Lecture Notes

Year 1 Semester 2

Ash Stewart

MSci Physics with Particle Physics and Cosmology

School of Physics and Astronomy
University of Birmingham

Table of Lectures

LC Classical Mechanics and Relativity 2

Tue 17 Feb 2026 11:00

Lecture 11 - Special Relativity I: Foundations and Time Dilation

1 Introduction

In this lecture:

- Recap from CMR1
- Einstein's two postulates
- Time dilation
- Galilean Transformations

Crucially, there are no special or absolute frames of reference - laws of motion are only sensible when we consider a frame of reference. The only slightly special frame is the one in which we are currently stationary.

Special relativity provides a theory of relative motion between inertial frames of reference. An inertial frame is one which is not accelerating relative to the other frames being considered.

2 Frames of Reference

We will generally consider two related frames of reference, denoted Σ and Σ'. Σ is our "stationary" frame, i.e the frame taken by an observer sat on earth. Σ' is our "moving frame", relative to the stationary observer and is moving with constant speed v.

Coordinates in the stationary frame are (x, y, z, t), while in the moving frame they add a prime, so are given by: (x', y', z', t').

We want to compare observations of position/experience of time/velocity in the moving frame with observations of the same categories in the stationary frame.

The frame in which an object is stationary is denoted its rest frame Σ_0.

3 Galilean Transformations

"Galilean Invariance": The laws of physics are invariant in all inertial (non-accelerating) frames.

"Galilean Transformation": Time is invariant and universal in all frames.

In Classical Mechanics, we treat time as invariant, but this stops being true in special relativity. We can transform between frames in such a manner that time is kept constant, this is a Galilean Transformation, but this is not true by default.

Consider our two reference frames Σ and Σ', where the latter moves with speed v relative to the former. Our coordinates in (x, y) become coordinates in (x', y').

An object moves at speed u' in the x-direction in reference frame Σ'. If we assume that time is invariant (as we're still currently doing classical physics without having introduced special relativity), then:

$$t = t'$$

And as the motion is entirely in the x-axis:

$$y = y'$$

After time t:

$$x = x' + \text{distance moved by } \Sigma' \text{ in time } t \text{ relative to } \Sigma$$

$$x = x' + vt$$

The object velocity is defined as:

$$u = \frac{dx}{dt} = \frac{d}{dt}(x' + vt)$$

$$= \frac{dx'}{dt} + v = u' + v$$

This holds if any only if time is the quantity we treat as being invariant. This agrees with our classical understanding.

3.1 Where does this break down?

Lets assume that the moving object is a photon, moving in Σ' with velocity c', hence, according to the previous derivation:

$$c = c' + v$$

Therefore observers in different frames will measure different values for the speed of light...oh no!

If we assume Galilean invariance, the laws of physics are the same in all reference frames, and yet the speed of light is included as a constant in many laws (i.e. electromagnetism). Therefore observers must measure the same speed of light in all reference frames.

This is a contradiction - we cannot assume that both time and the laws of physics are invariant.

3.2 Experimental Verification

The earth travels around the sun extremely quickly, and the sun is travelling even faster around the galactic centre. Therefore, the earth is moving in space.

If Galilean relativity is correct, we would measure the speed of light in one direction as $c + u$ and measure the speed of light in the other direction as $c - u$.

This was tested by the Michelson-Morley experiment, where incoming light was split by a half-silvered mirror. Half the light travels in one direction, and half the light is split off by 90°. They reflect off a pair of mirrors and are recombined in a splitter to be observed.

If the speed of light was different in different directions, the two beams would be out of phase and we would observe an interference pattern on combination. We would expect to see phase difference that varies with time, i.e. turning from destructive to constructive etc.

This was not observed and the interference pattern generated was constant. Therefore, the results showed no variation in the speed of light.[1]

4 The Solution - Einstein's Special Relativity

To fix this contradiction, Einstein came up with two postulates:

1. The laws of physics are the same in any inertial frame of reference.
    - This is the same as Galilean invariance and is easily believable.
2. The speed of light is constant in every frame of reference.
    - This was revolutionary and is much less intuitive.

The last postulate is difficult to understand intuitively, but makes everything work if we accept it as true!

[1] This is the same idea used in the LIGO experiment to discover gravitational waves, except this was used to show that the path length changed, and not the speed of light. If the path lengths changed, this was due to a gravitational wave (ripple in space time) propagating to the earth and interfering with the measurement.

4.1 Proving Time Dilation

Consider an observer on the earth in frame Σ. This observer watches two rockets both travelling side by side away from the earth in the x-direction with speed v. They are both travelling in x, with a constant y-difference y_0 between them.

Suppose the upper rocket A fires a laser beam directly at B (i.e. directly in the y-direction downwards).

Frame Σ: This is the rest frame of the earth (and the observer on earth) with coordinates (x, y).

Frame Σ': This is the rest frame of the rockets, with coordinates (x', y').

In the rockets rest frame Σ', the rockets are stationary and the time taken for a laser pulse to travel between them is:

$$t_0 = \frac{y_0}{c}$$

Or equivalently:

$$y_0 = ct_0$$

In the earth's rest frame Σ, the observer on the earth doesn't just see the light travelling in the y-direction. It also sees the light moving in the x-direction, as the whole Σ' frame containing the rockets relativistically move away.

Effectively, we have:

B moves distance x horizontally while it waits for the laser to hit it, and the length of the path of the laser beam in time t is given by D. If the time taken for light to reach B in Σ is t, then:

$$D = ct$$

As the speed of light is constant in all frames. As the rockets are moving away:

$$x = vt$$

This gives us a Pythagorean triangle, where:

$$D^2 = x^2 + y^2 = x^2 + y_0^2$$

And substituting in:

$$c^2t^2 = v^2t^2 + c^2t_0^2$$
$$t^2(c^2 - v^2) = t_0^2c^2$$
$$t = t_0\sqrt{\frac{c^2}{c^2 - v^2}}$$

Dividing through by c_2:

$$t = t_0\sqrt{\frac{1}{1 - \frac{v^2}{c^2}}}$$

Or:

$$t = t_0\gamma, \qquad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$\gamma > 1$ for all object speeds that don't exceed the speed of light (which is required), so the time observed in the frame Σ is longer than the "proper time" observed by the rocket.

5 The Twin Paradox

Consider two twins, an astronaut and a physics professor. The astronaut goes on a long trip close to the speed of light and reunites with his brother. Which brother is older?

Special relativity says that because there is no absolute frame of reference, the brother travelling away also sees the same problem in reverse, i.e. he sees the physics professor brother travelling away in the opposite direction at equal and opposite speed, i.e. they are the same age.

In practice, the astronaut is non-inertial and special rel isn't sufficient, so we can't solve it directly.

Thu 19 Feb 2026 15:00

Lecture 2 - Special Relativity II: Length Contraction

1 Proving Length Contraction

Consider a horizontal laser cavity. We fire a laser from the left of the cavity (point A) to the right of the cavity (point B). The laser bounces off a mirror at point B and bounces back.

The cavity has length L_0 as measured in the rest frame Σ_0. In this frame:

$$t_0 - \frac{2L_0}{c}, \qquad \text{Where } t_0 \text{ is the time taken for the A->B->A journey in } \Sigma_0$$

Now suppose the cavity moves at speed v relative to a stationary observer in a different frame (Σ).

The observer sees the cavity as having length L, how long does the A->B->A journey take as observed by a stationary observer in Σ.

$$t = t_{A\to B} + t_{B\to A}$$

The cavity is moving away from the observer, so the laser beam doesn't just travel L when going from $A \to B$. It must travel L plus the distance the whole cavity has moved away in this time (as B is moving away from the pulse). This extra distance is $vt_{A\to B}$, hence the distanced travelled when going from A to B is $L + vt_{A\to B}$.

As c is constant in all frames, we can say:

$$c = \frac{L + vt_{A\to B}}{t_{A\to B}}$$
$$t_{A\to B}(c - v) = L$$
$$t_{A\to B} = \frac{L}{c - v}$$

And now on the return journey from $B \to A$, as it travels back towards A, the cavity is moving in the same direction as the light, so A "catches up" to the pulse and results in a smaller required distance to be travelled, $L - vt_{B\to A}$

$$c = \frac{L - vt_{B\to A}}{t_{B\to A}} \implies t_{B\to A} = \frac{L}{c + v}$$

And returning to the total time:

$$t = t_{A\to B} + t_{B\to A}$$
$$t = \frac{L}{c - v} + \frac{L}{c + v}$$
$$= L\left(\frac{1}{c - v} + \frac{1}{c + v}\right)$$
$$= L\left(\frac{c + v + c - v}{(c - v)(c + v)}\right)$$
$$= L\frac{2c}{c^2 - v^2}$$
$$= \frac{2L}{c}\frac{c^2}{c^2 - v^2}$$
$$= \frac{2L}{c}\frac{1}{1 - \frac{v^2}{c^2}}$$

$$= \frac{2L}{c}\gamma^2$$

We now apply time dilation, which says that $t = \gamma t_0$, hence:

$$\gamma t_0 = \frac{2L}{c}\gamma^2$$

$$t_0 = \frac{2L}{c}\gamma$$

From the rest frame, we know that $t_0 = 2L_0/c$, so:

$$\frac{2L_0}{c} = \frac{2L\gamma}{c}$$

$$L_0 = L\gamma$$

As $\gamma > 1$, $\forall(u < c)$, the proper length measured from the rest frame will always be greater than the length measured by an observer, hence length is contracted for a moving object.

2 Example

In a particle accelerator, protons are accelerated to a measly $0.9c$. These protons pass through a tunnel of length 2km as viewed from the a laboratory rest frame at the accelerator.

Recall that:

$$t = \gamma t_0$$

$$L = \frac{L_0}{\gamma}$$

$$\gamma = \frac{1}{\sqrt{1-\beta^2}}, \qquad \text{where: } \beta = u/c$$

2.1 How long would the journey take, according to the rest frame?

Viewed in the lab frame, and $v = d/t$ so:

$$t = \frac{2 \times 10^3\text{m}}{0.9 \times 3 \times 10^8} = 7.4 \times 10^{-6}\text{s} = 7.4\mu\text{s}$$

No relativity needed!

2.2 How long would the journey take, according to the proton's frame?

We do need to consider relativity here, and use time dilation:

$$t = \gamma t_0$$

$$\gamma = \frac{1}{\sqrt{1-0.9^2}} = 2.3$$

$$t = \gamma t_0 \implies t_0 = \frac{t}{\gamma} = \frac{7.4\mu\text{s}}{2.3} = 3.2\mu\text{s}$$

2.3 How long is the tunnel, according to the proton's frame?

The rest frame now corresponds to the tunnel. It has proper length in its rest frame of 2km.

As viewed by the protons in their rest frame, it is the tunnel that is moving and is rushing towards them at $0.9c$. This is therefore a length contraction problem.

$$L = \frac{L_0}{\gamma} = \frac{2\text{km}}{2.3} = 870\text{m}$$

2.4 Checking Values

To check, we can reuse $v = d/t$ but this time in the proton rest frame.

$$d = 870\text{m}, \quad t = 3.2 \times 10^{-6}\text{s}$$

Hence

$$v = \frac{870\text{m}}{3.2 \times 10^{-6}\text{s}} = 2.7 \times 10^{8}\text{ms}^{-1} = 0.9c \text{ as required!}$$

3 Example II: Cosmic Rays

The highest energy cosmic rays are protons with massive energies $E \sim 10^{20}$eV, compared to the LHC with $E \sim 10^{12}$eV

This corresponds to a $\gamma = 10^{11}$, and our galaxy is $\sim 10^{20}$m across. [1]

As viewed on earth, these cosmic protons travel at $v \approx c$, hence,

$$t \approx \frac{d}{v} \approx \frac{10^{20}}{3 \times 10^{8}} = 3 \times 10^{11}\text{s} \approx 10^{5}\text{years}$$

As viewed by the protons:

$$t_0 = \frac{t}{\gamma} = \frac{3 \times 10^{11}}{10^{11}} \approx 3\text{s}$$

Which is starkly different!

4 Lorentz Transformations

In general, transforming between coordinates in two different frames can get extremely messy, more than can easily be handed in simple applications of length contraction or time dilation.

The Lorentz Transformations provide a general set of coordinate transformations between Σ and Σ', i.e. $(x, y, z, t) \to (x', y', z', t')$.

These transformations must be:

1. Symmetric about a change in sign of u, i.e. the transformation from $\Sigma \to \Sigma'$ with u and the transformation from $\Sigma' \to \Sigma$ with $-u$ must be the same.
2. They must be linear.

These transformations are:

We do not need to know the derivations, but do need to know the results.

[1] $\gamma = E/mc^2$ - will be expanded on in later lectures.

LC Electric Circuits

Fri 13 Feb 2026 11:00

Lecture 3

Fri 20 Feb 2026 11:00

Lecture 5

LC Electromagnetism I

Mon 19 Jan 2026 11:00

Lecture 1 - EM1 Intro and Electric Fields

1 Course Intro

Course Materials:

- Background material and derivations etc on PowerPoint.
- Worked examples etc are handwritten on visualiser, these are the bits we really need to know.

Why is EM important?

- Foundations of modern technology and the modern world.
- What gives elements their properties.
- Responsible for life itself.
- Everyday materials are held together by EM forces.
- Optics can only be understood through EM theory.

The course aim is to lay down the foundations, eventually leading us to Maxell's Laws.

1.1 Maxwell's Laws

Maxwell's four equations are:

$$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\epsilon_0}$$

$$\nabla \wedge \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}$$

$$\nabla \cdot \boldsymbol{B} = 0$$

$$\nabla \wedge \boldsymbol{B} = \mu_0 \left(\boldsymbol{J} + \epsilon_0 \frac{\partial \boldsymbol{E}}{\partial t} \right)$$

Where:

$$\nabla = \boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}$$

Together, these show that the electric and magnetic fields are related and are two aspects of a single force, the electromagnetic force. We don't have to properly understand them yet, but cannot learn them in EMII unless we fundamentally understand E and B fields from this module.

1.2 Course Structure

Part I: Electric Fields

- Charge and Coulomb's Law.
- The electric field.
- Gauss' Law.
- Capacitors.

Part II: Magnetic Fields

- Magnetic Fields

- Charged Particles in B-Fields
- Electromagnetic Induction.
- Magnetic Dipoles

In this lecture:

- Introduction to EM.
- Electric charge.
- Force between charges.
- The concept of the Electric Field (E-Field).

2 Electric Charge

First attributed to Thales circa. 624 - 546 BC. Experiments by Franklin and Coulomb expanded and showed that there was two types of charge, which they called positive and negative.

The "positive electricity" came from rubbing a glass rod with silk, and the negative from rubbing an ebonite (early plastic) rod with fur. They found that like charges repel and opposite charges attract.

We know that the elementary unit charge is the magnitude of charge of an electron/proton and everything else is a multiple of this [1]:

$$e = 1.6 \times 10^{-19} C$$

This has units of the Coulomb.

2.1 Charge Conservation

Electrons and protons are both stable (protons decay with a life greater than 10^{31} years). This means that the total charge of an isolated system is constant and can be conserved.

They have the same magnitude of charge, exactly:

$$|q_p| = |q_e| = e$$

2.2 Electrostatic Force

Like charges repeal and opposite charges attract, along the line of action given by a line drawn between the two charges. The force is proportional to the product of charges so:

$$F \propto q_1 q_2$$

Here, a negative force means attraction and a positive force means repulsion. Newton called this "force at a distance". Like gravity, two charges will exert a force on each other at a distance without any contact.

There must, therefore, be something between them that mediates this force. Later physics gives this as "virtual particles" which isn't a Y1 topic, so classically we say that this medium is the Electric Field.

2.3 Electric Field

A charge produces a field around it. Another charge also interacts with this field, and this interaction is what causes a force:

$$\underline{F} = \underline{E} q$$

Where F is the force exerted on a test charge of charge q by a charge Q producing a field E. The magnitude of the electric field has units NC^{-1} (force per units charge).

$$|\underline{E}| \propto Q \qquad |\underline{F}| \propto Qq$$

[1] While quarks have fractional charge, we don't get free quarks

Consider a point charge with a spherical electric field spreading out around it. As the distance from the charge increases, the surface area increases as $4\pi r^2$. Therefore the magnitude of the electric field must decrease with $4\pi r^2$

Therefore:

$$|\underline{E}| \propto \frac{Q}{4\pi r^2}$$

We need a (inverse) constant of proportionality. This depends on the medium, but for a vacuum we call it the permittivity of free space ϵ_0:

$$\epsilon_0 = 8.854 \times 10^{-12} C^2 m^{-2} N^{-1}$$

Hence:

$$\boxed{E = |\underline{E}| = \frac{Q}{4\pi\epsilon_0 r^2}}$$

2.4 Direction of the E-Field

Force is a vector, so the E-field must be too. Consider an E-field from a charge Q at distance r. We give E components E_r, E_θ, E_ϕ, where, since it's a sphere ϕ and θ represent the unit vectors in the two possible tangential directions.

If there was a component in E_θ this would be clockwise from one perspective, but anticlockwise from another (walking behind it). This is not possible, as the field must behave in the same manner from all viewpoints. Therefore:

$$E_\theta = E_\phi = 0$$

$$\underline{E} = \frac{Q}{4\pi\epsilon_0 r^2}\hat{t}$$

2.5 Force between two charges

Consider two charges q_1 and q_2. The force on q_2 due to the e-field from q_1 is:

$$\underline{F_1} = \underline{E_1} q_2$$

This is equal to the force on q_1 due to the e-field from q_2, given by:

$$\underline{F_2} = \underline{E_2} q_1$$

So the force between two charges is:

$$\underline{F} = \frac{q_1 q_2}{4\pi\epsilon_0 r^2}\hat{r}_{12}$$

2.6 Force between many charges

If we have more than two positive charges, we use the "principle of superposition". Effectively, you consider each pair of charges at the same time and vector sum of the forces together. I.e. if we have three points and we care about the net force on one, we take the vector sum of the two vectors from that point to the others.

In general:

$$\begin{aligned} \underline{F} &= \underline{F_1} + \underline{F_2} + \underline{F_3} + \cdots \\ &= q\sum_i \frac{q_i}{4\pi\epsilon_0 r^2}\hat{r_j} \end{aligned}$$

Where r_j is the distance between q_i and q, with unit vector $\hat{r_j}$ between them.

Since $\underline{F} = q\underline{E}$, the electric field at a test charge q must be:

$$\underline{E} = \sum_i \frac{q_i}{4\pi\epsilon_0 r_i^2}\hat{r_j}$$

2.7 Example

Say we have a square of side length a. Clockwise, these corners have charge Q, q, $-2Q$, $3Q$.

What is the net force exerted on the q charge?

LC Introduction to Particle Physics and Cosmology

Thu 22 Jan 2026 16:00

Lecture 1 - The Standard Model of Particle Physics

1 Course Introduction

Course Structure

- Particle Physics: 6 lectures in weeks 1 to 6.
  1. Introduction and the standard model.
  2. Experimental measurements.
  3. Interactions with matter
  4. Tracking detectors I
  5. Tracking detectors II
  6. Calorimeters and Particle Identification.
- Cosmology: 4 lectures in weeks 7 to 11.

Course Aims

- Overview of current methods in Particle Physics experiments.
- An emphasis placed on the questions and challenges.

For example, the LHC has already been programmed with experiments all the way up to 2041. Therefore, any detectors we design today only become relevant in over a decade, which makes good detector design decisions incredibly important. This course will equip us to understand what drives those design choices.

The course is assessed by a single one hour long exam, weighted half particle physics and half cosmology.

Recommended Texts

- Detectors for particle radiation (2nd edition), K. Kleinknecht (1998)
- Particle Physics, Martin and Shaw.
- High Energy Physics, D. H. Perkins (2nd through 4th editions)
- Feynman Lectures.
- Modern Particle Physics, M. Thomson.

2 Matter Particles

Fermions all have quantum spin 1/2. Spin is a purely inherent quantum property (like mass or charge) and has no classical representation, but is analogous to angular momentum. They are subject to Fermi-Dirac statistics, which means that no identical fermion in a system of fermions can have the same quantum number as any other. Fermions are divided into two types, quarks and leptons.

2.1 Quarks

There are three generations of quarks:

First Generation

- Up Quark (u), mass of ≈ 0.001GeV
- Down Quark (d), mass of ≈ 0.001GeV

Second Generation

- Charm Quark (c), mass of ≈ 1.3GeV
- Strange Quark (s) mass of ≈ 4.3GeV

Third Generation

- Top Quark (t), mass of ≈ 175GeV
- Bottom Quark (b), mass of ≈ 4.3GeV

"up-type quarks", u, c, t have electromagnetic charge $+2/3$ and "down-type quarks", d, s, b, have electromagnetic charge $-1/3$ (charges in units of e). Quarks do not ever exist alone in isolation.

2.2 Leptons

There are three generations of leptons, given by the electron e^-, muon, μ^- and tau τ^-. These all have charge of -1 (in units of e). These have masses (in MeV) of approximately $0.5, 105, 1800$.

These also have associated neutrinos, the electron neutrino, the mu neutrino and the tau neutrino, ν_e, ν_μ, ν_τ. These are not massless, but have a tiny mass many orders of magnitude smaller than their corresponding non-neutrino counterparts. These are neutrally charged.

Leptons are not subject to the strong interaction.

2.3 Hadrons

Quarks do not exist in isolation, but form bound states subject to the strong force called hadrons. There are two types of hadrons - baryons and mesons.

Baryons are formed from three quarks, which may be the same or different $q_1q_2q_2$.

Mesons are formed from a quark-antiquark pair, $q_1\bar{q}_2$. Examples of baryons include the proton and the neutrino, given by uud and udd.

3 Forces

As far as we know, there are four fundamental forces:

- Gravity
- Electromagnetism
- Strong
- Weak

For considering particle interactions, we disregard gravity as it becomes incredibly weak for small masses. Creating a complete theory that incorporates all four is an open question in physics. It's okay to neglect it, but it is unsatisfying.

We consider these forces as arising by the exchange (between two particles subject to a force between them) of particles called bosons. These have spin-1, so are called "gauge bosons". These are subject to Bose-Einstein statistics, which does not impose the same restriction as Fermi-Dirac for quantum numbers in a system.

For EM: The exchange particle is a photon, γ. This is represented on a Feynman diagram as a wiggly line. They are massless and couple to electric charge.

For Weak: The exchange particle is a $W^\pm$ or Z^0 boson. This is represented by a wiggly line or a dotted straight line. They are not massless, and have masses of approx. 80 and 90GeV respectively.

For Strong: The exchange particle is a gluon g. This is represented by a series of curls on a Feynman diagram. They are massless. They couple to "colour charge" which is just another quantum number analogous to electric charge. Just like electric charge has values $\pm$, colour charge has values we denote r, g, b

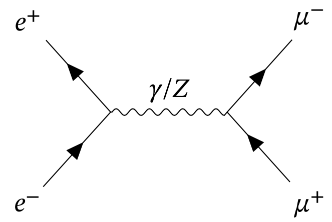
Quarks are subject to the strong, electromagnetic and weak interactions

Leptons are not subject to the strong interaction, but the e, μ, τ are subject to EM (ν is not as it is neutrally charged), and all are subject to the weak interaction. This makes neutrinos very difficult to detect as they are only affected by the weak interaction.

4 Feynman Diagrams

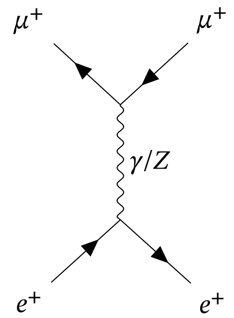
Feynman diagrams are space (y-axis), time (x-axis) diagrams to show allowed interactions between particles.

Consider a simple example of electron-positron annihilation. They travel towards each other, meeting and annihilating into either a photon or a Z boson. This is called a 'time-like exchange'. The boson then decays and we see pair production of two muons (one μ^- muon and one μ^+ antimuon).



Arrows in Feynman diagrams convey "fermion flow". This means that for a matter particle, the arrows aligns to the time axis. For antimatter particles, they antialign. Some conservation laws (i.e. charge) apply at the vertex level, while others only apply across whole processes.

We now consider a space-like exchange where the exchange is aligned with the vertical (space) axis. An antimuon scatters off a position like such:



5 Luminosity

We can determine the rate of interactions with the following:

$$W = \mathcal{L}\sigma$$

Where $W\,(\mathrm{s}^{-1})$ is interaction rate, $\mathcal{L}\,(\mathrm{cm}^{-2}\mathrm{s}^{-1})$ represents the luminosity (an attribute of the accelerator being used), and $\sigma\,(\mathrm{cm}^{-2})$ is the cross section, representing the underlying physics of the interaction.

These are investigated in greater detail in Lecture 03.

Mon 02 Feb 2026 12:00

Lecture 2 - Luminosity and Particle Signatures

1 Cross Sections

Consider a proton-proton interaction, producing some unknown particle X:

$$pp \to ppX$$

We have said that the rate of interaction is given by:

$$W = \mathcal{L}\sigma$$

Where $\mathcal{L}$ in $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is (coarsely) a parameter of the accelerator, describing its ability to produce collisions, and σ in cm^{-2} is a measure of interaction probability. Even though the particles are point-like, we treat them as having an effective area, and the magnitude of that area dictates how likely an interaction is to take place.

In this interaction, we have two protons (modelled as solid balls) passing immediately next to each other (one travelling clockwise and one counter-clockwise) around the accelerator. Assuming they pass immediately next to each other, and we model them as having radius $10^{-15}\mathrm{m}$, we have a a separation between the centres of each proton as $2 \times 10^{-15}\mathrm{m}$, therefore a cross section of:

$$\pi\left(2 \times 10^{-28}\mathrm{m}^2\right) \sim 0.12 \times 10^{-28}\mathrm{m}^2$$

To move this to a less annoying length scale, we define a new unit, the barn:

$$1\mathrm{barn} \equiv 10^{-28}\mathrm{m}^2 = 10^{-24}cm^2$$

In reality, this model may approximate a cross section, but it's not accurate. In reality, there's a much wider variation:

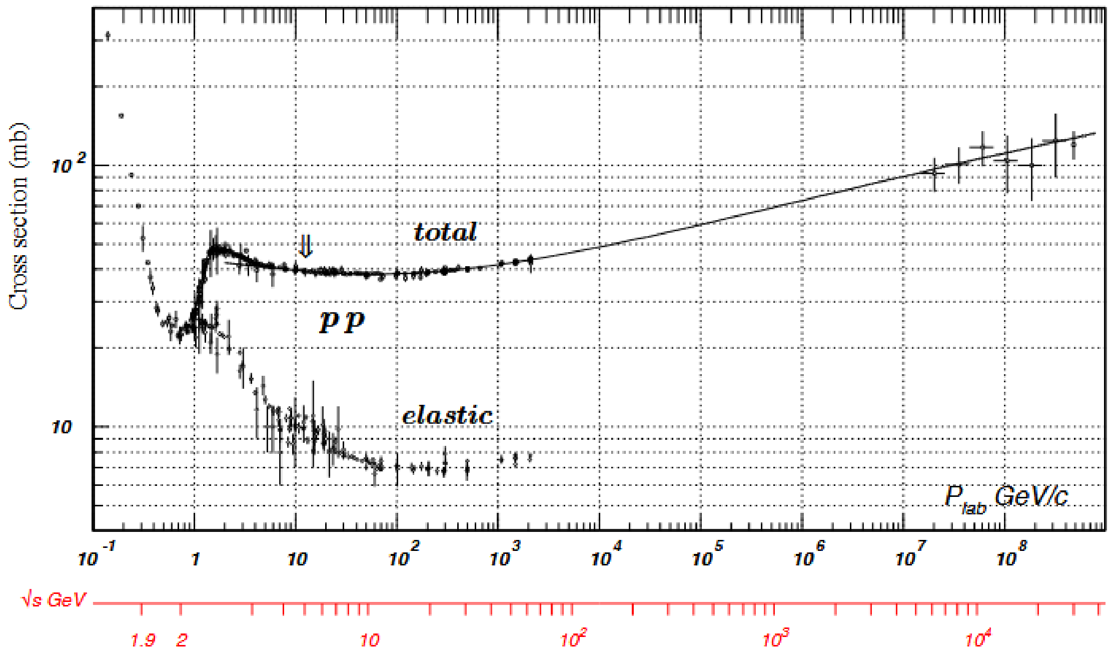


Figure 7.1

Here we have two x-axes running in parallel - the black axis is the momentum in a lab frame (as it hits some fixed target) while the red axis is the corresponding "centre of mass energy". How do we relate these two?

We want to know what the maximum mass of the particle we can generate is. In the lab frame, this requires us to take the momentum of the incoming and generated particle into account. We then have to take the final momentum of the system into account to conserve momentum, as the whole system must continue moving in the direction of motion for conservation.

Translating into the frame of reference given by the system centre of mass gives us a system where the two masses can be thought of as approaching each other with equal and opposite momenta. Since the total momentum is zero, the objects (incoming particle and the target) can theoretically hit each other and come to a complete stop. Because the system does not have to keep moving after the collision, all of the energy in this frame is free to be converted into the mass of a new particle.

While this may not be accurate in practice, it gives us a hard upper maximum for the possible energy available for production.

We can show relativistically that the energy in the centre of mass frame (labelled $\sqrt{s}$, the energy available for particle production) is:

$$E_{\text{com}} = \sqrt{s} \propto \sqrt{p_{\text{lab}}}$$

This can be seen in the final values of the x-axis, which (both starting approx. 1) are 10^8 and 10^4. This tells us that we reach diminishing returns with a fixed target collider - increasing energies by 8 orders of magnitude only increases the energy available for particle production by 4 orders. This is an inherent inefficiency of a fixed target collider.

This is much cheaper as its easier to align, we just fire a beam at a block of (for example) lead. It also makes it quite easy to change the target material. Changing materials in a colliding-beam collider (where two beams are fired in opposite directions, one clockwise and one anticlockwise, and collide with each other), i.e. to fire lead nuclei requires an extensive recalibration process.

In a dual-beam collider, only a very small proportion of the accelerated material from each beam actually interact with each other. In a fixed target collider, the target is much more dense, so we see a higher rate of interactions.

In summary, the advantages of a fixed target collider are:

1. Easier to collide.
2. Easier to change the target.

3. Very high density

2 Luminosity

2.1 Fixed Target Case

Consider a fixed target collider. We want to build an expression for the luminosity of this setup.

We have some incoming flux of particles (per second per unit area), J, incident on the block of material with density ρ, thickness t and mass of one nucleus m_A. The beam, modelled by a cylinder, illuminates some circular portion of the block, with area A.

Consider our interaction rate W. This is given by (where V is the volume of a cylinder from the illuminated beam, of thickness equal to the target):

$$W = \underbrace{JA}_{\substack{\text{incident particles}\\\text{per unit time}}} \times \underbrace{\frac{\rho}{m_A}V}_{\substack{\text{total number of}\\\text{target particles}}} \times \underbrace{\frac{\sigma}{A}}_{\substack{\text{probablity of}\\\text{interaction}}}$$

$$W = JA \times \frac{\rho}{m_A} At \times \frac{\sigma}{A}$$

$$W = J \times \frac{\rho}{m_A} At \times \sigma$$

$$W = \frac{J\rho At}{m_A}\sigma$$

Comparing to $W = \mathcal{L}\sigma$ gives the luminosity as:

$$\mathcal{L} = \frac{J\rho At}{m_A}$$

2.2 Colliding Beam Case

In a colliding beam case, the derivation is more (and too) complex. It is equal to:

$$\mathcal{L} = \frac{f_{\text{rep}} n_b N_1 N_2}{4\pi\sigma_x\sigma_y}$$

Where:

- f_{rep} is the repetition frequency, i.e the rate of the beam passing the collision point.
- n_b is the number of bunches in the beam.
  - A beam can be thought of as as string of pearls, rather than a single discrete constant beam - i.e. clusters of particles “bunches”, followed by empty space between them.
- N_1, N_2 are the number of particles per bunch for each beam
- σ_x, σ_y are the dimensions of the beam in the x- and y-direction, not a cross section as previously.

3 Examples of Detectors

We have some interaction point producing a spray of particles, and surround this with a series of different detector layers. Each produced particle will trigger a different subset of these layers, defining a unique signature we can use to identify produced particles.

Broadly, in some generic detector, we have:

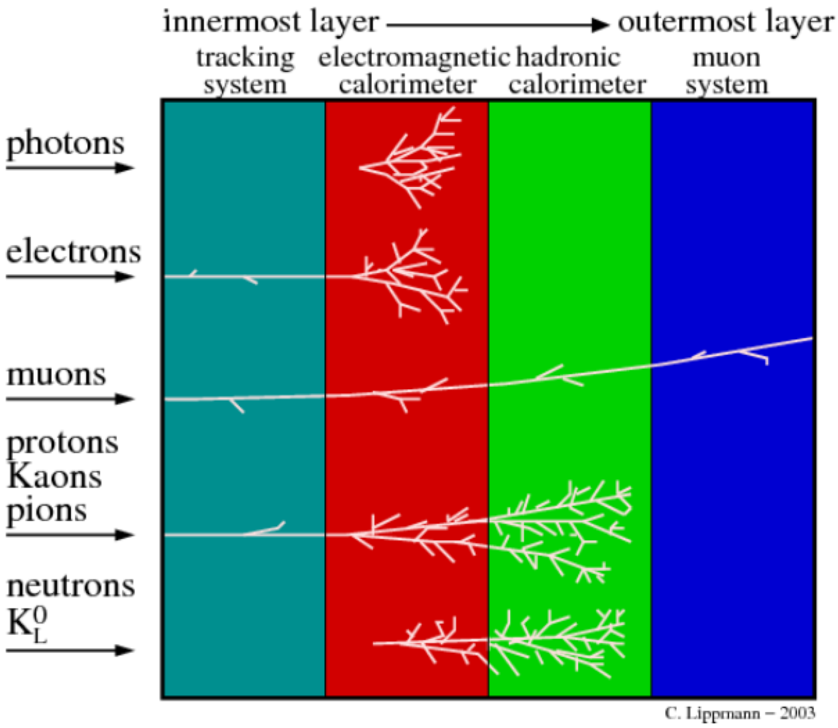


Figure 7.2

The tracking system is non-destructive. Formed of layers of silicon, a charged particle with ionise small portions of each layer. This can be turned into a signal. Neutral hadrons and photons will pass straight through, but charged particles will leave a deposit of charge and pass through unaffected.

We then move to destructive layers. Particles leave tree-like structures as they pass through these layers and create a shower of particles.

3.1 Decays

Most decays take place over a very low time scale, $< 10^{-8}$s and produce a final state made up of some subset of the following: $\gamma, \pi^+, \pi^-, \kappa^+, \kappa^-, p, n, \pi^0, e^+, e^-, \mu^+, \mu^-$.

So, in order to detect some exotic particle, we assume that it will either persist long enough to be detected itself, or decay into some subset of these known particles which will reach out detector.

Consider a parent particle A, for example B^0 ($\bar{b}d$) decaying into two child particles B, C, given by a "J Psi" ($c\bar{c}$), J/ψ and a "K Short", K_s^0 ($d\bar{s}$):

$$\begin{aligned} B &\to J/\psi \quad K_s^0 \\ \bar{b}d &\to c\bar{c} \quad d\bar{s} \end{aligned}$$

This decay has a lifetime of 10^{-12}s, via the weak interaction due to the change in quark flavour. The J Psi decays into e^+e^- or $\mu^+\mu^-$ via the EM interaction very rapidly in 10^{-21}s (its lifetime is governed by the strong interaction, which it may also use to decay via, even though we detect it via the EM decay path). The K Short decays into $\pi^+\pi^-$ or $\pi^0\pi^0$ again via the weak interaction with lifetime in 10^{-10}s.

The range of a particle is given by:

$$\text{Range} = \beta\gamma c\tau$$

Where τ is a time scale (lifetime), c is the speed of light, γ is the Lorentz Factor and β scales the range based on the speed actually being travelled. We also have:

$$E = \gamma m$$

$$p = \beta\gamma m$$

And familiarly:

$$E^2 - p^2 = m^2$$

If the B^0 has energy 20GeV and mass 5GeV, we have $\gamma = 4$ and this gives a range of $\approx$ 1mm. This is so small we will never observe it directly. The K_S^0 however has range $\approx$ 30cm, so is detectable.

Crucially:

- SI lifetimes: $\sim 10^{-21} - 10^{-24}$s
- EM lifetimes: $\sim 10^{-16} - 10^{-20}$s
- WI lifetimes: $\sim 10^{-12}$s

Thu 05 Feb 2026 16:00

Lecture 3 - Particle/Matter Interactions and Particle Signatures

In order to look at specific detectors and how they work, we need to consider how particles interact with matter. We'll consider categories of particles and their standard interactions, and use this to build a model for how we can build a detector.

We've looked at this previously:

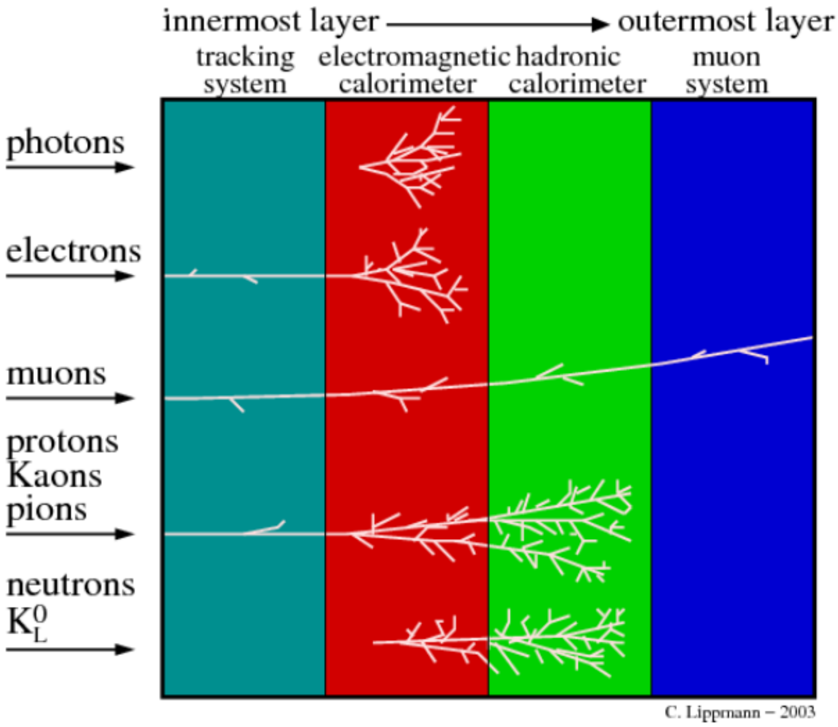


Figure 8.1

For example:

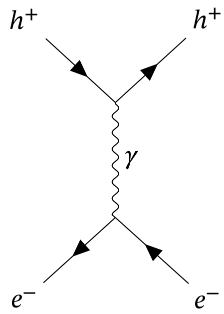
- Charged particles produce ionisation in the non-destructive tracking system, leaving deposits we can detect as signal.
- Electrons and photons leave distinctive signatures (shape of the shower of particles) in the electromagnetic calorimeters.
- Hadrons leave deposits in the electromagnetic calorimeter, but they dump all their energy (and are mostly identifiable by) the hadronic calorimeter.
- Muons make it through all of the previous layers, and are picked up at the very end by the outermost muon system.

1 Charged Particles

1.1 Ionisation

This is the process that powers the tracking system. A charged particle, h comes into the system (it could be a hadron or a lepton). It exchanges a photon with an electron in the system and excites the electron, ejecting it from the atom.

Strictly:



We can to try to characterise the rate of energy loss of the charged particle. The average rate of energy loss per unit distance is:

$$-\left\langle \frac{dE}{dl} \right\rangle \propto \ln E$$

Where E is energy, and the sign is negative as energy is being lost here.

For example, a real detector here might be two charged plates with a high voltage sandwiching a gas mixture. As the gas mixture is ionised, the ionised portion drifts towards one of the electrodes where it induces a detectable current.

1.2 Bremsstrahlung

Translates to "braking radiation". Consider a free electron radiating a photon. This is impossible for a free particle (as if we consider the electron's rest frame, emission would violate conservation of energy). We therefore need a source of external interference, in this case matter.

An electron is accelerated by nuclear change as it passes through material and is scattered. This scattering causes bremsstrahlung photon emission. The average rate of energy loss is given by:

$$-\left\langle \frac{dE}{dx} \right\rangle \propto \frac{E}{m^2} \propto \frac{E}{X_0}$$

An electron will then generate more bremsstrahlung than a muon, due to its much smaller mass. X_0 is called radiation length, and is covered in future lectures.

1.3 Cherenkov Radiation

Consider a charged particle moving through a (non-vacuum) material. It emits photons at some angle θ_c if the particle is moving faster than the speed of light in the medium (note this does not violate relativity, as the speed of light in a medium is less than the speed of light in a vacuum).

These emitted photons cause a coherent wavefront to form around the particle's trajectory, forming a cone around the direction of travel. The angle θ_c is given by:

Geometrically, after time t, the emitted photon has travelled ct/n and the particle vt, hence:

$$\cos\theta_c = \frac{ct/n}{vt} = \frac{c}{nv} = \frac{1}{n\beta}$$

Where n is the refractive index of the material. This is analogous to shock waves forming when an object goes faster than the speed of sound.

A planar detector will take a single cross-section through this cone, detecting rings around the point the charged particle passed through the material. By measuring this ring, we can work out the speed of the particle, and use this along with a measured momentum (in a tracking detector) to work out the mass of the particle.

Again:

$$-\left\langle \frac{dE}{dx} \right\rangle \propto z^2 \sin^2 \theta_c$$

Where z is the particle charge in units of $|e|$. It is important to note that this is a very small energy loss for the particle. It may emit $10^3\gamma$/cm, and only lose a few keV/cm

2 Photons

2.1 Photoelectric Effect

We have a photon strike a atom, transferring energy and forcing an electron to be ejected. The max kinetic energy of the electron is given by the photon energy minus some amount of work to eject it:

$$E_{kmax} = hf - \phi$$

Where ϕ, the work function, is the energy required to liberate one electron from the atom's surface.

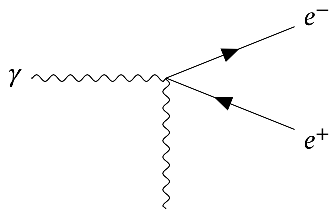
2.2 Compton Scattering

See QM1. A photon scatters off a quasi-free electron in an atom. The photon and electron are scattered with a change in energy:

$$\gamma + e^- \to \gamma' + e^{-\prime}$$

Thomson scattering is a low-energy form of scattering where the energies do not change, and Rayleigh scattering is a very low energy limit, where the interaction takes place between a photon and multiple atomic electrons.

2.3 Pair Production



A photon interacting with a nucleus can (leaving the nucleus unscathed) produce a particle and the corresponding antiparticle, typically a positron and an electron. This has a minimum photon energy of $E_\gamma > 2m_e c^2$[1] to ensure conservation of energy isn't violated and must occur in the presence of matter (for the same reason as Bremsstrahlung). The photon is not present in the final state.

However, say a photon has precisely the energy required to create a pair. This would (as it stands) create a pair of electrons with zero momentum (or very small momentum). However, a photon always has a momentum given by its energy, so momentum is not conserved. If the pair travel to attempt to resolve this, they now have some kinetic energy too, which means the photon must have a higher energy, and hence a higher momentum. This creates mismatch, we cannot conserve both energy and momentum in this situation.

By this interaction taking place in the presence of a nucleus, the nucleus can absorb some recoil (via exchange of a virtual photon) to ensure conservation of energy and momentum are satisfied.

This must take place in a Coulomb field to contribute this photon. The present nucleus has charge $z|e|$.

Lets consider the cross section of this interaction:

$$\sigma_{\text{pprod}} \propto \frac{7}{9} \frac{A}{N_A} \frac{1}{X_0}$$

[1] $m_e c^2 = 0.511$MeV

X_0 is the radiation length. It is a complex property of the material but is proportional to $1/z^2$.

3 Neutrinos

Neutrinos are really difficult to detect, as they are only effected by the weak interaction and have no ionisation/pair production etc - they are effectively non-interacting. We detect them by detecting the products of a weak force interaction.

For example:

- A muon neutrino can exchange a W boson with a proton, which becomes a neutron. The muon neutrino becomes a muon and the up quark in the proton becomes a down quark in the neutron.
  - We can detect the muon that's produced as the neutrino passes through the detector and interacts.
  - We build a massive multi-tonne detector that puts a large amount of mass in the way of the neutrino, often water.
  - We do this in the hope that it will interact with some of the matter and produce the more-detectable muon.
- The neutrino can scatter off a nucleus.
  - A mall amount of energy is exchanged with the nucleus, depositing a small amount of heat energy in the detector's matter.
  - Nothing changes/decays/is produced etc other than a small amount of heat.
  - We build a very sensitive detector capable of detecting this very small amount of heat.

4 Hadrons

Hadrons are really complex in their interactions. They can have inelastic nuclear interactions, with large energy deposits at a small number of sites (compared to an ECal shower where we'd expect to see many smaller deposits). This arises as a result of the nucleus becoming excited or breaking apart entirely and producing a chain of secondary hadrons/particles and a change in the nucleus.

These secondary particles can go on to interact again. This are complex and messy objects which can fragment off to cause many secondary impacts.

5 Conclusion

In summary:

- The key interactions we care about are ionisation, Bremsstrahlung and Cherenkov radiation for charged particles.
- For photons, we care about pair production.
- For hadrons, hadron showers are messy and complex. We add inelastic nuclear interactions on top of an electromagnetic component from ionisation and everything becomes rather tricky.

Mon 09 Feb 2026 12:00

Lecture 4 - Vertexing and Tracking Systems

In our familiar layered model:

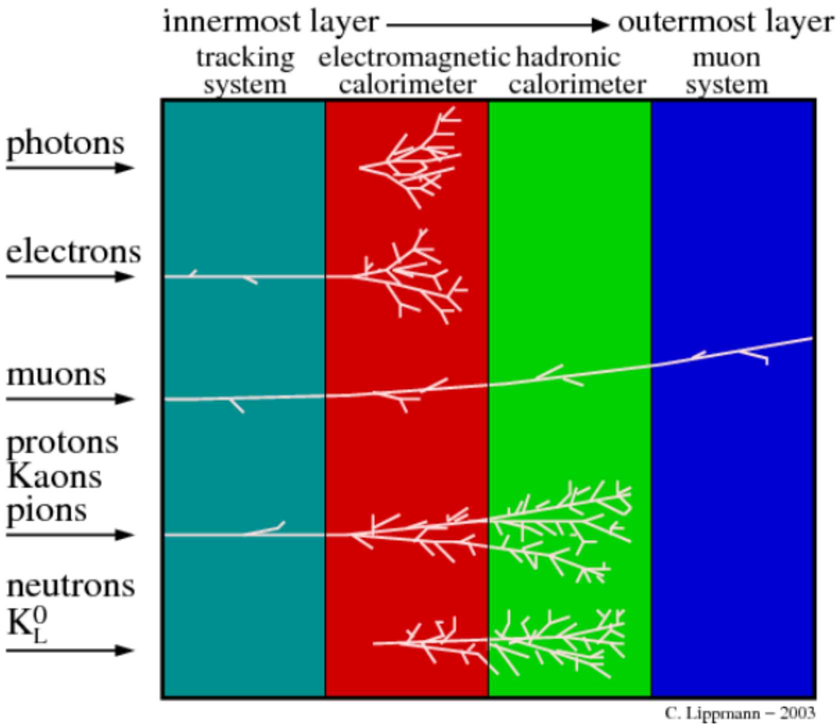


Figure 9.1

The first subdetector type is the initial tracking system which provides a non-destructive estimate of a particle's trajectory.

1 Tracking Systems

Purpose: To determine the trajectory of charged particles (usually in the presence of a magnetic field), in order to infer the momentum of the particle. This happens by ionisation, whereby the particle leaves small deposits of charge in either a gas or layers of a semiconductor sensor.

Position: As the following layers are destructive (i.e. calorimeters will absorb the particle entirely in order to make an energy measurement, while a muon system will block any other particle), we need to place the tracking system first for it to be effective.

We also want it placed near the primary interaction point. Since we want to use position measurements in the tracking system to extrapolate backwards and determine the origin point (where the collision took place), keeping them as close as possible to this origin point reduces the overall uncertainty of the track.

For a long path, a small uncertainty in position points propagates to a much larger uncertainty in track the further away we are from the initial collision point.

1.1 Material Budget

- In an ideal world, the tracking detector would have a perfectly massless and lightweight material, in order to reduce the risk of scattering. However, we need some mass in order to measure ionisation, so it's a constant trade off between the two.
- We want as small a number of radiation lengths X_0s within / upstream of the tracking system as possible.
- Some material is unavoidable, for example in the LHC there needs to be a conductive shield around the beam to prevent the large magnetic fields from inducing currents that would interfere with the sensitive measurement equipment. It also separates the highly pure vacuum of the beam pipe from the slightly less pure vacuum of the outer portion containing the LHCs electronics.

Radiation Length This is an inherent property of each material and is a measure of energy loss. A particle of energy E_0 passes through a distance of one radiation length and loses a factor of energy $1/e$.

$$E(x) = E_o \exp\left(\frac{-x}{X_0}\right)$$

For example:

- For Cu: $X_0 = 15\text{mm}$.
- For Be: $X_0 = 35.2\text{cm}$.

The processes by which energy is lost depends on the particle species in question. For example, an electron loses energy by Bremsstrahlung much more rapidly compared to a muon, due to the differences in mass (average energy loss $\propto 1/m^2$). To define radiation length, we use an electron as a scale.

This lets us talk about "material budget", as adding more material causes a higher uncertainty. For each particle path we care about, we want the fewest radiation lengths possible.

2 Measuring Trajectories

For a particle passing through a tracking system, we reconstruct its trajectory by measuring individual energy deposits (called "hits") caused by ionisation in a large volume (typically $1\text{m} \times 1\text{m} \times 1\text{m}$)

This volume is made up often of many layers, so we can gain an idea of the particles position as it passes through each layer and use this to extrapolate. A tracking system has a resolution of a few $100\mu\text{m}$ and may leave $10 - 100$ hits.

In a "vertex detector", we use a smaller system to reconstruct tracks at/around the interaction point with a higher precision - typically aiming for $\sim 10\mu\text{m}$. It is typically a silicon detector (layers of silicon detector sheets) and generally records fewer hits (< 10).

In a "tracking detector" we use a much larger system commonly further away from the primary interaction point. In a vertex system, there's no magnetic field so while we can use it to extrapolate back to find the PIP, we cannot use it to estimate momentum.

The LHC-b experiment for example has both a vertexer very close to the PIP (to determine the PIP location), and a much more substantial tracking system behind a magnet further away (to determine the deflected track in a magnetic field and hence the momentum.)

3 Measuring Momentum

The motion of a charged particle in a magnetic arises due to the Lorentz force and is proportional to charge:

$$\underline{F} = q(\underline{E} + \underline{v} \times \underline{B})$$

Where $\underline{F}$ is the force, $\underline{E}$ is the electric field, $\underline{v}$ is the particle velocity and $\underline{B}$ is the magnetic field. Assuming we are only dealing with a magnetic field, and taking magnitudes, we have:

$$\boxed{\underbrace{p}_{\text{GeV}} = 0.3 \times \underbrace{B}_{\text{T}} \times \underbrace{q}_{|e|} \times \underbrace{r}_{\text{m}}}$$

Where the magnetic field produces curvature of radius r in the plane perpendicular to the B-field. Motion parallel to $\underline{B}$ is unchanged.

In 3D, and assuming we are contained within the magnetic field, the particle will follow a helix of constant radius of curvature. This assumes that there are no non-conservative forces (i.e. scattering). In the ideal case, there is no work done, meaning in our force:

$$\underline{F} = q\underline{v} \times \underline{B}$$

The force and the velocity are perpendicular, so $\underline{F} \cdot \underline{v} = 0$.

3.1 Quantitatively

Consider a particle path with a constant radius of curvature R. Consider three hits, with vertical separation between the top and the bottom being $L/2$. The distance between the top/bottom hit and the origin is l, and the final distance between this straight line and the deflected path is the "sagitta" s.

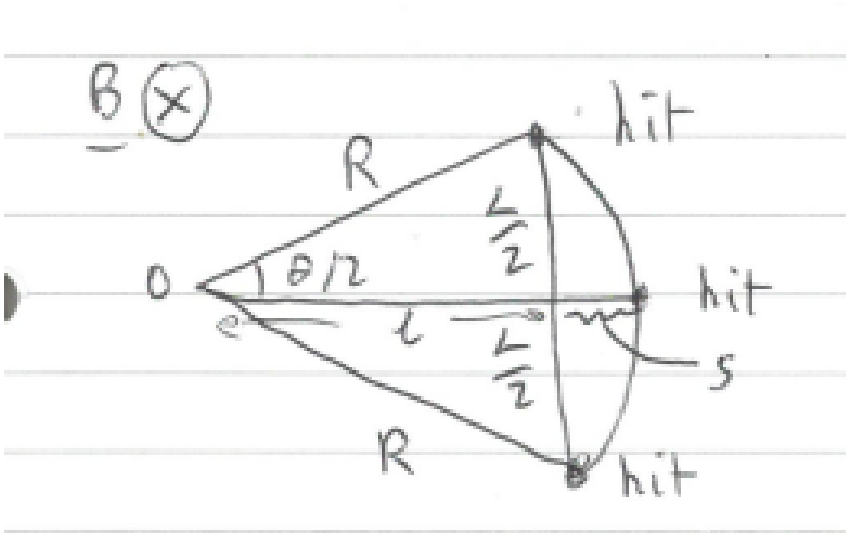


Figure 9.2

We know $L/2$ (the vertical separation between hits) as we've build the detector so know the resolution, and we want to measure the sagitta, the deviation from a straight line.

We know that:

$$p = 0.3BqR$$

And from the diagram:

$$R = l + s$$

$$l = R\cos\left(\frac{\theta}{2}\right)$$

Putting these together:

$$s = R\left(1 - \cos\frac{\theta}{2}\right)$$

Assuming the bending is gentle in a detector, so θ and s are small. Hence we apply a small angle approximation:

$$\cos\frac{\theta}{2} \approxeq 1 - \left(\frac{\theta}{2}\right)^2 \frac{1}{2!} + \left(\frac{\theta}{2}\right)^4 \frac{1}{4!} + \ldots$$

Taking the first order:

$$s \approxeq \left(\left(\frac{\theta}{2}\right)^2 \frac{1}{2!}\right) = \frac{R\theta^2}{8}$$

We can also use the small angle approximation for sine:

$$\sin\frac{\theta}{2} = \frac{L}{2R}$$

For small theta:

$$\sin\frac{\theta}{2} \approxeq \frac{\theta}{2}$$
$$\theta \approxeq \frac{L}{R}$$

So:

$$s = \frac{L^2}{8R}$$

Hence, finally, we have:

$$p = 0.3Bq\frac{L^2}{8s}$$

3.2 Uncertainties

We want to find the uncertainty on momentum, σ_p/p:

$$\frac{\sigma_p}{p} = \frac{\sigma_s}{s} = \frac{8p}{0.3BqL^2}\sigma_s$$

However we want to use our uncertainties on individual hits, x_1, x_2, x_3 etc. We have (and will not derive):

$$\frac{\sigma_p}{p} = \frac{\sigma_{xy}\,P}{0.3BL^2}\sqrt{\frac{720}{N+4}}$$

Mon 09 Feb 2026 12:00

Lecture 5 - Calorimetry

Now that we've finished talking about determining the principal interaction point and the momentum of produced particles in a tracking detector, we move onto calorimeters:

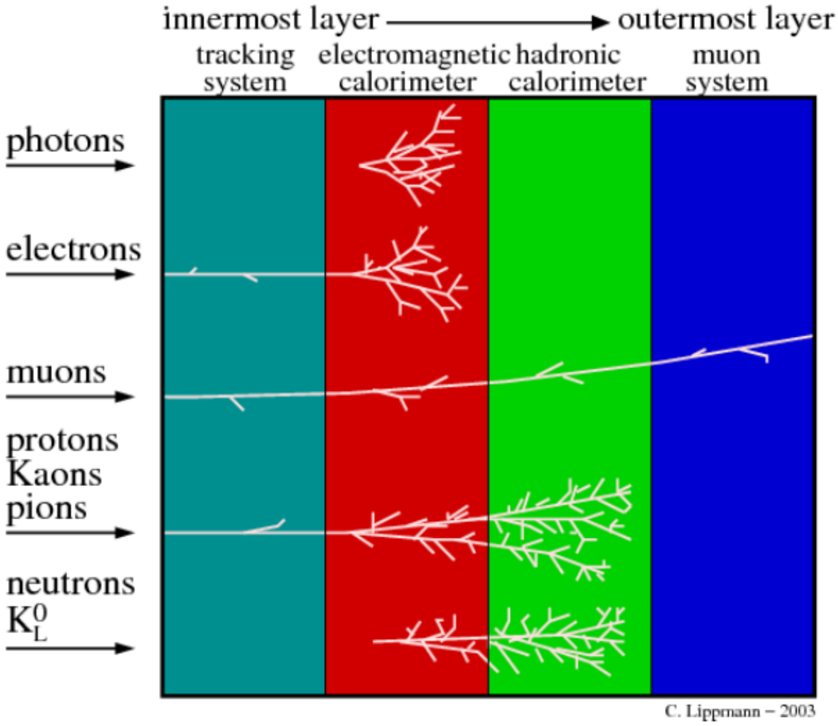


Figure 10.1

1 Materials and Energy Loss

Energy loss in a material is described by the Bethe formula, which gives mean energy loss by ionisation for unit path length:

$$\left\langle -\frac{dE}{dx} \right\rangle = Kz^2 \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 W_{\text{max}}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right]$$

This is:

- Valid for $\beta\gamma < 1000$ within a few percent precision.
- Includes dependence on the medium, I, Z, A etc.

The good news is that we do not need to know this formula. We do however need to know some key features:

- $\frac{dE}{dx} \propto \frac{1}{\beta^2}$ below some minimal value of $\frac{dE}{dx}$
- $\frac{dE}{dx} \propto \ln(\beta^2\gamma^2)$ above this minimal value of $\frac{dE}{dx}$. This is called "relativistic rise".
- This minima happens at approximately $\beta\gamma \sim 3 - 4$.

- At large $\beta\gamma$, polarisation of the medium causes saturation (effectively a plateau).

Broadly, we expect to see a fall up until some minimal value, then a relativistic rise, and then a plateau:

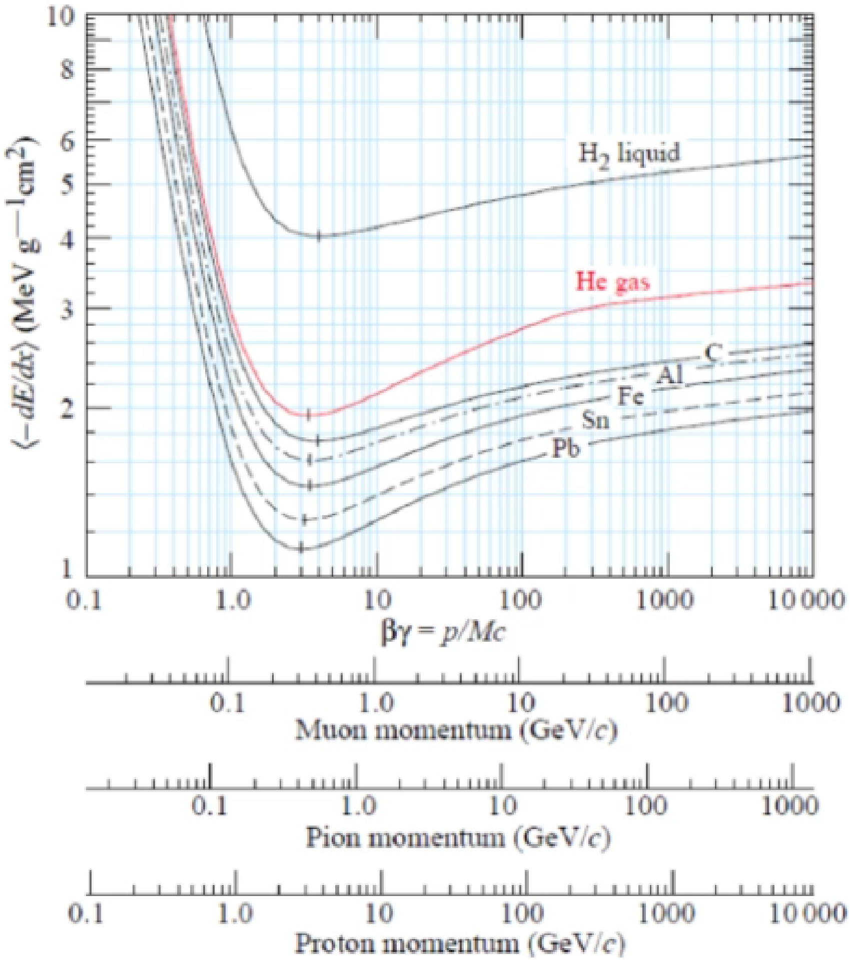


Figure 10.2

2 Calorimeters

The purpose of a calorimeter is to determine the total energy of an incident particle. It does so by absorbing the particle entirely, creating a measurable signal which is proportional to the energy of the incident particle.

LC Mathematics for Physicists 1B

Wed 21 Jan 2026 11:00

Lecture 1 - Course Welcome and Introduction to Partial Differentiation

1 Course Welcome

1.1 Recommended books:

- Mathematical Techniques 4e, Jordan & Smith
- Engineering Mathematics 8e, Stroud
- Calculus (Schaum), 6e, Ayres & Mendelson
- Advanced Calculus (Schaum), 6e, Ayres & Mendelson

1.2 Assessment details:

- Maths 1A/1B form a single 20 credit module.
- 80% assessed by a 3 hour exam - Section 1 is 36% with 6 short questions and Section 2 is 64% with 4 long questions.
- 20% assessed by problem sheets.

1.3 Course structure:

1. Partial Differentiation
   - Definition, total differential, chain rule, gradient.
   - Taylor series, stationary points, Lagrange multipliers.
2. Differential Equations
   - Definition, 1st order separable, exact and homogenous.
   - Linear equations: general solution, 1st order and constant coefficients.
3. Integration
   - Definition as area under the curve, fundamental theorem of calculus.
   - Integration by: substitution, parts, partial fractions and tricks.
4. Multiple Integrals
   - Multiple and repeated integrals. Change of order of integration.
   - Change of variables and the Jacobian. Arc length. Solids of revolution.

2 Multivariate Functions

Lots of physics involves functions of more than one variable. A physical quantity defined at every point in space is called a field. We can have both scalar fields and vector fields.

For example, some scalar fields are:

- $V(x, y, z)$: Electrostatic potential. This is often easier to work with compared to the full electric (vector) field.
- $T(x, y, z)$: Temperature.
- $p(x, y, z)$: Pressure.

While some vector fields are:

- $\underline{E}(x, y, z)$: Electric Field.
- $\underline{B}(x, y, z)$: Magnetic Field.
- $\underline{v}(x, y, z)$: Velocity Field (i.e in fluid mechanics).

2.1 Partial Derivatives

Consider a function of two variables. The partial derivative of a function with respect to one variable is the rate of change of a function wrt that variable, while keeping other variables constant. Effectively, we carry out a derivative while treating the other variables as if they were constants.

Suppose we have a function $f(x, y)$. The definition of a partial derivative is:

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \frac{f((x_0 + h), y_0) - f(x_0, y_0)}{h}$$
$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{k \to 0} \frac{f(x_0, (y_0 + k)) - f(x_0, y_0)}{k}$$

Just like we denote $\frac{df}{dx}$ as the derivative of a function of a single variable, we denote $\frac{\partial f}{\partial x}$ as the partial derivative of a function of several variables.

Note that this is not delta f and delta y, i.e. not $\frac{\delta f}{\delta x}$

In theory, we'd explicitly notate:

$$\left(\frac{\partial f}{\partial x}\right)_y$$

With the subscript y explicitly stating that y is being kept constant. This is rarely, but sometimes, needed.

Consider $f(x, y, z) = x^2 \sin yz$. We have:

$$\frac{\partial f}{\partial x} = 2x \sin yz$$
$$\frac{\partial f}{\partial y} = x^2 z \cos yz$$
$$\frac{\partial f}{\partial z} = x^2 y \cos yz$$

2.2 Higher Orders

Higher derivatives are defined as they were previously, but they can now be mixed. For example, with $f(x, y) = x^2 \sin y$, we can write:

$$\frac{\partial f}{\partial x} = 2x \sin y \qquad \frac{\partial f}{\partial y} = x^2 \cos y$$
$$\frac{\partial^2 f}{\partial x^2} = 2 \sin y$$

We can also have:

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = 2x \cos y$$

Shorthand notation exists, i.e. $f_{xx} = \frac{\partial^2 f}{\partial x^2}$ or $f_{yx} = \frac{\partial^2 f}{\partial y \partial x}$

For most cases, but not all, mixed derivatives are often independent of the order of partial derivatives, so: $f_{xy} = f_{yx}$

Thu 22 Jan 2026 09:00

Lecture 2 - Partial Differentiation II

1 The Total Differential

In order to generalise the chain rule, we need to define the total differential. Consider the change in a function of two variables, $f(x, y)$ when we move from some point (x, y) to some point $(x + dx, y + dy)$.

The partial derivative only tells us what happens when we change one variable, but we're changing two here. The total differential sums these two in order to get the full total change.

$$df = f(x + dx, y + dy) - f(x, y)$$

We have to look at the change in a single variable at a time, so we split it into two pieces where only x changes in the first, and only y changes in the second.

$$df = \underbrace{[f(x + dx, y + dy) - f(x, y + dy)]}_{\text{isolates change in } x} + \underbrace{[f(x, y + dy) - f(x, y)]}_{\text{isolates change in } y}$$

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

More generally for a function of $f(x_1, x_1, \dots, x_n)$:

$$df = \frac{\partial f}{x_1}dx_1 + \frac{\partial f}{x_2}dx_2 + \cdots + \frac{\partial f}{x_n}dx_n = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}dx_i$$

The total change in the function $f(x_1, x_1, \dots, x_n)$ is the sum of partial changes due to changing a single variable.

2 The Chain Rule

Recall that if $y = y(x)$, $x = x(t)$, then:

$$dy = \frac{dy}{dx}dx = \frac{dy}{dx}\frac{dx}{dt}dt \implies \frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$$

Now, if $f = f(x, y)$ and $x = x(t)$, $y = y(t)$, we can adjust the chain rule to say:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

$$= \frac{\partial f}{\partial x}\frac{dx}{dt}dt + \frac{\partial f}{\partial y}\frac{dy}{dt}dt$$

$$\implies \frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

Example

$$f(x, y) = x^2 + y^2, \qquad x(t) = t^2, \qquad y(t) = t^3$$

Hence:

$$f(t) = t^4 + t^5 \implies \frac{df}{dt} = 4t^3 + 6t^5$$

By rewriting in terms of one variable:

$$\frac{\partial f}{\partial x} = 2x = 2t^2 \qquad \frac{\partial f}{\partial y} = 2y = 2t^3$$

$$\frac{dx}{dt} = 2t \qquad \frac{dy}{dt} = 3t^2$$

And instead using the new chain rule:

$$\begin{aligned} &\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= (2t^2 + 2t) + (2t^3)(3t^2) \\ &= 4t^3 + 6t^5 \end{aligned}$$

Hence the new chain rule works!

2.1 Polar Coordinates

Suppose our x and y are now functions of two different variables themselves, so:

$$f = f(x, y) \qquad x = x(r, \theta) \qquad y = y(r, \theta)$$

From r, θ we want to calculate x, y and then from x, y we want to calculate f.

$$\begin{aligned} df &= \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \\ dx &= \frac{\partial x}{\partial r}dr + \frac{\partial x}{\partial \theta}d\theta \\ dy &= \frac{\partial y}{\partial r}dr + \frac{\partial y}{\partial \theta}d\theta \end{aligned}$$

Hence:

$$df = \frac{\partial f}{\partial x}\left(\frac{\partial x}{\partial r}\,dr + \frac{\partial x}{\partial \theta}\,d\theta\right) + \frac{\partial f}{\partial y}\left(\frac{\partial y}{\partial r}\,dr + \frac{\partial y}{\partial \theta}\,d\theta\right)$$

$$df = \left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r}\right)dr + \left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta}\right)d\theta$$

We also know (if we substitute x, y into the original function to get a function of r, θ):

$$df = \frac{\partial f}{\partial r}dr + \frac{\partial f}{\partial \theta}d\theta$$

We can read off the final partial derivatives:

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial x}{\partial r}, \qquad \frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta}$$

As expected!

2.2 Generalising

Suppose we have two functions which map $\mathbb{R}^m \to \mathbb{R}^p$, and $\mathbb{R}^p \to \mathbb{R}^n$, respectively:

$$(x_1, x_2, \dots, x_m) \to (y_1, y_2, \dots, y_p) \to (z_1, z_2, \dots, z_n)$$

Then we have:

$$\begin{aligned} dz_i &= \sum_{k=1}^{p} \frac{\partial z_i}{\partial y_k}dy_k \\ dy_k &= \sum_{l=1}^{m} \frac{\partial y_k}{\partial x_l}dx_l \end{aligned}$$

And substituting:

$$dz_i = \sum_{k=1}^{p} \sum_{l=1}^{m} \frac{\partial z_i}{\partial y_k} \frac{\partial y_k}{\partial x_l} dx_l$$

And (as the two sums are independent), we can pull out the inner sum:

$$= \sum_{l=1}^{m} \left[\sum_{k=1}^{p} \frac{\partial z_i}{\partial y_k} \frac{\partial y_k}{\partial x_l} \right] dx_l$$

$$= \sum_{l=1}^{m} \frac{\partial z_i}{\partial x_l} dx_l$$

The partial derivatives $\partial z_i / \partial x_j$ are therefore given by:

$$\frac{\partial z_i}{\partial x_j} = \sum_{k=1}^{p} \frac{\partial z_i}{\partial y_k} \frac{\partial y_k}{\partial x_j}$$

Fri 23 Jan 2026 12:00

Lecture 3 - Partial Differentiation III

Fri 13 Feb 2026 12:00

Lecture 12 - End of Partial Differentiation & Start of ODEs

Recap of lecture 11:

- The tangent plane to a surface $f(x, y, z) = 0$ at (x_0, y_0, z_0) is given by:

$$\left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0) + \left(\frac{\partial f}{\partial z}\right)_0 (z - z_0) = 0$$

  Such that $\underline{\nabla} f(x_0, y_0, z_0)$ is the normal vector to the plane
- The parametric representation of a curve $\underline{r}(t)$ has:
  - Unit tangent: $\underline{\hat{T}} = \frac{dr}{dt} / \left|\frac{dr}{dt}\right|$
  - Arc length $s(t)$: $\frac{ds}{dt} = \left|\frac{dr}{dt}\right| \to \underline{\hat{T}} = \frac{dr}{ds}$.
  - Unit normal and curvature:

1 Orthonormal Triads

We can create an *orthonormal triad* by introducing a new normal vector called the unit binormal, $\underline{\hat{B}} = \underline{\hat{T}} \times \underline{\hat{N}}$. Since $\underline{\hat{N}} \times \underline{\hat{N}}$, differentiating wrt s gives:

$$\underline{\hat{N}} \cdot \frac{d\underline{\hat{N}}}{ds} = 0$$

TODO

We have:

$$\frac{d\underline{\hat{T}}}{ds} = \kappa \underline{\hat{N}}$$
$$\frac{d\underline{\hat{N}}}{ds} = -\kappa \underline{\hat{T}} + \tau \underline{\hat{B}}$$

Hence:

$$\begin{aligned}\frac{d\underline{\hat{B}}}{ds} &= \frac{d}{ds}\left(\underline{T} \hat{\times} \underline{\hat{N}}\right) \\ &= \frac{d\underline{\hat{T}}}{ds} \times \underline{\hat{N}} + \underline{\hat{T}} \times \frac{d\underline{\hat{N}}}{ds} \\ &= \kappa \underline{\hat{N}} \times \underline{\hat{N}} + \underline{\hat{T}} \times \left(-\kappa \underline{\hat{T}} + \tau \underline{\hat{B}}\right) \\ &= \tau \underline{\hat{T}} \times \underline{\hat{B}} = \tau \underline{\hat{T}} \times \left(\underline{\hat{T}} \times \underline{\hat{N}}\right) \\ &= \tau \left[\left(\underline{\hat{T}} \cdot \underline{\hat{N}}\right) \underline{\hat{T}} - \left(\underline{\hat{T}} \cdot \underline{\hat{T}}\right) \underline{\hat{N}}\right] \\ \\ &= \tau \underline{\hat{N}}\end{aligned}$$

This gives the Frenet-Serret Formulae:

This concludes partial differentiation! :D

2 Ordinary Differential Equations

A differential equation is any equation that involves derivatives. We care, because most laws of physics manifest themselves in the form of differential equations. For example:

$$\text{Newton's Second Law:} \qquad \underline{F} = m\frac{d^2 r}{dt^2}$$

$$\text{3D Time-Independent Schrödinger Eqn:} \qquad -\frac{\hbar}{2m}\left(\frac{\partial^2\psi}{dx^2} + \frac{\partial^2\psi}{dy^2} + \frac{\partial^2\psi}{dz^2}\right) + V(x,y,z)\psi = E\psi$$

$$\text{3D Wave Eqn:} \qquad \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}$$

$$= \nabla^2\phi = \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}$$

$$\text{Gauss' Law:} \qquad \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = \frac{\rho}{\epsilon_0}$$

$$\text{Navier-Stokes Eqn:} \qquad \rho\left(\frac{\partial \underline{v}}{\partial t} + (\underline{v}\cdot\underline{\nabla})\,\underline{v}\right) = -\underline{\nabla}p + \rho\underline{g} + \mu\nabla^2\underline{v}$$

In this course, we will only solve DEs of a single variable, i.e. Ordinary Differential Equations (ODEs). We don't look at Partial DEs of multiple variables yet.

In order to think about solving these, we need to classify them. Most DEs aren't soluble in closed form with elementary functions and need to be solved numerically. Here, we only consider nice soluble functions, but this is a vast minority in reality. We want to identify classes of DEs we can reasonable solve with a method for each.

We can generally solve linear equations by breaking them into small chunks and solving them individually, for example.

3 Types of DEs

3.1 Partial vs. Ordinary

In the examples above, only the first was an ODE, and the rest PDEs. Ordinary Differential Equations (ODEs) involve only a single variable.

Consider a vector $\underline{r}(t) = (x(t), y(t), z(t))$. t is called the independent variable, with x, y, z being dependant variables. While we have 3 dependant variables, we only have one independent variable (so only one thing to differentiate wrt), so this would end up being ordinary.

PDEs involve equations of two or more variables and hence involve partial derivatives.

3.2 Order

The order of a DE is given by the order of the highest derivative involved, so Newton's 2nd Law is a second order DE, as the highest order derivative is a second derivative.

3.3 Degree

The degree of a DE is a less important measure than the others. It is given by the highest power of the highest order derivative. For example, Newton's 2nd is a first degree, while an equation containing a^3 would be third degree (and second order, as a is a second derivative).

Ideally, we want this to be 1 for ease of solving, and higher degrees are rare but they do exist. For example, from Lagrangian Mechanics we have:

$$\frac{1}{2m}\left[\left(\frac{\partial s}{\partial x}\right)^2 + \left(\frac{\partial s}{\partial y}\right)^2 + \left(\frac{\partial s}{\partial z}\right)^2\right] + V(x,y,z) = \frac{ds}{dt}$$

3.4 Homogenous and Inhomogeneous

A homogenous DE is a DE that does not have any terms of only the independent variable(s), while an inhomogeneous DE does.

For example, Newton's 2nd is homogenous as there is no term that involves t alone. This would be inhomogeneous:

$$\frac{\partial^2 x}{\partial t^2} = t + x$$

While this would be homogenous:

$$\frac{\partial^2 x}{\partial t^2} = tx$$

As t is a coefficient and not a pure term in its own right.

3.5 Linear and Non-Linear

A DE is linear if the dependant variable(s) and all of its/their derivatives occur purely as linear functions. For example:

$$\left(1 - x^2\right)\frac{d^2 y}{dx^2} - 2x\frac{dy}{dx} + n(n+1)y = 0$$

This is linear, as the dependant variable y never has a power greater than 1.

$$\frac{dy}{dx} + xy = 0$$

Is also linear, while this is not:

$$\frac{dy}{dx} + xy^2 = 0$$

This is also non-linear (as shown by the Taylor Expansion of sine):

$$\frac{d^2\theta}{dt^2} = -\frac{g}{l}\sin\theta$$

3.6 Examples

$$(1) \quad \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 = u^2$$

Homogenous first-order second-degree non-linear PDE.

$$(2) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = x^2 + y^2 + z^2$$

Inhomogeneous second-order first-degree linear PDE.

$$(3) \quad \frac{\partial y}{\partial x} + y^2 = x$$

Inhomogeneous first-order second-degree non-linear ODE.

Fri 20 Feb 2026 11:56

Lecture 15 - Differential Equations III

Summary of last lecture:

- "Homogeneous" Equations
$$\frac{dy}{dx} = f\left(\frac{x}{y}\right) \to x\frac{dv}{dx} = f(v) - v \text{ where } y(x) = xv(x)$$
- Linear Equations
$$\sum_{k=0}^{n} a_k(x)\frac{d^k y}{dx^k} = f(x) \to y(x) = \sum_{k=1}^{n} \alpha_k y_k(x) + y_{PI}(x)$$
The solution is a sum of the complementary function (the general solution of the homogenous equation) and the particular integral (one solution of the inhomogeneous equation).
- First Order Linear Equations
$$\frac{dy}{dx} = P(x)y = Q(x) \to \frac{d}{dx}\left[y(x)e^{\int P(x)dx}\right] = Q(x)e^{\int P(x)dx}$$

1 Examples

1.1 Example I

$$(1-x^2)\frac{dy}{dx} - xy = 1$$

Rewriting in the standard form:
$$\frac{dy}{dx} - \frac{x}{1-x^2}y = \frac{1}{1-x^2}$$

Hence:
$$P(x) = \frac{-x}{1-x^2}$$
$$I(x) = \exp\left(-\int \frac{xdx}{1-x^2}\right) = \exp\left(\frac{1}{2}\ln(1-x^2)\right) = \sqrt{1-x^2}$$

Multiplying through by the integrating factor:
$$\sqrt{1-x^2}\frac{dy}{dx} - \frac{x}{\sqrt{1-x^2}}y = \frac{1}{\sqrt{1-x^2}}$$

The L.H.S is now the derivative of a product:
$$\frac{d}{dx}\left(y\sqrt{1-x^2}\right) = \frac{1}{\sqrt{1-x^2}}$$

And integrating both sides:
$$y\sqrt{1-x^2} = \arcsin x + c$$
$$\implies \boxed{y = \frac{c}{\sqrt{1-x^2}} + \frac{\arcsin x}{\sqrt{1-x^2}}}$$

1.2 Example II

2 Linear ODEs with Constant Coefficients

In the most general form, we have:

$$a_n \frac{d^n y}{dx^n} + a_{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1 \frac{dy}{dx} + a_0 y = f(x)$$

Firstly, we solve the homogenous equation:

$$a_n \frac{d^n y}{dx^n} + a_{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \cdots a_1 \frac{dy}{dx} + a_0 y = 0$$

Let $y(x) = e^{\lambda x}$:

$$e^{\lambda x} \left(a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0 \right) = 0$$

This reduces to finding the zeroes of the corresponding nth order polynomial. If this polynomial has n distinct zeroes, then the complimentary function is given by:

$$y_{CF}(x) = \alpha_1 e^{\lambda_1 x} + \alpha_2 e^{\lambda_2 x} + \cdots + \alpha_n e^{\lambda_n x}$$

If these roots contain a repeated root, this will reduce the number of unique solutions by one. If we have any complex solutions, they wil come in complex conjugate pairs:

$$a^{(\pm ib)x} = e^{ax} \left(\cos bx \pm i \sin bx \right)$$

This therefore has independent real solutions:

$$e^{ax} \cos bx$$

$$e^{ax} \sin bx$$

3 Equidimensional Equations

These have coefficients which do depend on x, but where the coefficients are functions of x such that the nth derivative has a coefficient of $a_n x^n$

The general form of a homogeneous equidimensional equation is:

$$a_n x^n \frac{d^n y}{dx^n} + a_{n-1} x^{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1 x \frac{dy}{dx} + a_0 y = f(x)$$

This has a solution in the general form $y(x) = x^\lambda$. When we differentiate k times with respect to x we lose k powers of x, as each differentiation decreases the power. We need to restore this therefore with a x^k prefactor.

4 Mass on a Spring (SHM)

A mass, m on a spring is displaced from its equilibrium position by some distance x. There is a restoring force given by $F = -kx$.

$$F = ma \implies F = m \frac{d^2}{dt}$$

$$m \frac{d^2 x}{dt^2} + kx = 0$$

$$\frac{d^2 x}{dt^2} + \omega^2 X = 0 \qquad \text{where } \omega^2 = \frac{k}{m}$$

We have:

$$x(t) = e^{\lambda t}$$

$$(\lambda^2 + \omega^2)e^{\lambda t} = 0$$
$$\lambda^2 + \omega^2 = 0$$

Hence:

$$\lambda = e^{\pm i\omega t}$$

LC Temperature and Matter

Tue 20 Jan 2026 12:00

Lecture 1

Tue 20 Jan 2026 12:00

Lecture 1

Tue 20 Jan 2026 12:00

Lecture 1

Tue 20 Jan 2026 12:00

Lecture 1

Tue 20 Jan 2026 12:00

Lecture 1

Tue 20 Jan 2026 12:00

Lecture 1

Tue 20 Jan 2026 12:00

Lecture 1

Tue 20 Jan 2026 12:00

Lecture 1

Tue 17 Feb 2026 12:00

Lecture 9 - Laws of Thermodynamics

1 0th Law of Thermodynamics

Consider two blocks next to each other, one with T_h and one with T_c. The blocks with thermalise to their mean $(T_h + T_c)/2$ in the trivial case where the block are identical.

Net heat will flow between thw two systems until they have the same energy density and hence temperature.

If we have three systems: A, B, C. If we know that A and B are in thermal equilibrium, and A and C are in thermal equilibrium, then B and C are also in thermal equilibrium.

It's a fairly trivial axiom, so we denote it the zeroth law. Effectively, all objects in a system in equilibrium share the same temperature.

2 Ideal Gases

An ideal gas is defined as a collection of molecules (or atoms, if monatomic) that are non-interacting with each other (no interatomic forces) and collide elastically with each other.
The internal energy of the gas is dependant on the velocities of the molecules, and hence on the temperature, and not on pressure or volume.

Boyle's Law: "The absolute pressure exerted by a given mass of an ideal gas is inversely proportional to the volume is occupies, if the temperature and amount of gas remain unchanged within a closed system."

Effectively:

$$P \propto \frac{1}{V}$$

Charles' Law: "When the pressure of a sample of an ideal gas is held constant, the Kelvin temperature and volume will be in direct proportion."

Effectively:

$$T \propto V$$

Ideal Gas Law: Since $P \propto \frac{1}{V}$ and $T \propto V$, we have: $PV = kT$, where k varies with context, for example when considering moles:

$$\boxed{PV = nRT}$$

Where n is the number of moles and $R = 8.314$ is the gas constant.

Or for molecules:

$$\boxed{PV = Nk_bT}$$

Where $N = N_a n$ is the number of molecules, and $k_b = R/N_a$ is the Boltzmann Constant.

This is called an equation of state and allows us to describe the gas' state macroscopically. We generally express this on a P/V diagram, where each point on the plot represents a specific gas state. If we keep the amount of gas present constant, we can use any two of the variables to determine the third.

2.1 Joule's Second Law

Consider a system with a box divided in two. An ideal gas is contained in the leftmost half, and the rightmost half is a vacuum. We separate the two with a divider constituting an impermeable membrane.

We remove the membrane laterally, doing no work on the gas. Joule observed that the gas stayed at the same temperature as it diffused into the vacuum.

Consider the internal energy of the gas, denoted U, assuming that U is a function of two of the state variables. We know the temperature did not change, and the volume did, so $U(T, V)$.

We would expect:

$$dU = \frac{\partial U}{\partial T}dT + \frac{\partial U}{\partial V}dV$$

Temperature was observed experimentally to not change, hence $dT = 0$, so:

$$dU = 0 + \frac{\partial U}{\partial V}dV$$

Volume did change, so $dV \neq 0$. We did no work on the gas, as the divider was removed laterally, hence there was no change in internal energy, so:

$$dU = \frac{dU}{dV}dV = 0$$

And since $dU = 0$, we must have:

$$\frac{dU}{dV} \neq 0$$

This means that there is no dependence on volume for internal energy, hence U is dependent only on T, as found by Joule. This means that the gas Joule chose was well approximated by an ideal gas.

This is easier at higher temperatures, as a high temperature leads to high kinetic energies, so the interatomic forces becomes less significant and easier to disregard in reality.

We see here that it the ideal gas law becomes a better approximation has we increase temperature, and a worse and worse approximation as pressure increases.

From a LJP perspective, increasing the pressure decreases the average distance between gas molecules. This means that the potential between gas molecules is no longer negligible, and the assumption of zero potential no longer holds.

If we increase pressure even further, the force becomes repulsive and PV/nRT returns to being positive. At higher temperatures, this little dip isn't observed.

At low pressure or high temperature, the ideal gas law is a good assumption as:

- At low pressure, interatomic spacing is large enough to disregard interatomic forces.
- At high temperature, kinetic energy is large enough to comparatively disregard interatomic forces.

We deal solely with Maxwell-Boltzmann gases in this source that follow classical laws. We also have Fermi and Bose gases (made entirely of fermions and bosons respectively), but they're quantum mechanical, exotic and outside of this course.

2.2 Changing Energies

Say we want to go from T_1 to T_2. We need to change the internal energy of the gas somehow. We have two general ways to change this:

- Add heat to the system by transferring heat to the gas at a constant volume.
- Do some work on the gas, i.e. by crushing it and reducing volume.

Consider a piston of area A, applying constant force F with extension ΔL. The pressure on the gas is $P = F/A$. The work done on the gas is:

$$\Delta W = F\Delta L$$
$$\implies \Delta W = P(A\Delta L)$$
$$\implies \Delta W = P\Delta V$$

Moving from infinitesimal Delta values:

$$W_{\text{on gas}} = \int dW = \int P\,dV$$

In the opposite case considering work done by the gas, conservation of energy says it must be oppositely signed:

$$W_{\text{by gas}} = -\int P\,dV$$

2.3 Mass Drop Experiment

The falling mass spins the paddles in the water, the work as potential energy is converted to kinetic energy raising the temperature of the water.

3 1st Law of Thermodynamics

The change in internal energy of a system, ΔU is increased with increasing heat transfer into the system, Q_{in} and work done on the system, W_{on}:

$$\Delta U = Q_{\text{in}} + W_{\text{on}}$$

This is just conservation of energy.