understand why as the interaction energy (momentum) goes up, the smaller the interaction strength becomes. In effect, what happens is that the original red charge is no longer just associated with the quark. It is effectively spread out in space. The large wavelength (low momentum) interaction sees the whole spread-out charge with the net charge R and hence the interaction is strong. On the other hand, as the wavelength becomes smaller and smaller (higher momentum), the gluon containing a $\bar{R}$ sees smaller and smaller volume which contains less and less red charge. (In contrast, the electron charge is always located at the position of the original electron.) Hence as the interaction energy goes up, the interaction strength goes *down*. This property is known as the asymptotic freedom. Note that this happens because unlike the photon, gluons carry color charges. Hence, by emitting a gluon, a quark can actually change its charge.

In perturbative QCD, this can be calculated:

$$\begin{aligned} \alpha_S(Q^2) = \frac{g^2}{4\pi} &= \frac{1}{((11N_c - 2n_f)/12\pi)\ln(Q^2/\Lambda_{\rm QCD}^2)} \\ &= \frac{1}{((33 - 2n_f)/12\pi)\ln(Q^2/\Lambda_{\rm QCD}^2)} \end{aligned} \tag{166}$$

where Q^2 is the virtuality of the exchanged gluon and $\Lambda_{\rm QCD} \approx 0.2\,{\rm GeV}$ is the intrinsic QCD scale that determines the typical energy scale of hadrons. It is determined by

$$\alpha_S(M_Z^2) \approx 0.118 \tag{167}$$

Now in any field theory calculation, the basic calculational tool is the perturbation theory where physical quantities are expanded in terms of the small coupling constant. What the above fact about QCD is telling us is that the high energy part of any QCD process may be calculated within perturbation theory. However, the low energy part of any QCD process is out of our reach as the strength of the interaction for that part is nowhere near small.

Hadron-Hadron Cross Section

Now consider a hadron-hadron collision at high energy. There are two known facts. One, experimentally the total cross-section behaves like

$$\sigma \sim a + b \ln \sqrt{s} + c(\ln \sqrt{s})^2 \tag{168}$$

Two, perturbative calculation of quark-quark scattering cross-section behaves like, for instance,

$$\frac{d\sigma_{ud \to ud}}{dt} = \frac{4\pi\alpha_S^2}{9} \left(\frac{s^2 + u^2}{s^2 t^2} \right) \tag{169}$$

where we have introduced the Mandelstam variables

$$s = (p + k)^2 \tag{170}$$

and

$$t = (p - p')^2 = (k - k')^2 \tag{171}$$

and

$$u = (p - k')^2 \tag{172}$$

where p and k are the incoming momenta and p' and k' are the outgoing momenta satisfying the energy-momentum conservation

$$p + k = p' + k' \tag{173}$$

These variables satisfy, in the massless limit

$$\begin{aligned} s + t + u &= p^2 + k^2 + 2pk + p^2 + (p')^2 - 2pp' + p^2 + (k')^2 - 2pk' \\ &= 2pk - 2pp' - 2pk' \\ &= 2p(k - p' - k') \\ &= -2p^2 = 0 \end{aligned} \tag{174}$$

Hence the total cross-section is

$$
\begin{aligned}
\int dt \frac{d\sigma_{ud\to ud}}{dt} &= \int dt \frac{4\pi\alpha_S^2}{9}\left(\frac{s^2+(s+t)^2}{s^2t^2}\right) \\
&= \int dt \frac{4\pi\alpha_S^2}{9}\left(\frac{2s^2+t^2+2st}{s^2t^2}\right) \\
&= \frac{4\pi\alpha_S^2}{9}\int_{-s}^{0} dt\left(\frac{2}{t^2}+\frac{1}{s^2}+\frac{2}{st}\right) \\
&= \frac{4\pi\alpha_S^2}{9}\left(\frac{2}{|t_{\min}|}-\frac{|t_{\min}|}{s^2}-\frac{1}{s}-2\frac{\ln(s/|t_{\min}|)}{s}\right) \\
&\approx \frac{4\pi\alpha_S^2}{9}\frac{2}{|t_{\min}|}
\end{aligned}
\tag{175}
$$

where $t_{\min}$ is the IR cut-off. This does not behave at all like what the experiment measures unless somehow the cut-off behaves

$$
|t_{\min}| \sim 1/(a + b\ln s + c\ln^2 s) \tag{176}
$$

But there is no good reason the cut-off should behave this way.

Note that in the CM frame,

$$
\begin{aligned}
s &= (p+k)^2 \\
&= 2E_pE_k(1-\cos\theta_{pk}) \\
&= 4E_p^2
\end{aligned}
\tag{177}
$$

and

$$
\begin{aligned}
t &= (p-p')^2 \\
&= -2E_pE_{p'}(1-\cos\theta) \\
&= -\frac{s}{2}(1-\cos\theta_{pp'})
\end{aligned}
\tag{178}
$$

The maximum value of $|t|$ is therefor s.

These two facts combined together indicates that the total cross-section of hadron-hadron collisions is dominated *not* by pQCD interaction. It must be then dominated by the interaction of the soft part of the hadrons.

For the exprimentally measured cross-section, Richard Feynman argued as follows: Suppose the amplitude to emit small $x = 2p/\sqrt{s}$ gluon is $\dfrac{1}{x^{1+\lambda}}$.

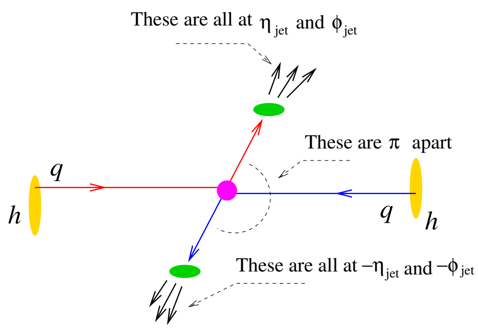
Then the soft-soft scattering cross-section is

$$\sigma \sim \left| \int \frac{dx_a}{x_a^{1+\lambda}} \right|^2 \left| \int \frac{dx_b}{x_b^{1+\lambda}} \right|^2 \sim (E_a E_b)^{2\lambda} \sim s^{2\lambda} \tag{179}$$

For small λ, this behaves like a series in $\ln s$. Hence, there must be a lot of soft gluons in a high energy hadron and they dominate the collision process.

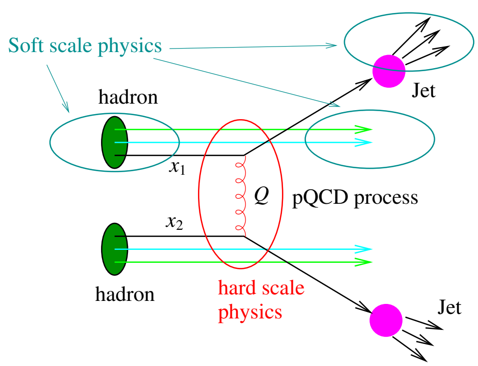
What about the perturbative parton-parton scattering, then? Does it have no relevance in describing hadron-hadron collisions? Fortunately, the large x partons once in a great while do undergo a large angle scattering. This part of the scattering cross-section, then should be describable by perturbative QCD (pQCD). These are the *jets*. In most of the hadron-hadron collisions, jets do not appear. Only in a small fraction of events, hard scatterings ocurr and jets appear. When this happens, it is clear that the jets are very distinct from what's called the underlying events (UE). They appear as a clean shower of collimated hadrons whose total energy can be a good fraction of $\sqrt{s}$. Since the origin of the jets is the hard parton-parton scattering, one can think of a jet as a hard colored parton that eventually become a shower of hadrons as it tries to become color neutral.

What is a jet?



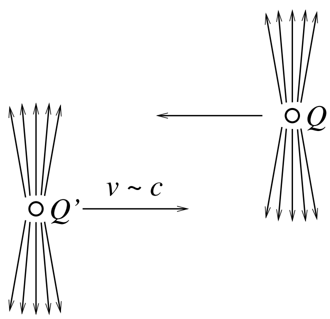
- A jet is a phenomenon where a lot of final state energy is concentrated in a small angle around a common axis
- Origin: Hard collisions of partons $\Longrightarrow$ pQCD applies
- Usually dijet, sometimes triple-jet (Radiation of a hard gluon at a large angle)

Applying QCD



- *Only* the hard scattering is hard pQCD. Everything else is soft.

How to think about the initial state factorization – QED analogy



- Weizsäcker-Williams field – Highly contracted in the z direction
- Coulomb potential in the rest frame of the charge

$$\varphi = Q/|\mathbf{r}|$$

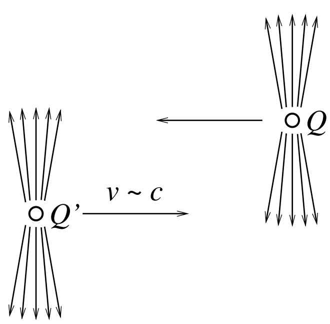
- In the moving frame

$$A^{\mu}(x') = \Lambda^{\mu}_{\nu} A^{\nu}(x(x'))$$

- The coordinate in the moving frame $x' = (t, x, y, z)$. This corresponds to the rest frame position $x = (t\gamma - z\gamma v, x, y, z\gamma - t\gamma v)$.

Jeon (McGill) | Hard Probes | Beijing 2013 | 27 / 139

How to think about the initial state factorization – QED analogy



- Weizsäcker-Williams field – Highly contracted in the z direction
- Coulomb potential in the rest frame of the charge

$$\varphi = Q/|\mathbf{r}|$$

- In the moving frame

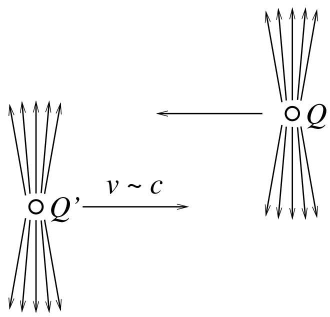
$$A^{\mu} = \frac{Q(\gamma, 0, 0, \gamma v)}{\sqrt{(z - vt)^2\gamma^2 + \Delta\mathbf{x}_{\perp}^2}}$$

- Pure gauge in the $v \to 1$ limit

$$A^{\mu} \approx \frac{Q(1, 0, 0, 1)}{|z - vt|} = Q\partial_{\mu} \ln |z - vt|$$

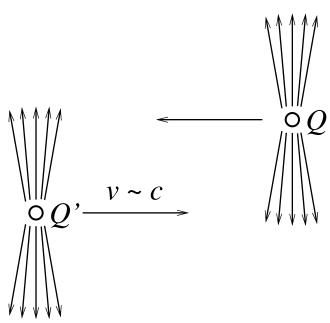
Jeon (McGill) | Hard Probes | Beijing 2013 | 28 / 139

How to think about the initial state factorization – QED analogy



- Weizsäcker-Williams field – Highly contracted in the z direction
  $F^{\mu\nu} \approx 0$ unless $z \approx vt$
- In the rest frame: Coulomb field is made up of space-like virtual photons
  $q^{\mu} q_{\mu} = -\mathbf{q}^2$ with $q_0 = 0$.
- In the Lab frame:
  $q'^{\mu} = (q^z \sinh\eta, \mathbf{q}_{\perp}, q^z \cosh\eta)$
- For large η,
  $|\Delta E| = |q^- - |\mathbf{q}|| \sim e^{-\eta}\mathbf{q}^2/q_z$
  $\Longrightarrow \Delta t \sim 1/|\Delta E| \sim e^{\eta} q_z/\mathbf{q}^2 \Longrightarrow$ virtual photons look almost like real photons.

How to think about the initial state factorization – QED analogy



- Weizsäcker-Williams field – Highly contracted in the z direction
  $F^{\mu\nu} \approx 0$ unless $z \approx vt$
- To a first approximation, the approaching particles *do not* know about each other until they are on top of each other.
- Initial photon momentum distribution *factorizes*: $F(x_1, x_2) = f(x_1)f(x_2)$
  but this is not exact.
- In QCD, color neutrality of hadrons help.

So how does one understand the hadron structure? Let's think about a proton. In the rest frame of the proton, It is a bound state of 3 quarks, (uud). That means, if I had magical microscope that can look inside the proton, I will see that 3 point like quarks interacting with force (potential energy) mediated by gluons. That is, the whole system is made of 3 real particles (quarks) and a whole bunch of virtual particles (gluons). These gluons, however, are not real. It's just like a Coulomb potential. The Coulomb potential is made of photons, yet, a charged particle at rest does not shine. To shine, it must accelerate, thus interrupting the static state and make the signal propagate.

Now just as the Coulomb field is made of virtual photons which only has the spatial momentum, the force field binding the 3 quarks together as a proton is also made up of gluons that only has the spatial momentum. Now, suppose we boost the proton to a high energy. That is, $\gamma \gg 1$. In the moving frame, the energy is

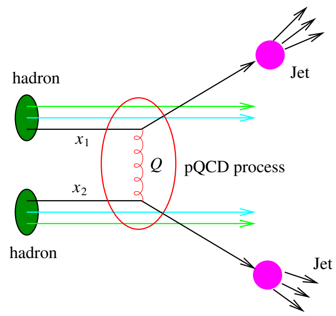
$$\begin{aligned}
(q^0)' &= \gamma q^0 + \gamma v q^z \\
&= \gamma v q^z \\
(q^z)' &= \gamma q^z + \gamma v q^0 \\
&= \gamma q^z
\end{aligned} \tag{180}$$

since $q^0 = 0$ in the rest frame. Notice that in this frame, even though the momentum remains space-like, the gluon appears to have both the energy and momentum and if γ is sufficiently large, then this virtual particle would appear to be almost real. As such, a highly boosted hadron can be thought of being composed of the valence quarks (the original quarks in the rest frame) and a cloud of virtual gluons promoted to be (almost) real. This is called the Color Glass Condensate (CGC) picture.

Now special relativity tells us that a highly boosted system exhibits two important features. It is Lorentz contracted and it is time dilated. Hence, in this frame, the hadron does not appear as a round object, but as a thin 2D sheet. Furthermore, due to the time dilation effect, the movement of particles in the system slows down by a factor of γ. In effect, they are frozen. Now a hadron is not just a classically bound state of quarks. It is also quantum mechanical system where vacuum bubbles of $q\bar{q}$ pairs and gluons pairs pop in and out all the time. Hence, in addition to the valence quarks and cloud of virtual gluons, the highly boosted hadron also contains sea-quarks that are frozen in before they had a chance to go back to vacuum.

Factorization

Factorization Theorem



- Hadron-Hadron Jet production scheme:

$$\frac{d\sigma}{d\hat{t}} = \int_{abcd} f_{a/A}(x_a, Q_f) f_{b/B}(x_b, Q_f) \times \frac{d\sigma_{ab\to cd}(Q_R)}{d\hat{t}} D_{C/c}(z_C, Q'_f)$$

- pQCD is used to calculate $\sigma_{ab\to cd}$ and the *evolution* of the parton distribution functions and the fragmentation functions.

$f_{a/A}(x, Q_f)$: PDF, $\frac{d\sigma_{ab\to cd}(Q_R)}{d\hat{t}}$: PQCD x-section, $D_{C/c}(z, Q'_f)$: FF

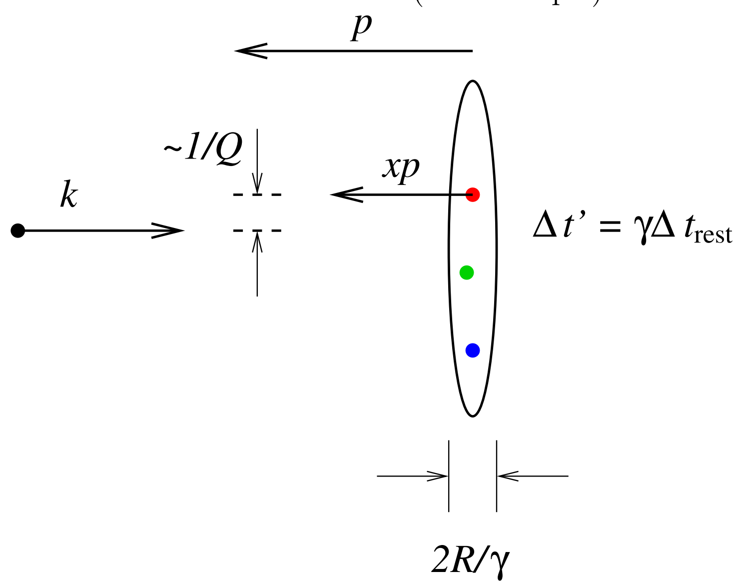
Factorization Theorem

- How realistic pQCD calculations are done

$$\sigma_{hh'\to C+X} = \int_{abcd} dx_1 dx_2 f_{a/h}(x_1, Q_f) f_{b/h'}(x_2, Q_f) \sigma_{ab\to cd}(Q_R) D_{C/c}(z_C, Q'_f)$$

- pQCD controls the *evolutions* of $f_{a/h}(x_1, Q_f)$ and $D_{C/c}(z_C, Q'_f)$. But pQCD cannot determine the initial data because this is dominated by IR processes.
- pQCD *can* calculate $\sigma_{ab\to cd}(Q_R)$ when the renormalization scale Q_R can be set high (that is, when $\sqrt{s}$ is large)

As explained above, only the hard part of QCD interactions are analytially calculable using perturbation theory. But any interactions that involves hadrons either in the initial state or in the final state must have soft interactions just because hadrons are soft objects. Then how can we calculate anything if we can't really calculate the soft part of QCD from first principles? Fortunately, quantum mechanics and relativity help us in this regard. Consider a collision between an electron (to be simple) and a hadron.



As explained above, the hightly boosted hadron has a very thin longitudinal extent and it is full of quarks, anti-quarks and gluons which are basically static. Hence, the collision of electron and the hadron would proceed in the following way.

1. Before the collision, the electron and the hadron do not know each other's existence.

2. An electron can only directly interact with a quark by exchanging a virtual photon with the virtuality Q.

3. If only the single parton-electron interaction matters, then this process can be described with the probability of finding a quark within $1/Q$ with a certain momentum fraction x of the original hadron.

4. If quarks are randomly distributed in a disk of radius R, then the probability density that two particles are at the distance r is roughly

the area they span divided by the total area.

$$P(r < d) \sim \frac{d^2}{R^2} \sim \frac{1}{(QR)^2} \tag{181}$$

Hence, if $Q \gg 1/R$, then this probability is small and we can use the single particle picture.

5. The interaction happens in the time/length scale of $\tau_{\text{crossing}} = R/\gamma$ because that's the longitudinal size of the hadron. The interactions that produces hadrons happen in the time scale of $1/\Lambda_{\text{QCD}} \sim 1\,\text{fm}$ in the rest frame. In the moving frame, this is also dilated by γ.

6. By the time hadrons start to form from the interrupted remnant of the original hadron the projectile electron is long gone.

Hence, *provided* that there is a large scale difference between the hadronic scale $\sim 1\,\text{GeV}$ and the collision energy $\sqrt{s}$, one can construct a probabilistic picture that starts with the probability of finding a parton (quark or gluon) inside the highly boosted hadron with a certain momentum fraction x, the probability to have a hard collision, and then the probability to form a certain species of hadrons after the electron is gone.

So here is the factorization formula again for the hadron-hadron collisions:

$$\sigma_{hh' \to C+X} = \int_{abcd} f_{a/h}(x_1, Q_f) f_{b/h'}(x_2, Q_f) \sigma_{ab\to cd}(Q_R) D_{C/c}(z_C, Q'_f) \tag{182}$$

4 Heavy Ion Collisions and QGP production

I think we have enough background now to talk about heavy ion collisions. A heavy ion collision is a collision between two large nuclei. The nuclei we have used so far include Au^{197}_{79}, Pb^{208}_{82}, Cu^{64}_{32}, U^{238}_{92} etc.

To think about the nucleus before collisions, we again appeal to the Color Glass Condensate framework. A nucleus is a collection of protons and neutrons. A highly boosted nucleus can be viewed as a collection of large x partons ($\sim$valence quarks) and the small x partons (sea-quarks and Coulomb like gluons) radiated by the large x partons.

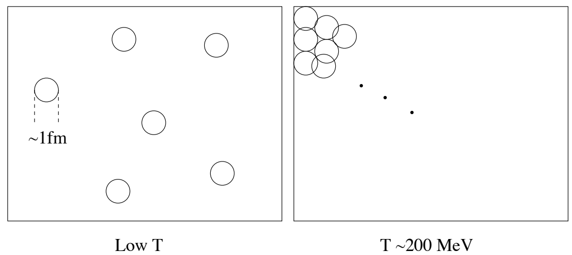
So in the beginning, two think pancakes approach. Each pancake is made up of few hard partons and many soft partons. The density combined with the running QCD coupling then tells you that majority of the interaction is going to be between two clouds (CGC gluons) of soft partons. This part of interaction occurs every time there is a collision regardless of what hard partons are doing. Within the time scale of soft interaction $\sim 1/\Lambda_{\rm QCD}$, these part of interaction is going to become chaotic and thermalized. This becomes the Quark-Gluon Plasma if there is sufficient energy in the system.

While this is going on, once is a great while, two hard partons from two nuclei may collide at a large angle exchaging momentum of $O(\sqrt{s})$. As we have argued before, this process is perturbative. The complecation here is that this is happening while the soft part is thermalizing to become QGP and the hard partons subsequently need to traverse QGP.

So why is this useful? This is because our primary goal is to study QGP. The production of two hard partons which will eventually become jets follow pQCD and calculable. We can also measure it in the proton-proton collisions which presumably do not produce QGP droplets. Hence by comparing the properties of jets (hard partons) that emerges out of heavy ion collisions (through QGP) and those of the pp jets, we can learn a lot about the QGP medium.

5 Properties of QGP

Another estimate of $T_{\text{transition}}$



- Density: Consider a pion gas.

$$n = 3 \int \frac{d^3p}{(2\pi)^3} \frac{1}{e^{E_p/T} - 1} = 0.37\, T^3$$

As T becomes larger, more and more pair creation results.

- Inter particle distance:

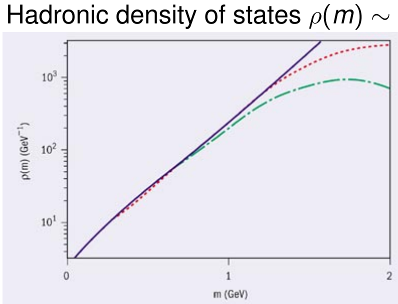
$$l_{\text{inter}} = n^{1/3} = 1.4/T$$

At $T = 200\,\text{MeV}$, $l_{\text{inter}} \approx 1.4\,\text{fm} \approx r_\pi$

Jeon (McGill) Hard Probes Beijing 2013 46 / 139

Hagedorn Temperature

Hadronic density of states $\rho(m) \sim e^{m/T_H}$:



The smoothed mass spectrum of hadronic states as a function of mass. Experimental data: long-dashed green line with the 1411 states known in 1967; short-dashed red line with the 4627 states of 1996. The solid blue line represents the exponential fit yielding T_H=158 MeV. *CERN Courier, Sept, 2003*

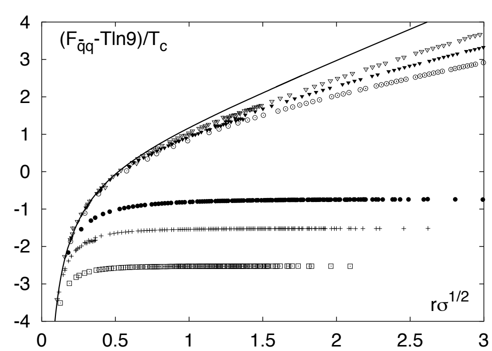
- $\sum_m \int_p \rho(m) e^{-E_p/T}$: Not well defined when $T > T_H$ for hadronic matter.
- Phase transition around T_H: Hagedorn temperature $\approx 160\,\text{MeV}$

Jeon (McGill) Hard Probes Beijing 2013 47 / 139

Story so far

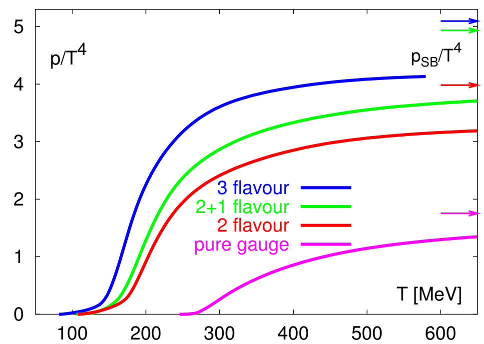
- Perturbative calculation possible much above $\mu = \Lambda_{\mathrm{QCD}}$
- $\mu \sim T$ at high T
- If T is much above the binding energy of hadrons
  $\Longrightarrow$ Deconfinement
- At high enough T, the system is a plasma of weakly interacting quarks and gluons
- All the above arguments are plausible but not a proof

Lattice QCD Evidence



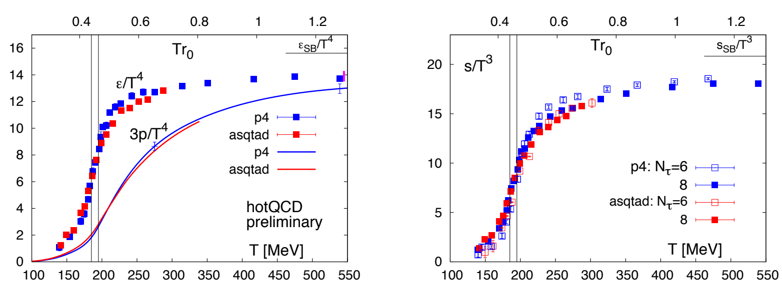
- F. Karsch, hep-lat/0403016. The color averaged heavy quark free energy at temperatures $T/T_c = 0.9,\ 0.94,\ 0.98,\ 1.05,\ 1.2,\ 1.5$ (from top to bottom) obtained in quenched QCD.

Lattice QCD – QGP



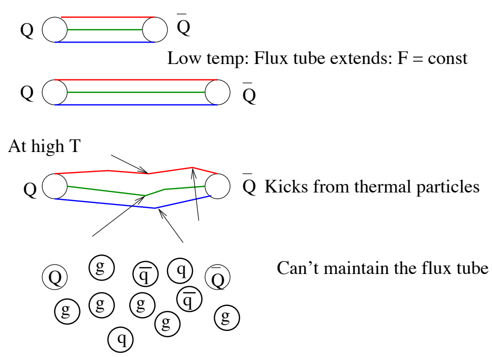
- QCD is an asymptotically free theory - High $T \Longrightarrow$ Free quarks and gluons
- Phase transition happens – Hadrons should 'melt' at around $T = 170\,\text{MeV} = 2 \times 10^{12}\,\text{K}$ [F.Karsch et al.] "Cross-over"

Lattice QCD Evidence of QGP



- From HotQCD Collaboration (C. DeTar, arXiv:0811.2429)
- "Cross-over" between 185 - 195 MeV

What it means



http://www.physics.adelaide.edu.au/

~dleinweb/VisualQCD/Nobel/

Expected properties

- High number density

$$n \approx (24+16)\int \frac{d^3p}{(2\pi)^3}\, e^{-p/T} \approx 4\,T^3$$
$$= 4\left(\frac{T}{200\,\text{MeV}}\right)^3 \text{fm}^{-3}$$

- High energy density

$$\varepsilon \approx (24+16)\int \frac{d^3p}{(2\pi)^3}\, p\, e^{-p/T} \approx 12\,T^4$$
$$= 2.4\left(\frac{T}{200\,\text{MeV}}\right)^4 \text{GeV/fm}^3$$

Simple Estimates

With $\hbar = c = 1$

- 1 mole of hydrogen atom: 6.02×10^{23} atoms $= 1$ g (Avogadro's number)
- 1 hydrogen atom $m_p \approx (1/6) \times 10^{-23}\,\text{g} = (1/6) \times 10^{-26}\,\text{kg}$
- $m_p = 940\,\text{MeV} \approx 1\,\text{GeV}$
- $E = mc^2$: $1\,\text{GeV} \approx (1/6) \times 10^{-26}\,\text{kg}$

$$\begin{aligned} 2.4\,\text{GeV}/\text{fm}^3 &= 0.4 \times 10^{-26}\,\text{kg}/(10^{-13}\,\text{cm})^3 \\ &= 0.4 \times 10^{-26+39}\,\text{kg/cm}^3 \\ &= 4 \times 10^{12}\,\text{kg/cm}^3 \end{aligned}$$

- Typical human: $\sim 100\,\text{kg}$

$$2.4\,\text{GeV}/\text{fm}^3 \sim 4 \times 10^{10}\,\text{human/cm}^3$$

Simple Estimates

With $\hbar = c = 1$

- Another way of looking at the energy density

$$2.4\,\text{GeV}/\text{fm}^3 = 4 \times 10^{12}\,\text{kg}/\text{cm}^3$$

- Restoring $c = 3 \times 10^8$ m/s,

$$2.4\,\text{GeV}/\text{fm}^3 = 4 \times 10^{12} \times (9 \times 10^{16})\,\text{J/cm}^3 = 3.6 \times 10^{29}\,\text{J/cm}^3$$

- World energy consumption (2008):

$$144\,\text{pWh} = 144 \times 10^{15} \times 3.6 \times 10^3\,\text{J} = 5.2 \times 10^{20}\,\text{J}$$

- A cubic centimeter of QGP can power the world for about 70 million years.

Simple Estimates

With $\hbar = c = 1$

- Pressure $P \approx \epsilon/3$

$$P = 0.8\,\text{GeV}/\text{fm}^3 \approx 1.3 \times 10^{12}\,\text{kg}/\text{cm}^3 = 1.3 \times 10^{18}\,\text{kg}/\text{m}^3$$

- SI Unit for pressure: $\text{Pa} = \text{N/m}^2 = \text{kg/m/s}^2$
- Restoring $c = 3 \times 10^8$ m/s,

$$P \approx 1.3 \times 10^{18} \times (9 \times 10^{16})\,\text{kg}/\text{m/s}^2 \approx 10^{35}\,\text{Pa} \approx 10^{30}\,\text{atm}$$

How do you achieve high temperature?

- Temperature = energy (1 eV $\approx$ 12,000K)
- More usefully, the energy density:

$$\varepsilon = g \int \frac{d^3p}{(2\pi)^3}\, E_p\, e^{-E_p/T} \approx \frac{3g}{\pi^2} T^4$$

- To get high temperature: Get high energy density $\Longrightarrow$ Cram maximum possible energy into the smallest possible volume while randomizing the momenta $\Longrightarrow$ Relativistic heavy ion collisions.
- What to expect: $dN/d\eta$ and $dE/d\eta$ grow something like $(\ln s)^n$ with $n \sim 1$ $\Longrightarrow$ T should behave something like $(\ln s)^n$ with $n \sim 1$