Mechanics and relativity

Course supplement and recapitulation notes

Markuss G. Ķēniņš

University of Groningen

October 26, 2024

CONTENTS

I Introduction **3**

§1 Units, vectors, and the centre of mass 4

II Relativity **9**

§2 The principle of relativity 10
§3 Coordinate time 13
§4 The spacetime interval 16
§5 Proper time 19
§6 Dimensional analysis 20
§7 First exam preparation 22
§8 Coordinate transformations 24
§9 Speed of light and the Doppler effect 26
§10 Guestimation 28
§11 Four-momentum 29

CHAPTER I
INTRODUCTION

§1. Units, vectors, and the centre of mass

§ Approximations. Much of the theory of elementary classical physics is concerned with making suitable approximations and simplifications; this point shall be prevalent throughout this entire course and the reader is advised to adopt it promptly, as well as to achieve proficiency at it. A mathematician of pure degree may argue that any such solution utilising an approximate step is not fully correct; this is rebutted by the simple principle in physics of an ever-more precise theory – no matter what theory, it captures the quantities only up to some finite precision; beyond this precision the theory is meaningless.[1]

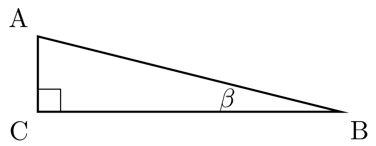
Several commonly used approximations are given below.

1. Any infinitely-differentiable function $f : \mathbb{R} \to \mathbb{R}$ is successively approximated around $x_0 \in \mathbb{R}$ (with notation $u = x - x_0$) by

   (i) a constant $\bar{f}(u) = x_0$,
   (ii) a line $\bar{f}(u) = x_0 + f'(x_0)u$,
   (iii) a parabola $\bar{f}(u) = x_0 + f'(x_0)u + \frac{1}{2}f''(x_0)u^2$

   and so on (the coefficient of u^k is $(1/k!)f^{(k)}(x_0)$ – this is the famous Taylor series). For example, $\sin\theta \approx \theta$ and $\cos\theta \approx 1 - \frac{1}{2}\theta^2$ for small $-\pi/2 < \theta < \pi/2$; $e^x \approx 1 + x$, $(1+x)^n = 1 + nx$ for small x.

2. For a right triangle



   with angle β small, the adjacent edges AB and CB are of approximately equal length. For a very small angle β these edges are approximately parallel.

The reader will encounter many more examples in their journey through physics.

§ Units. In physics the quantities encountered very often contain units; all units must be expressed in terms of some basis set of units: the internationally accepted is the *Système international d'unités*, or SI units. The quantities are meters, kilograms, seconds, and a few others, as the reader surely knows; conversion to SI units is crucial for simplification of units. Units are *not* italicised when typesetting.[2]

A quick way to simplify some complicated unit, which any sensible person does not take upon themselves to recall, is to use elementary formulas to reduce it, for example:

1. Newton N arises as the unit of force $F = ma$, where mass has unit kilogram kg and the unit of acceleration is meters per second squared $\mathrm{m/s^2}$, hence

$$[\mathrm{N}] = \left[\frac{\mathrm{kg}\cdot\mathrm{m}}{\mathrm{s}^2}\right],$$

   where [ ] denotes a unit within it.

[1] As an example consider the movement of a particle; if it describes the centre of a tennis ball, all is well and 0.1 mm precision is more than sufficient. If we consider an electron, however, the result of classical theory up to arbitrarily small precision becomes nonsense – it is not a classical system.

[2] In LaTeX the `siunitx` package, loaded as `\usepackage[separate-uncertainty=true]{siunitx}`, is invaluable for typesetting units; the usual command is `\SI{2.06(3)e5}{\kilo\gram\per\meter}`, which produces $(2.06 \pm 0.03) \times 10^5\,\mathrm{kg\,m^{-1}}$, substituting for applicable numbers and units in general, of course.

2. Similarly, for Joule: $W = Fx$, where $x = [\mathrm{m}]$ and $F = [\mathrm{N}]$, hence
$$W = [\mathrm{J}] = [\mathrm{N \cdot m}] = \left[\frac{\mathrm{kg \cdot m^2}}{\mathrm{s^2}}\right],$$

3. And power: $P = W/t$, giving
$$P = [\mathrm{W}] = \left[\frac{\mathrm{J}}{\mathrm{s}}\right] = \left[\frac{\mathrm{kg \cdot m^2}}{\mathrm{s^3}}\right],$$

There are two styles of performing unit conversion; the author dislikes the style given in the book – for it the reader is advised to consult the book; the preferred style shall be presented here. The rule is simple:

> 'First, rewrite composed units in terms of basis units. Then convert each basis unit to the corresponding SI unit(s). Only then do cancellation.'

The following example is presented. Consider a galaxy observable from Earth; the quantity that describes the brightness of the surface of this galaxy per unit area, Σ, is called *surface brightness*, which is given by $\mathrm{L_\odot/pc^2}$, where $\mathrm{L_\odot}$ is the solar luminosity (luminosity of our[3] Sun) and pc is a parsec – the distance at which one astronomical unit subtends an angle of one arcsecond ($1/3600°$). Suppose the surface brightness of Andromeda has been measured to be $1.92 \times 10^{-7}\ \mathrm{L_\odot/pc^2}$; now for the conversion to SI units.

The value of one solar luminosity is $1\ \mathrm{L_\odot/pc^2} = 3.828 \times 10^{26}\,\mathrm{W}$, whereas a parsec is approximately $3.26\,\mathrm{ly}$ (light-years). A year consists of 365 days, 24 hours per day, 60 minutes per hour, and 60 minutes per second, thus $1\,\mathrm{y} = 3.1 \times 10^7\,\mathrm{s}$; now, a light-year is the distance light (with speed $c = 3 \times 10^8\,\mathrm{m/s}$) travels during a year, hence $1\,\mathrm{ly} = 9.5 \times 10^{15}\,\mathrm{m}$. A parsec is thus
$$1\,\mathrm{pc} = 3.1 \times 10^{16}\,\mathrm{m} \quad \Longrightarrow \quad 1\,\mathrm{pc}^2 = 9.5 \times 10^{32}\,\mathrm{m}^2.$$

Finally, substituting and simplifying the units,
$$\left[\frac{\mathrm{L_\odot}}{\mathrm{pc^2}}\right] = \frac{3.828 \times 10^{26}}{9.5 \times 10^{32}}\left[\frac{\mathrm{W}}{\mathrm{m^2}}\right] = 4.03 \times 10^{-5}\left[\frac{\mathrm{kg \cdot m^2}}{\mathrm{s^3}}\frac{1}{\mathrm{m^2}}\right] = 4.03 \times 10^{-5}\left[\frac{\mathrm{kg}}{\mathrm{s^3}}\right],$$
thus
$$1.92 \times 10^{-7}\frac{\mathrm{L_\odot}}{\mathrm{pc^2}} = 1.92 \times 10^{-7} \cdot 4.03 \times 10^{-5}\frac{\mathrm{kg}}{\mathrm{s^3}} = 7.7 \times 10^{-12}\frac{\mathrm{kg}}{\mathrm{s^3}},$$
and we are done. 'Elementary, my dear Watson.'

§ Vectors. Vectors are the principal (algebraic) objects of study in the subject called Linear Algebra, of which there are several approaches – a non-abstract approach considering and working only with common examples of vectors, and an *algebraic* approach, which gives profound generality, strength, and applicability of the theory; it comes with a catch, however, – it is abstract and, for the untrained mind, difficult. The collection of vectors is called a vector space.

The space we live in, based on classical physics, allows us to move in three spatial dimensions – choosing some point as the origin of our coordinate system (vector space), as well as the three spatial directions, we may describe any point in space by three coordinates $x, y, z \in \mathbb{R}$. It is convenient to define an object
$$\mathbf{r} := \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

[3]If you reside on Earth, that is; if this is not the case, please try your best to make contact with us – we have been searching for you for decades, but without avail.

called a vector; the set of all such objects, that is, arising from all real numbers $x, y, z \in \mathbb{R}$, is called the Euclidean 3-space, denoted $\mathbb{R}^3$. In general, something only becomes a space if there is some way to combine two elements of the set into another element of the set, as well as scale elements by numbers (this is the mathematical 'definition'). This is accomplished by *defining*[4] vector addition $+_{\text{vect}} : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ coordinate-wise, that is, for vectors $\mathbf{r}, \mathbf{r}' \in \mathbb{R}^3$, their sum is defined as

$$\mathbf{r} +_{\text{vect}} \mathbf{r}' = \begin{bmatrix} x \\ y \\ z \end{bmatrix} +_{\text{vect}} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} := \begin{bmatrix} x + x' \\ y + y' \\ z + z' \end{bmatrix},$$

where $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the 'normal' real number addition. And scalar-vector multiplication $\cdot_{\text{sv}} : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$ as multiplying component-wise, that is, for a real number $\lambda \in \mathbb{R}$ and a vector $\mathbf{r} \in \mathbb{R}^3$, their product is defined as

$$\lambda \cdot_{\text{sv}} \mathbf{r} = \lambda \cdot_{\text{sv}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} := \begin{bmatrix} \lambda \cdot x \\ \lambda \cdot y \\ \lambda \cdot z \end{bmatrix},$$

where $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the 'normal' real number multiplication. For notational simplicity, we (almost) always write $+_{\text{vect}}$ simply as $+$ and, likewise, $\cdot_{\text{sv}}$ as $\cdot$, which is not expected to cause confusion to the reader if the following important convention is followed:

> 'Vectors must **ALWAYS** be denoted with arrows on them.'

(Whereas in typesetting we use bold.) Unless the reader happens to be a mathematician, there are no exceptions. Both square brackets [...] and angular parentheses (...) may be used to denote vectors in $\mathbb{R}^3$ – the choice is purely æsthetic.

The notation $\mathbf{r} - \mathbf{r}'$ is understood to mean $\mathbf{r} +_{\text{vect}} (-1) \cdot_{\text{sv}} \mathbf{r}'$.

In Euclidean geometry the distance between two points (which are equivalent to vectors from the origin as the reader may observe) are separated by some distance; intuitively, this distance should be the length of the line connecting them. By simple geometrical argumentation it follows that the distance between two points (a, b, c) and (a', b', c') is

$$d = \sqrt{(a' - a)^2 + (b' - b)^2 + (c' - c)^2}.$$

Intuitively, there is only one line segment (with direction) from point (a, b, c) to point (a', b', c'); if we move this line without changing the direction to start from the origin, then clearly what is obtained is a vector $\mathbf{u}$ in Euclidean 3-space defined as

$$\mathbf{u} = \begin{bmatrix} a' - a \\ b' - b \\ c' - c \end{bmatrix}.$$

And the metric for the magnitude $u = |\mathbf{u}|$ of this vector will simply be the distance between the two points, that is, $|\mathbf{u}| = d$ as before. Generalising this example, for any vector $\mathbf{r} = [x, y, z] \in \mathbb{R}^3$, the magnitude is defined to be $|\mathbf{r}| := \sqrt{x^2 + y^2 + z^2}$.

§ Centre of mass. Consider masses $m_1, m_2, ..., m_N$ (measured in kilograms, of course) located at points $\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N$ in Euclidean 3-space. If the masses are connected by arbitrarily light but rigid rods, then this 'body' will move as one; the reader is expected to already be familiar with the fact that its translational motion may be given by the translation of its centre-of-mass. Let the total mass be $M = m_1 + m_2 + \cdots + m_N = \sum_{i=1}^{N} m_i$.

[4]Yes, is is necessary to define it. It is, of course, an extraordinarily 'natural' definition, however, it is *not* unique – there exist other perfectly valid definitions; nor does it follow from the 'regular' addition (plus) on integers – vectors are not real numbers. Even worse, if you consider a different vector space, then addition *has* to be different for the addition function (binary operation) to make sense.

The *centre of mass*, denoted $\mathbf{R} \in \mathbb{R}^3$, is defined as the point where the contributions from each weighted position vector $m_i\mathbf{r}_i$ sum to zero – looking from this point, the 'contribution' of mass is equal in all directions; in mathematical terms this condition becomes

$$\sum_{i=1}^{N} m_i(\mathbf{r_i} - \mathbf{R}) = 0.$$

Solving for $\mathbf{R}$,

$$\sum_{i=1}^{N} m_i\mathbf{r_i} - m_i\mathbf{R} = 0 \quad \Longrightarrow \quad \sum_{i=1}^{N} m_i\mathbf{r_i} - \sum_{i=1}^{N} m_i\mathbf{R} = 0 \quad \Longrightarrow \quad \sum_{i=1}^{N} m_i\mathbf{r_i} = \mathbf{R}\sum_{i=1}^{N} m_i,$$

where the second sum is simply M, and hence

$$\mathbf{R} = \frac{1}{M}\sum_{i=1}^{N} m_i\mathbf{r_i}.$$

§ Supplement – algebraic theory of vector spaces. This material is **purely** for readers interested in the rigorous theory – it is absolutely **not** necessary for parsing course material. The theory begins with a vector space – the set where all the objects of study, namely, *vectors*, lie in; the vector space is essentially a set with two operations $\oplus$ and $\odot$ on it that behave 'nicely' with respect to each other. The emphasis is on *any* set – there is no restriction on the form of the objects; they may be 3-tuples of real numbers (i.e. the vectors covered before in this text as the reader may recall), functions from *some* set onto the real numbers, polynomials, matrices, and even more wild examples.

A vector space consists of two parts – firstly, a set V with a binary relation $\oplus : V \times V \to V$ (namely, a function that takes a pair of elements of the set V and returns an element in V; we define this function – it is not given by some sentient being), where there exists an element $\mathbf{0} \in V$ called the[5] *zero vector* that behaves as an identity with respect to $\oplus$, that is, for any vector $\mathbf{v} \in V$ we have $\mathbf{v} \oplus \mathbf{0} = \mathbf{v}$ and $\mathbf{0} \oplus \mathbf{v} = \mathbf{v}$ (adding the zero vector to any other vector from either side does not change the vector).

As the reader may suspect, a plus always comes with a minus, where two such vectors cancel – formally, for any vector $\mathbf{v} \in V$ there must exist some other vector $\mathbf{u} \in V$ such that $\mathbf{u} \oplus \mathbf{v} = \mathbf{0}$ (and, similarly, the other way $\mathbf{v} \oplus \mathbf{u} = \mathbf{0}$) – that is, the vector $\mathbf{u}$ is the *additive inverse* of $\mathbf{v}$; what would be suitable notation for it? The reader is reminded that for any integer n, adding $-n$ to it would give zero; here the same notation is adopted: we denote $\mathbf{u}$ by $-\mathbf{v}$.

So much for the existence of special vectors; now, addition should behave 'nicely'; in this context[6] it means that it does not matter where we put parentheses: for any three vectors $\mathbf{v}, \mathbf{u}, \mathbf{w} \in V$ the expressions $\mathbf{v} \oplus (\mathbf{u} \oplus \mathbf{w})$ and $(\mathbf{v} \oplus \mathbf{u}) \oplus \mathbf{w}$ mean the same thing

$$\mathbf{v} \oplus (\mathbf{u} \oplus \mathbf{w}) = (\mathbf{v} \oplus \mathbf{u}) \oplus \mathbf{w},$$

and hence parentheses may be omitted and the product simply denoted by $\mathbf{v} \oplus \mathbf{u} \oplus \mathbf{w}$. And the other property is that the order of $\mathbf{v}$ and $\mathbf{w}$ around a plus does not matter: $\mathbf{v} \oplus \mathbf{w} = \mathbf{w} \oplus \mathbf{v}$ for all vectors $\mathbf{v}, \mathbf{w} \in V$. All these properties make V into an algebraic object consisting of a set V with *structure* (the operation and its properties) called an *Abelian group structure* – the triple $(V, \oplus, \mathbf{0})$ is called an *Abelian group*.[7]

[5] The definite article *the* is applicable here because it may easily be shown that the zero vector is unique – if two vectors behave as zero vectors, then they must be identical.

[6] And if the student pursues studies of abstract algebra (at this university, the courses are Group Theory, Algebraic Structures, and Advanced Algebraic Structures of the mathematics bachelor programme), it will be evident that this applies to an absurd variety of contexts.

[7] The reader may wonder if all the properties make V into an Abelian group, what then is just a *group*? The answer is simple: the last property is key – in an Abelian group the order around $\oplus$ does *not* matter, whereas in a regular group, it does ($\mathbf{v} \oplus \mathbf{w} \neq \mathbf{w} \oplus \mathbf{v}$ in general).

This does not yet make V into a vector space, however – we arrive at the second part of a vector space. As the reader saw before, Euclidean 3-space vectors may be 'multiplied' by a real number (a scalar); in general, the number does not even have to be real, however, for a vector space to have some nice properties,[8] the scalars have to behave almost identically to real numbers. The definition of what exactly is meant by this is quite complicated and not too illuminating, hence it shall be omitted. We call a triple $(F, +, \cdot)$ a *field* (of scalars), where F is a set of objects called scalars, $+ : F \times F \to F$ and $\cdot : F \times F \to F$ are binary operations if they behave basically as addition and multiplication of real numbers. In further study, the reader is advised to treat F as the set of real numbers $\mathbb{R}$.

Now, for the second part of a vector space. We also define an operation $\odot : F \times V \to V$ that to each pair of a scalar and vector assigns some vector in V – 'multiplies it by a scalar' as it were. The 'structure' this operation must satisfy is that multiplying a vector by a product of scalars does the same as first multiplying it by one and then the other: for all $\lambda, \mu \in F$ we have $(\lambda \cdot \mu) \odot \mathbf{v} = \lambda \odot (\mu \odot \mathbf{v})$. Similar results hold for a sum of scalars and a sum of vectors:

$$(\lambda + \mu) \odot \mathbf{v} = \lambda \odot \mathbf{v} \oplus \mu \odot \mathbf{v},$$

and

$$\lambda \odot (\mathbf{v} \oplus \mathbf{w}) = \lambda \odot \mathbf{v} \oplus \lambda \odot \mathbf{w}$$

for all $\lambda, \mu \in F$ and $\mathbf{v}, \mathbf{w} \in V$, of course. For F to be a field, there must exist a scalar $1_F \in F$ such that multiplying any other scalar λ by it gives back λ (similarly to the real number 1). We likewise require that it does not change the vector: $1_F \odot \mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in V$. We say that this structure on V makes it into a *left-F-module*. And we now have a vector space; voila!

1.1 Definition. A vector space V over F is a sextuple $(V, F, \oplus, \odot, \mathbf{0}, 1_F)$ with V some set, $(F, +, \cdot)$ a field, $\mathbf{0} \in V$ and $1_F \in F$ and operations $\oplus : V \times V \to V$ and $\odot : F \times V \to V$ that satisfy the following properties for all $\mathbf{v}, \mathbf{u}, \mathbf{w} \in V$ and $\lambda, \mu \in F$:

(M1) 'commutativity of addition' $\mathbf{v} \oplus \mathbf{w} = \mathbf{w} \oplus \mathbf{v}$;

(M2) 'associativity of addition' $\mathbf{v} \oplus (\mathbf{u} \oplus \mathbf{w}) = (\mathbf{v} \oplus \mathbf{u}) \oplus \mathbf{w}$;

(M3) 'zero vector is additive identity' $\mathbf{0} \oplus \mathbf{v} = \mathbf{v}$;

(M4) 'existence of additive inverses' for all $\mathbf{v} \in V$ there exists $-\mathbf{v} \in V$ such that $\mathbf{v} \oplus (-\mathbf{v}) = \mathbf{0}$;

(M5) 'associativity of multiplication' $(\lambda \cdot \mu) \odot \mathbf{v} = \lambda \odot (\mu \odot \mathbf{v})$;

(M6) 'scalar unit is multiplicative identity' $1_F \odot \mathbf{v} = \mathbf{v}$;

(M7) 'distributivity I' $\lambda \odot (\mathbf{v} \oplus \mathbf{w}) = \lambda \odot \mathbf{v} \oplus \lambda \odot \mathbf{w}$;

(M8) 'distributivity II' $(\lambda + \mu) \odot \mathbf{v} = \lambda \odot \mathbf{v} \oplus \mu \odot \mathbf{v}$.

And we are done.

[8] Having nice properties is equivalent to saying that there *exists* a proof for a lemma describing the property. Proofs of several extremely important theorems requires well-behaved scalars.

CHAPTER II
RELATIVITY

§2. The principle of relativity

§ Galilean transformations. All processes in nature must be described in some coordinate system called the *frame of reference*; such a frame satisfying a property of not being acted upon by external forces is called *inertial*. Given any two reference frames S and S' whose origins coincide at time zero, where S' moves uniformly with respect to S, the following implication is true: S' is intertial *if and only if* S is intertial. The reader may observe that it is always possible to construct another inertial reference frame given some other.

Consider a particle with position $\mathbf{r} = [x, y, z]$ in coordinate system S (inertial reference frame). Let another intertial reference frame S' move away from S in the positive x direction with velocity β; at time $t = 0$ the frames coincide. The position of the particle in the S' frame is thus given by

$$\begin{aligned} t' &= t \\ x' &= x - \beta t \\ y' &= y \\ z' &= z \end{aligned}, \tag{2.1}$$

which are known as the *Galilean transformations*. The reader may argue that the equations are not of true generality for they assume motion of S' to be in the x direction as opposed to in an arbitrary (non-cardinal) direction; this is in fact not so for given such S' we may simply reorient the coordinate systems to have the x-axes point in the direction of S' velocity from which the given scenario follows. The time derivative and the repeated time derivative yield the following equations.

$$\begin{aligned} v'_x &= v_x - \beta & \qquad a'_x &= a_x \\ v'_y &= v_y & a'_y &= a_y \\ v'_z &= v_z & a'_z &= a_z \end{aligned}$$

The implication is that forces act identically in inertial reference frames; velocities are additive by Galilean relativity. The latter point is challenged by Einstein's relativity, the effects of which become noticeably apparent only at very high velocities.

§ The principle.[1] The *principle of relativity of Einstein* may be stated as:

> 'Equations expressing the laws of nature are *invariant* under coordinate and time transformations between inertial systems.'

Forces between distant bodies are not instantaneous; there exists some interval of time that change due to the force may be observed. This leads to the *maximum velocity of the propagation of the interaction*; if motion of a particle beyond this velocity were possible, then an interaction occurring via this particle would manifest itself with velocity faster than the maximal, which is a contradiction to the existence of maximal velocity, hence such motion is impossible.

2.1 Proposition. *The maximum velocity of propagation is the same in all inertial reference frames.*

Proof. Consider frames S and S' with maximal velocities c and $c' = \sigma c$, respectively ($\sigma \in \mathbb{R}$). Position a source of signal at the origin of S and let S' move with the signal in the positive

[1] The following account is based on the following references: T.A. Moore, *Six ideas that shaped physics*, D. Morin, *Special relativity for the enthusiastic beginner*, and partly from the early sections of L.D. Landau and E.M. Lifshitz, *Course of theoretical physics* (aka the course), vol. 2, *The classical theory of fields*. The reader is advised to consult Morin for deepening their knowledge and Landau–Lifshitz for an illuminating account of the theory once the basics have been learnt with proficiency.

direction. Consider two detectors 1 and 2 moving away from the source in the positive direction with velocities c and $(1/2)c$, respectively. In the S frame the positive detector is never reached, whereas the negative is.

The signal is stationary in the S' frame. As observed from the S' frame, the detectors are either stationary or move. Since the laws must be invariant, so must be their consequences, hence detector 1 must either be stationary or move away from the origin in S', whereas detector 2 must move towards the origin. Let the velocity (in terms of c) in the S' frame be given by $f(\epsilon)$, where ϵc is the velocity in the S frame. It follows that $f(1/2) \leq \sigma \leq 1$.

Inductively repeating the argument on the detector 2 velocity being $(1 - 1/2^k)c$ for k natural, it follows that for any number $1 - 1/2^k$ smaller than 1 to an arbitrarily tiny degree, σ is bounded between $f(1 - 1/2^k)$ and 1; since $f(\epsilon) \to 1$ as $\epsilon \to 1$, it follows that this is possible only if $\sigma = 1$, completing the proof.[2] □

§ Michelson–Morley. The Michelson–Morley experiment is presented henceforth. By Galilean relativity, it had been postulated that there exists a fixed medium permeating all space – the *æther*; it is hence 'the' primary stationary reference frame of the universe, with respect to which time is absolute. The Michelson–Morley experiment, however, disproved the existence of it, as will be evident soon; from this the other postulate of relativity follows.

Now, onto the experiment. Consider a source and detector on a moving platform of length L in some fixed space; in the extreme cases, the platform may move parallel to the direction of the signal or perpendicular to it. Let the signal velocity be v_s and the platform velocity be v_p. The reader is now asked to find the time for the signal to each the detector in each case (let the times be t_1 and t_2, respectively); from that the difference in time may easily be computed – the experiment is concerned with this quantity.

The time in the parallel case is simply the sum of times resulting from transformed velocity in the forward and backward direction, that is, $v_s - v_p$ and $v_s + v_p$, respectively. It is obtained that

$$t_1 = \frac{L}{v_s - v_p} + \frac{L}{v_s + v_p} = \frac{2Lv_s}{v_s^2 - v_p^2}.$$

In the perpendicular case the velocity component in the horizontal direction may easily be computed by using the Pythagorean Theorem to be $\sqrt{v_s^2 - v_p^2}$; clearly, the trips are equally long, hence

$$t_2 = \frac{2L}{\sqrt{v_s^2 - v_p^2}}.$$

Notably, the times are not equal. Therefore the difference in time is given by

$$t_1 - t_2 = \frac{2L}{v_s}\left(\frac{1}{1 - v_p^2/v_s^2} - \frac{1}{\sqrt{1 - v_p^2/v_s^2}}\right).$$

The result may be simplified in the limiting case, where the source velocity is far larger than the platform velocity $v_s \gg v_p$; this is the case with sound on a train and light on Earth moving around the Sun. Using the approximation $(1+x)^n \approx 1 + nx$ given in §1, it is found that $(1-\epsilon)^{-1} \approx 1 + \epsilon$ and $(1-\epsilon)^{-1/2} \approx 1 + \epsilon/2$. Hence the time difference in this limiting case becomes

$$t_1 - t_2 \approx \frac{2L}{v_s}\left[\left(1 + \frac{v_p^2}{v_s^2}\right) - \left(1 - \frac{v_p^2}{2v_s^2}\right)\right] = \frac{Lv_p^2}{v_s^3}.$$

The experiment was conducted by using two optical rails of $L = 10\,\mathrm{m}$ at right angles to each other using light as the signal carrier; using an interferometer, the time difference

[2] The statement of the proposition appears in Landau & Lifshitz, however, without proof – it is of my own making. The critical step was revealed to me in a dream.

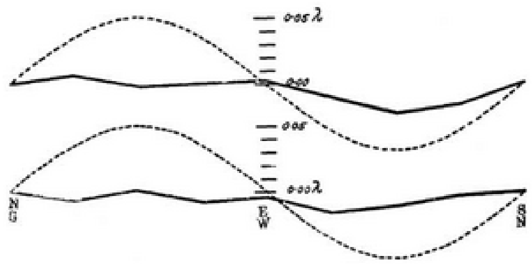


Figure 2.1: Results of the Michelson–Morley experiment. The solid lines indicate data, whereas the dashed – the theoretical prediction *rescaled to 1/8 (smaller)* of the values. The upper curve is for the observations done at noon, whereas the lower – in the evening.

is measured directly and is highly insensitive to error in apparatus – small differences in the lengths of the rails do not change the result significantly. Hence in this scenario $v_s = c$ and $v_p = v$ is the (unknown) velocity of the movement of Earth through the æther. Should such æther exist, then the observed time difference should be Lc^1/v^3. The experiment may measure the time difference by measuring how much the wavefronts are shifted with respect to each other – the *fringe shift* (to parse the details, great familiarity with optics is required). The expected fringe, factoring in only the movement of Earth in different directions during the year, is approximately 0.40 fringe for the yellow light used; the observed fringe was approximately 0.02 fringe and independent upon the rotation of the apparatus in the room, time of day (rotation of Earth upon its axis), and the time of the year (position and velocity of Earth on its orbit) as shown in Fig. 2.1.

The reader may now rightly draw the conclusion that no such (stationary) æther exists.[3] Hence the following proposition is experimentally validated and shall serve as a principal postulate henceforth.

2.2 Proposition. *All inertial reference frames are 'equivalent' – there is no reference frame of universal time for time is measured separately in each frame.*

The founding principles of mechanics following the principle of relativity (of Einstein) have now been laid; the reader is now set to begin the long and perilous but cathartic journey of exploration of such mechanics – *relativistic mechanics.*

[3] For further details on ruling out alternative formulations of an æther see Morin, § 1.1.2.

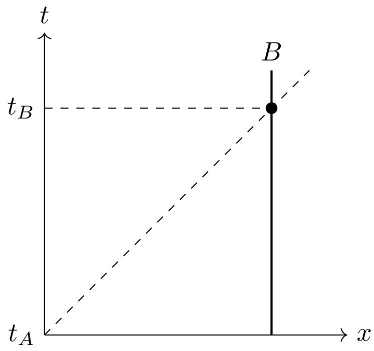
§3. Coordinate time

§ Clocks. In relativity measurements of time are those fraught with more trouble than any other quantity; the problem lies in defining what exactly is meant by time in the context. From the principle of relativity (of Einstein), the time coordinate must be treated in exactly the same manner as a space coordinate – this leads to *coordinate time* as the time value in some inertial reference frame.

Realistically speaking, a clock may only measure the time of an event if the event occurs at the clock; hence we would like our thought-experimental setups to contain a clock at every spatial coordinate, where the clocks should, of course, agree on the time (otherwise there is no point of the clock). This is achieved by synchronising the clocks:

> 'Two clocks are defined to be synchronised if the distance between them is computed correctly according to their time readings of a light flash sent from one to the other.'

That is, for clocks A in the same intertial reference frame, if clock A measures time to be t_A and clock B measures t_B, then the computation of the distance between the clocks using their readings must give $c(t_A - t_B)$. If the distance is not computed to be so, then the clocks are not synchronised – if t_B is an hour (let alone a year) behind, then the computation would yield nonsense.



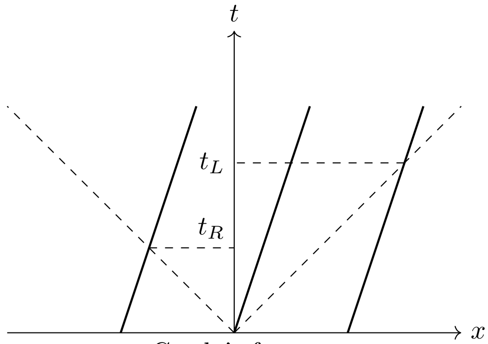
In order to plot light on the diagonal in the spacetime diagram, it is convenient to pick the distance unit to be *light seconds* (time remains in seconds) – one light second is the distance light travels in a second. In the diagram above – a *spacetime diagram* – the points occupied by an object at any instant of time are called its *worldline*. Each observer comes equipped with its own intertial reference frame and a plethora of synchronised clocks in its frame.

§ Fundamental effect – loss of simultaneity. Consider a train wagon moving in the $+x$ direction (also denoted $\hat{x}$); in the middle of the platform lies a light source, and observer Markuss stands on the platform, whereas on the ground observer Ceyda watches the platform of length $2\ell'$ pass by with velocity v. Clearly, Markuss observes the light hit both end walls at the same time; what about Ceyda? The reader should now attempt to draw the setup in Ceyda's frame.

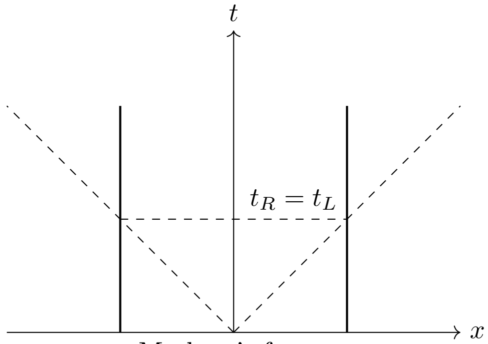
The maximal velocity of propagation in Ceyda's frame must remain c, as given by the first postulate (Proposition 2.1 in §2); nevertheless, the walls of the wagon are observed to move with velocity v in the $+x$ direction. The time required for light to reach the wall will depend on the separation velocity between the photon and the wall, which is $c + v$ for the posterior and $c - v$ for the anterior (front) wall of the wagon.[4] The reader now has

[4]Emphasis is given to the fact that this velocity, which may exceed c, is the *separation* velocity, that is, it is not the velocity of any particle or light signal in the frame for such velocity it may not exceed the maximal value as given by the first postulate. The separation velocity, however, may assume values up to $2c$.

sufficient information to compute the times t_L and t_R to reach the posterior and anterior walls, respectively, to be $t_L = \ell'/(c+v)$ and $t_R = \ell'/(c-v)$; notably, $t_R \neq t_L$ – *simultaneous events in one frame are no longer simultaneous in another*. The result is illustrated by the diagrams below.



Ceyda's frame



Markuss's frame

§ Fundamental effect – rear-clock-ahead. Suppose Markuss positions synchronised clocks at each of the walls; we wish to place the light source such that Ceyda observes both walls to be hit simultaneously; let the length of the wagon in Markuss's frame be L. The reader may now easily compute that in Markuss's frame the lengths are $L\frac{c+v}{2c}$ and $L\frac{c-v}{2c}$ for the rear and front ends, respectively (and their sum is L).

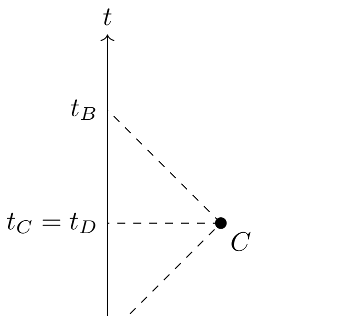
In the train frame to reach the rear clock light must travel for an additional $t_\Delta = \Delta L/c$ seconds, where ΔL is the length difference, namely,

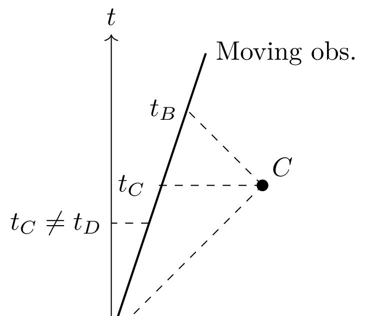
$$t_\Delta = \frac{1}{c}\Delta L = \frac{1}{c}\left[L\frac{c+v}{2c} - L\frac{c-v}{2c}\right] = \frac{Lv}{c^2}.$$

Now, if Ceyda observes the clocks upon being hit by photons, the time values will differ by Lv/c^2 seconds – *the rear clock is ahead*.[5]

There are two other fundamental effects of relativity – time dilation and length contraction –, that follow immediately from the two postulates (Propositions 2.1 and 2.2) discussed previously, however, the discussion of which shall have to wait until the next session. The enthusiastic reader is advised to consult Morin §§ 1.3.2 & 1.3.3.[6]

§ The radar method. A very specific type of diagram will now be discussed as given below; let the diagram on the left be the Other frame.





[5]In some sense this effect is equivalent to the loss of simultaneity; for pedagogical reasons the two have been separated in these notes. Note also that L is the distance *in the frame of the train*!

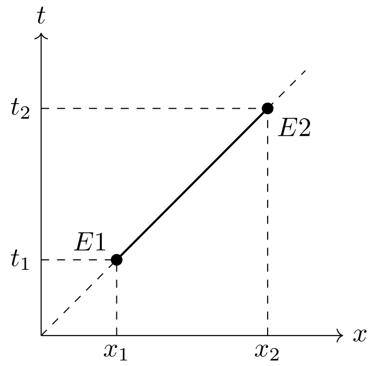
[6]A formal treatment is, of course, contained in The course, Landau & Lifshitz, vol. 2, ch. I, §§ 1–5 (see Footnote 1 in §2); the chapter in its 23 pages, the last 8 of which already contain material beyond this course, contains this relativity course in its entirety. Should the reader wish to engage in the formality, it must be noted that all effects will have to be derived manually – the book does not spoon-feed.

Event A marks a laser signal being sent in the $+x$ direction until it reaches a spaceship at event C and is reflected back to the observer at event B. In the frame of the observer on the left, the event C occurs half-way between events A and B, hence $t_C = \frac{1}{2}(t_B + t_A)$ and $x_C = \frac{1}{2}(t_B - t_A)$ in the frame of the observer of the Other frame (in SR units); now in the Home frame, the observer is moving towards the spaceship and from the diagram the reader is forced to conclude that event C occurs later than half-way between A and B.

It is now evident that the time difference between two events – the *coordinate time* (difference) – depends on the frame of reference. Upon drawing many a diagram, the reader shall become accustomed to this concept.

§4. The spacetime interval

§ Minkowski spacetime. Consider intertial reference frames S and S' of coordinates (x, y, z, t) and (x', y', z', t'), where S' moves in the $+x$ direction and the frames coincide at $t = t' = 0$. Consider a light flash sent from (x_1, y_1, z_1, t_1) marked by event $E1$ to (x_2, y_2, z_2, t_2) – event $E2$ – in the S frame as shown in the diagram below.



The reader may now easily compute the separation between the two events in space to be given simply by the Euclidean 3-space distance $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$; recalling the definition of synchronised clocks from §3, the separation is also given by the distance that light travelled during this time $c(t_2 - t_1)$. The equality obtained yields

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - c^2(t_2 - t_1)^2 = 0. \tag{4.1}$$

The trivial replication of reasoning in the S' frame is left to the reader, namely, to show that for events $E1$ and $E2$ in the S' frame the following equality

$$(x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2 - c^2(t'_2 - t'_1)^2 = 0. \tag{4.2}$$

is likewise obtained. The reader will now spot the curious invariance (in coordinates) between the two formulae – the left-hand side is zero no matter the frame; it is hence suitable to define the quantity

$$s_{12} := \sqrt{c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2}, \tag{4.3}$$

called the *spacetime interval* (or simply *interval* for short) between *any two* events 1 and 2; the reader will observe that this quantity may become imaginary – taking event 1 to be $E1$, the reader is asked to indicate in the diagram above the regions of events producing positive, zero, and imaginary values (with $t_2 \geq t_1$). For an infinitesimally short separation between events 1 and 2, the differential form

$$ds^2 = c^2t^2 - dx^2 - dy^2 - dz^2, \tag{4.4}$$

is obtained, which is called the Minkowski metric.

4.1 Lemma. *The spacetime interval is frame-independent. That is, given frames S and S' as above, it follows that $ds^2 = ds'^2$, and thus $s = s'$.*

Proof. [7] Consider frames S and S' as given above with the Minkowski metric transformed by the function a, that is, $ds^2 = a ds'^2$; by the principle of relativity, or specifically Proposition 2.2, a cannot depend on the coordinates in the frames; similarly, dependence on angles is out, leaving only the possibility of dependence on the magnitude of the velocity of S'.

[7] The following elementary proof is reproduced from § 2, Landau & Lifshitz.

Let S be the home inertial reference frame, whereby S_1 and S_2 move with velocity vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ relative to S from the origin; let the transformations of the Minkowski metrics between frames be $ds^2 = a(v_1)ds_1^2$, $ds^2 = a(v_2)ds_2^2$, and $ds_1^2 = a(v_{12})ds_2^2$, where v_{12} is the velocity of the S_2 frame in the S_1 frame. Hence $a(v_{12}) = a(v_1)/a(v_2)$; now, there is angle-dependence in v_{12}, which has been ruled out from $a(v_1)$ and $a(v_2)$, leaving only the possibility of a being a constant function; the ratio thus reduces to unity, from which it follows that $ds_1^2 = ds_2^2$. By symmetry also $ds^2 = ds_1^2 = ds_2^2$; and returning to S and S', $ds^2 = ds'^2$. Now, since s (and s') can only be positive or imaginary, it follows that $s = s'$, completing the proof. □

Let us slightly explore the geometry of the Minkowski metric before continuing; it is clear that the value of the spacetime interval lets us classify events depending on whether the value s_{12} is real, zero, or imaginary. This classification leads nicely to an answer to the following problem, *which the reader is now asked to solve.*

4.2 Problem. Given two events 1 and 2 at (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2), respectively, in some frame S, do there exist intertial reference frames such that

(a) the events occur at the same place;

(b) the events occur at the same time?

Namely, what are the conditions on the coordinates and time in each case for the frame to exist? It may be helpful to consider the Euclidean distance (squared) $\ell_{12}^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$ and time separation $t_{12} = t_2 - t_1$; that is, $s_{12}^2 = c^2 t_{12}^2 - \ell_{12}^2$.

Upon solving (or *at the very least* attempting to do so), the reader will find the answer given in the footnotes.[8] The third option, namely, when $s_{12} = 0$, has already been covered in text before – the events are connected by a light ray. This classification shall be explored further later on in the course.

§ Fundamental effect – time dilation. In this paragraph the famous 'light clock' shall be explored. Consider a train moving with velocity v in the ground frame with mirrors on the floor and ceiling separated by a distance h and light source between them pointed upwards; a light pulse thus bounces between the mirrors indefinitely. In the frame of the train the light takes $t_{\mathrm{tr}} = h/c + h/c = 2h/c$ seconds to traverse a round trip; now, in the ground frame the path of light becomes a triangle with the vertical velocity component being $\sqrt{c^2 - v^2}$, obtaining a round trip time of $t_{\mathrm{gr}} = 2h/\sqrt{c^2 - v^2}$. The ratio of the times hence becomes

$$\frac{t_{\mathrm{gr}}}{t_{\mathrm{tr}}} = \frac{c}{\sqrt{c^2 - v^2}} = \frac{1}{\sqrt{1 - v^2/c^2}} =: \gamma,$$

where we denote the dimensionless fraction on the right by γ (note that $\gamma \geq 1$),[9] called the *Lorentz factor*, hence obtaining that $t_{\mathrm{gr}} = \gamma t_{\mathrm{tr}} \geq t_{\mathrm{tr}}$, meaning that *as observed in the ground frame, the time on the train runs slower than observed on the train.*

Beware that the time ratio applies only for events occurring at the same place in the train's frame; likewise, for a (stationary) light clock on the ground, the observer on the train would observe it running slower than the clock on the train by the same factor γ. Recall that from the second postulate of special relativity (Proposition 2.2 in §2), all frames must be equivalent and thus see one-another's clocks run slow.[10]

[8] Here follows the answer to Problem 4.2; *the reader is strongly advised* against *reading the following lines if they have not attempted the problem!*

Consider some other system S' (with t'_{12} and ℓ'_{12}), which we will attempt to describe in each case. (a) Spatially unseparated events must satisfy the following condition $\ell'_{12} = 0$, hence $s'^2_{12} = c^2 t'^2_{12} > 0$; because of the invariance of the spacetime interval $s = s'$, it follows that the condition is that s_{12} must be *real*. (b) Similarly, simultaneous events require $t'_{12} = 0$, from which the condition that s_{12} be *imaginary* is obtained. The case $s_{12} = 0$ is given in text.

[9] In fact, it is evident that $\gamma = 1$ if and only if $v = 0$, from which it follows that both frames coincide.

[10] The interested reader may wish to ponder the following question (aka *the twin paradox*): consider twins Lotte and Renske; at some point in her life, Renske flies on a spaceship to a nearby star and back (quickly, that is); upon Renske's return, are the twins the same age? If not, who is older?

§ Fundamental effect – length contraction. Consider a similar setup as before, but with mirrors on the rear and front walls; let the length of the train be L_{tr} as measured on the train, and similarly, L_{gr} on the ground (for a moving train). As the reader may easily compute, for a round trip of a light pulse, $t_{\mathrm{tr}} = 2L_{\mathrm{tr}}/c$, whereas

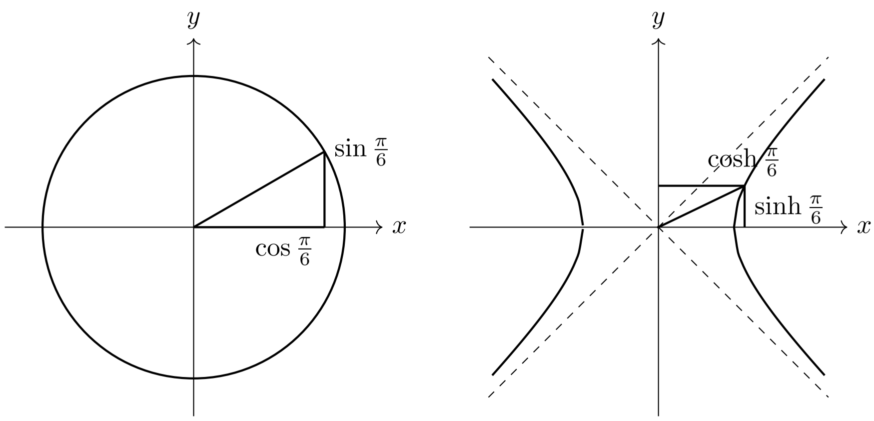
$$t_{\mathrm{gr}} = \frac{L_{\mathrm{gr}}}{c-v} + \frac{L_{\mathrm{gr}}}{c+v} = \frac{2cL_{\mathrm{gr}}}{c^2-v^2} = \gamma^2\frac{2L_{\mathrm{gr}}}{c}.$$

Combining this result with $t_{\mathrm{gr}} = \gamma t_{\mathrm{tr}}$, we obtain

$$\gamma^2\frac{2L_{\mathrm{gr}}}{c} = \gamma\frac{2L_{\mathrm{tr}}}{c},$$

and hence $L_{\mathrm{gr}} = \frac{1}{\gamma}L_{\mathrm{tr}}$, namely, *the length of the train, as observed from the ground, is smaller than as measured on the train.* This phenomenon is called the *Lorentz–FitzGerald contraction.*

§ Supplement – hyperbolic trigonometry. The reader may recall that ordinary trigonometrical functions of $\sin t$ and $\cos t$ relate principally to lengths in the unit circle; they are functions entirely of the counter-clockwise angle attained with respect to the $+x$ direction. Importantly, the curve $(\cos t, \sin t)$ traces out the unit circle $x^2 + y^2 = 1$ for $0 \le t < 2\pi$; similarly, $(\cosh t, \sinh t)$ trace out the left side of the unit hyperbola $x^2 - y^2 = 1$ as $-\infty < t < +\infty$. The transition to the hyperbolic setting is nothing more than replacing the unit circle with a unit hyperbola, however an important detail is revealed – the argument of the hyperbolic functions is *not* an angle, but *the area traced out by the angle* (from 0 to t bounded by radius, x-axis, and the hyperbola); let the reader compare the following diagrams with $t = \pi/6$.



Now for a more algebraic treatment. The reader may show that the functions

$$\sinh t = \frac{1}{2}(e^x - e^{-x}) \qquad \cosh t = \frac{1}{2}(e^x + e^{-x}) \tag{4.5}$$

trace out the $(x > 0)$ unit hyperbola $x^2 - y^2 = 1$ as $-\infty < t < +\infty$. Notice that $\sinh t$ is an odd, whereas $\cosh t$ even function, and $\sinh t + \cosh t = e^t$; it is hence clear that the two hyperbolic functions split the exponential function into an odd and even part. Various trigonometric identities hold for hyperbolic functions, with the most fundamental $\cos^2 t + \sin^2 t = 1$ replaced by $\cosh^2 t - \sinh^2 t = 1$; given the explicit formulae in Eq. (4.5), the reader may show that various properties hold.[11]

The reader is now in a position of sufficient knowledge and ability to solve the exercise set of the previous week 'Exploring Minkowski space'.[12]

[11] For further reading, see `https://www.math.purdue.edu/~pvankoug/math266/sinh_cosh_sin_cos.pdf`

[12] The reader may find the answers (before their official publication) in § 4 of vol 2., Landau & Lifshitz.

§5. Proper time

§ Proper time. We call the time measured by a clock travelling in whatever way from event $E1$ to event $E2$ the *proprietary time* (or *proper time* for short) – that is, the clock is present *at* both events (note that this is not necessarily possible for all events). Let S be the ordinary Home frame and consider the frame S' of such a moving clock; now, during an infinitesimally short interval of time dt in S, the motion of the clock is in a straight line, travelling a distance of $\sqrt{dx^2+dy^2+dz^2}$, hence during this interval it describes an intertial reference frame; note that in the S' frame, the clock is at rest $dx' = dy' = dz' = 0$. Using the invariance of the spacetime interval,

$$c^2dt^2 - dx^2 - dy^2 - dz^2 = ds^2 = ds'^2 = c^2dt'^2,$$

whereupon rearrangement, the expression

$$dt' = dt\sqrt{1-\frac{dx^2+dy^2+dz^2}{c^2dt^2}} = dt\sqrt{1-\frac{v^2}{c^2}}$$

is obtained with v denoting the speed of the clock in the S frame; the integral over an interval of time t_1 to t_2 in the S frame yields the proper time $\Delta\tau$ on the moving clock to be

$$\Delta\tau = t_2' - t_1' = \int_{t_1}^{t_2}\sqrt{1-\frac{v^2(t)}{c^2}}\,dt, \tag{5.1}$$

with emphasis on the fact that the speed of the clock $v(t)$ may vary with time – the clock may take any (valid) path.

The reader may recall from §4 the fundamental effect of time dilation; now, observing Eq. (5.1), the proper time $t_2' - t_1'$ is computed to be necessarily smaller than $t_2 - t_1$; in fact, from the definition of the Lorentz factor in §4, the proper time is $\int_{t_1}^{t_2}\gamma(t)^{-1}\,dt$, hence we recover the time dilation $t_2' - t_1' = \gamma\cdot(t_2 - t_1)$, derived before using elementary methods, in the case of constant velocity $v(t) = \text{const}$.

§ Small speed approximation. The reader may recall from §1 the ubiquitous approximation $(1+x)^n \approx 1 + nx$ for small x; applying it to the case where $v \ll c$, we obtain that $(1 - v^2/c^2)^{1/2} \approx 1 - \frac{1}{2}(v^2/c^2)$. For constant v^2, Eq. (5.1) becomes

$$\Delta\tau = t_2' - t_1' = \sqrt{1-\frac{v^2}{c^2}}(t_2 - t_1) \approx \left(1 - \frac{1}{2}\frac{v^2}{c^2}\right)(t_2 - t_1).$$

§ Ranking time. Here we show that $\Delta t \geq \Delta s \geq \Delta\tau$: the time measured between two events (in a light cone) in some coordinate system (inertial ref. frame; *coordinate time*) is greater or equal than the time measured by a clock travelling in a straight line between them (equal to the *interval*), which is bigger or equal than the time recorded by any other clock travelling from one event to the other (*proper time*). Recall that in SR ($c = 1$) units, $\Delta s = \sqrt{\Delta t^2 - \Delta\ell^2} \leq \Delta t$. We now show the other inequality.

5.1 Proposition. *The spacetime interval is the maximal possible value of proper time connecting two events; this occurs for a straight line connecting them.*

Proof. Consider a clock with its own intertial reference frame S centred on the clock that travels from point A to B in a straight line, and a frame S' of a clock travelling between the same points in some other way; this implies that $v(t)$ of the moving clock observed in S is nonzero (can be zero for some but not all t). Integrating,

$$\Delta s_{AB} = t_B - t_A = \int_{t_A}^{t_B} 1\,dt > \int_{t_A}^{t_B}\sqrt{1-\frac{v^2(t)}{c^2}}\,dt = \Delta\tau_{AB},$$

where $\Delta\tau_{AB}$ is the proper time of the non-intertially moving clock, thus proving the statement. □

§6. Dimensional analysis

§ Quantities. Consider a scalar[13] quantity θ with some unit in the SI system (for simplicity, we shall work with SI units rather than arbitrary quantities) that depends in whatever way on other quantities $q_1, q_2, ..., q_n$ also with some units; namely, $\theta = \beta f(q_1, ..., q_n)$ for some function f and dimensionless constant $\beta \in \mathbb{R}$. In the simplest of cases, the function will simply raise the quantities to some powers combined in products (by multiplication; division is multiplication by negative exponents), for example $f(F, \ell, t) = F^1 \ell^1 t^{-1} = [\mathrm{kg\,m\,s^{-2}} \cdot \mathrm{m} \cdot \mathrm{s}^{-1}]$, where F force, ℓ length, and t time. We shall now see that in some cases the function f is unique and determined entirely by the units – knowing the units of quantities $q_1, ..., q_n$ under f, there is only one way to exponentiate them to obtain the correct unit for θ.

6.1 Theorem (π-*theorem*, Bertrand–Buckingham, 1878–1914)**.** *Given n physical variables $q_1, ..., q_n$ with units, such that $1 \leq k \leq 7$ is the total number of basis SI units of the variables, then the number p of dimensionless variables (constructed from $q_1, ..., q_n$) needed to rewrite the equation $\theta = \beta f(q_1, ..., q_n)$ is $p \leq n - k$.*

Equality is obtained if the vectors of the units are independent (k is the rank of the dimension matrix).

The proof is rather difficult to follow and requires linear algebra, working with the units in the style given in Footnote 13. The exponents are unique if $p = 0$. The π-theorem implies that there is no unique way to write the function of the number of variables exceed the number of SI units present, which is rather obvious; on the other hand, the units must be independent – if two variables q_1, q_2 have the same unit, then q_1/q_2 is unitless, and so is $(q_1/q_2)^\alpha$ for every exponent α, hence there is no unique way to write f as a power law.

To algebraically determine uniqueness, the rank of the dimension matrix must be computed – this requires linear algebra. Note that this is rather overkill formalism – in practice, with 2–5 variables, by simply setting up the equations, it will become obvious whether a unique solution exists.

6.2 Example. The radius up to the event horizon (where light may no longer escape the pull) of a black hole is called the Schwarzschild radius R. Come up with variables it should depend on and find an expression (up to dimensionless constant) for it.

i. The question concerns light, and is clearly rather intimately tied to its speed; hence we postulate that one of the variables is the speed of light c. Since the effect is gravitational, we must somehow tie in the gravitational constant G and the mass M of the black hole.

ii. We check whether $R = k \cdot f(c, G, M)$ gives a unique result (k is a unitless real number). We posit

$$R = kc^\alpha G^\beta M^\gamma,$$

where α, β, γ are rational exponents. Then, by plugging in the units[14], we obtain

$$[\mathrm{m}] = \left[\left(\frac{\mathrm{m}}{\mathrm{s}}\right)^\alpha \left(\frac{\mathrm{m}^3}{\mathrm{kg\,s}^2}\right)^\beta (\mathrm{kg})^\gamma \right] = \left[\mathrm{m}^{\alpha+3\beta} \mathrm{kg}^{\gamma-\beta} \mathrm{s}^{-\alpha-2\beta}\right],$$

from which it follows that the exponent for second must be zero: $-\alpha - 2\beta = 0$ or $\alpha = -2\beta$, then $\gamma - 2\beta = 0$ implies $\gamma = -\alpha/2$ and, lastly, the units for meter give $\alpha + 3\beta = 1$, thus

[13] A statement could be made that there are in fact no scalar quantities if we take each quantity to be the pair $(\lambda, \mathbf{J})$, where λ is a real number (the scalar value of a quantity) and $\mathbf{J}$ is 'the canonical SI vector', which consists of 7 numbers that correspond to exponents of the 7 basis SI units $(\mathrm{s}, \mathrm{m}, \mathrm{kg}, \mathrm{K}, \mathrm{A}, \mathrm{mol}, \mathrm{cd})$; e.g. $(-2, 1, 1, 0, 0, 0, 0)$ corresponds to $\mathrm{s}^{-2}\,\mathrm{m\,kg} = \mathrm{N}$. *This is, of course, rather silly.*

[14] If you have (like myself during the tutorial) forgotten the units of G, we proceed as in §1 – the gravitational constant is determined by $F = G \cdot m_1 m_2 / r^2$, where the units are

$$\left[\frac{\mathrm{kg\,m}}{\mathrm{s}^2}\right] = G\left[\frac{\mathrm{kg} \cdot \mathrm{kg}}{\mathrm{m}^2}\right],$$

from which it follows that G must have kg and s^2 in the denominator, with m^3 in the numerator, namely, $G = [\mathrm{m}^3\,\mathrm{kg}^{-1}\,\mathrm{s}^{-2}]$.

$\beta = 1$, $\alpha = -2$, and $\gamma = 1$. Therefore $R = kGM/c^2$, obtaining the correct expression up to a (small) dimensionless constant k.

Dimensional analysis in physics is really a game of examples. If you do not understand, do some more examples and you will soon catch the trick. Mathematically, there is a whole field of dimensional analysis, however, it is far too complicated and for our purposes unnecessarily abstract (see the theorem above...).

§7. First exam preparation

§ How to get a 10 on the exam.[15] In order to prepare successfully for exams, the following quote from a professor here at the University of Groningen is recalled.

> 'Intuition is trained by **intensive** active study during the semester.' – AVK.

And as such the reader may expect to reach proficiency with the content *only if* they have solved numerous problems of the highest difficulty they are able to handle, have actively revised definitions and the theory linking them, including the derivation of all quantities and formulae from first principles (principle of relativity and Propositions 2.1 and 2.2) on the spot (perhaps on a bit of toilet paper, while sitting on the toilet), and, finally, the reader has practised *at least three* past exams (recording the time and their mistakes, of course).

The following exam preparation checklist has been compiled; the reader wishing to do well on the exam is advised to complete this checklist – it is recommended to begin preparation *at least 7 days* before the exam. Upon completing an item, the reader may inscribe a checkmark (*or a smiley face* ☺) in the corresponding box.

1. ☐ Read these supplementary and recapitulation notes in full. If some point proves difficult, attempt to figure out where to seek information in the book(-s) (see Footnote 1); if you fail to find an answer, ask a teaching assistant (*before* the exam).

2. ☐ Given some formula, ask yourself whether you understand its meaning; do you know and understand the derivation? Close the book and these notes and derive it. Do this for all important formulae.

3. ☐ Open an assigned problem at random. Can you solve it? You are given 15 minutes. If you find yourself stuck, begin the following procedure (**without opening the book**):[16]

   i. Do you understand what statement you have to show to be true or what quantity must you find? If not, return to steps 1 & 2.

   ii. Write down every formula or relation you know that contains this quantity or expression (for the first exam there will not be many). Are there quantities in them that have been given to you? If yes, explore whether you could use those formulae.

   iii. If you cannot see any connections between what is given and what must be reached, draw the problem, consider some simpler case, try to get an intuitive understanding on what is happening; what physical processes are involved in your thought-experimental/intuitive setup? Do you know how to state them mathematically (formulae)? Are you able to use such formulae in your problem? Proceed if you can.

   iv. If you reached this point, *it follows that you are desperately stuck.* Not to worry, you will find a way out! **Do not panic!** Recall how you solved exercises in your tutorial sessions; begin writing out things in this manner, and keep writing until you see something – it is important to keep writing and to keep considering new things since you are looking for a connection, and this requires exploring possibilities, i.e. brainstorming. Keep doing this until you find a way. *On the exam itself, it may be helpful to move on to the next problem if you find yourself stuck.*

[15] Without bribing the professor, of course. P.S. Daan enjoys Coca Cola.

[16] That is, you must simulate exam conditions; if you become too comfortable with peeking in your notes, the book, or let alone Google, you will be completely unprepared for dealing with a problem you cannot immediately solve – I will bet €20 that *every single* student will encounter at least one such problem on the exam (or a part of a problem for the best students...).

4. ☐ Pick a past exam, close all your notes and notebooks and set a time for 2 hours (1.5 h if the exam is set for 90 minutes). Attempt to solve this exam; if the timer rings before you are done, change it to a stopwatch and continue, recording how long was required to complete it. Reread your solutions and see if you can spot mistakes; once you are satisfied, *declare your attempt done.* Then look up the solutions and check your answers – a silly mistake you did not spot before is still a mistake (no changing your answers now!).

   Did you solve everything correctly? Excellent. If not, try to figure out the cause for your mistakes – be it algebra, an incorrect diagram, not understanding some concept, failure to remember something, or something other. Read up on this topic until you understand it; solve some book problems on it. Can you now solve the exercise you could not before?

5. ☐ Pick another past exam. Repeat the previous step.

6. ☐ Pick another past exam. Repeat the previous step.

7. ☐ Attempt to identify whether there is anything to improve; attempt to do so until the exam. Solve some more past papers, book problems, or read additional material until you feel prepared; *do not get overconfident*, however: expect that you will get stuck somewhere on the exam – do you think you will find a way out?

This is the end of the checklist. At this point you should be well-prepared for the exam; cross your fingers and hope all works out. I hope the notes and the checklist have proven helpful. And, of course, **the best of luck!**

'May the odds be ever in your favour.' – Effie Trinket.[17]

[17] If Katnis was able to beat the Games, surely you can beat an exam.

§8. Coordinate transformations

§ Drawing the two observer diagram. A frame S' moves with v from S Hone frame. We begin with a simple t-x coordinate system; the t' axis is trivially with slope $1/v$. Then consider a radar sent from $-T'$, which is then reflected and received at T' in the S' frame's $x' = 0$; then, in the S' frame the reflection point is determined to be at $t' = 0$. By continuously varying T', we obtain a straight line of the x'-axis, which lies at a slope of v.

§ Lorentz transformations – geometrical derivation. We begin with a two-observer diagram with an event B and draw the parallelogram $Ot'_BBx'B$ with primes coordinates on the primed axes. The height of B is the time, denoted b, from x' vertically up to B plus the time from $x = 0$ to x', denoted a; by time dilation, $b = \gamma t'_B$, whereas by length contraction the length on the x axis from 0 to x of the point x' is $\gamma x'_B$, and thus $a = \gamma x'_B \tan\theta = \gamma x'_B v$, where θ is the angle from the x axis to x'. Then $t_B = a+b = \gamma v x'_B + \gamma t'_B = \gamma(t'_B + v x'_B)$. Similarly, by geometrical reasoning, the length from x of the point x' until x of B is $b\tan\theta$, hence $x_B = \gamma x'_B + bv = \gamma x'_B + \gamma t'_B v = \gamma(x'_B + v t'_B)$.

By solving the system of equations for x'_B and t'_B in terms of x_B and t_B we easily obtain the Lorentz transformations

$$\begin{aligned} t' &= \gamma(t - vx) \\ x' &= \gamma(x - vt) \end{aligned}. \tag{8.1}$$

§ Lorentz transformations – algebraic derivation. Consider an event B. The spacetime interval from the origin to B must be the same in frames S and S' by the frame invariace; hence we have

$$c^2t^2 - x^2 = s^2 = s'^2 = c^2t'^2 - x'^2. \tag{8.2}$$

Claim. The solution to the equation above is

$$\begin{aligned} x &= x'\cosh\psi + ct'\sinh\psi \\ ct &= x'\sinh\psi + ct'\cosh\psi \end{aligned}, \tag{8.3}$$

where ψ may be interpreted as the angle by which the S' frame is rotated with respect to S'.

Proof. We prove the claim by checking that the solution Eq. (8.3) satisfies Eq. (8.2) and then argue its uniqueness. Recall the Pythagorean theorem for hyperbolae

$$\cosh^2\psi - \sinh^2\psi = 1.$$

Computing

$$\begin{aligned} c^2t^2 - x^2 &= (x'\sinh\psi + ct'\cosh\psi)^2 - (x'\cosh\psi + ct'\sinh\psi)^2 \\ &= x'^2(\sinh^2\psi - \cosh^2\psi) + c^2t'^2(\cosh^2\psi - \sinh^2\psi) \\ &= c^2t'^2 - x'^2, \end{aligned}$$

we see that the correct form is obtained. Now, an equation of two variables may at most have one solution with a single free parameter (by algebraic considerations); we see that Eq. (8.3) satisfies this scenario, and thus is the unique solution. □

The following hyperbolic identities are useful (proofs are left to the reader):

$$\sinh\psi = \frac{\tanh\psi}{\sqrt{1-\tanh^2\psi}}, \qquad \cosh\psi = \frac{1}{\sqrt{1-\tanh^2\psi}}.$$

We divide the two equations, obtaining $v/c = x/ct = \tanh\psi$, from which

$$\sinh\psi = \frac{v/c}{\sqrt{1-(v/c)^2}} = \gamma\frac{v}{c}, \qquad \cosh\psi = \frac{1}{\sqrt{1-(v/c)2}} = \gamma,$$

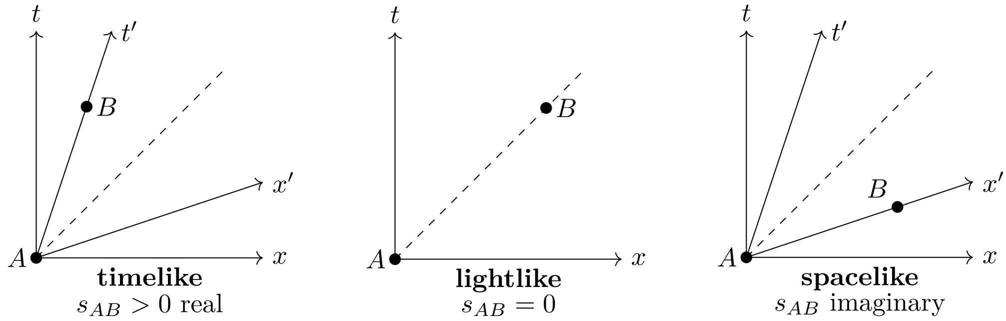
and finally, by plugging back into Eq. (8.3), we obtain the inverse Lorentz transformations $x = x'\gamma + ct'\gamma v/c = \gamma(x' + vt')$ and $ct = x'\gamma v/c + ct'\gamma = c\gamma(t' + vx'/c^2)$ (note that in SR units the c^2 term to equalise dimension would be equal to 1. And we are done.[18]

[18] We have thus reproduced the Landau–Lifshitz derivation.

§9. Speed of light and the Doppler effect

§ Types of intervals. Recall that in §2 the (second) postulate of relativity given by Proposition 2.1 emerged from the observation that 'there exists a maximum velocity of propagation'; the proof (from the principle) rested on the fact that the existence of causality is determined entirely by the maximum velocity of propagation – given a 'source of signal' point at some time and a particle at some other point in spacetime, the signal may affect the particle (at that point) if and only if the signal may reach the point with the maximum velocity of propagation; if this is impossible, there may be no causal relation between the events.

We differentiate between three types of intervals:



It is evident that in the S' frame of lightlike interval (left above), A and B occur *at the same place*; in the S' frame of a spacelike interval (right) – *at the same time*; whereas, in a lightlike interval (centre), the events are connected by a flash of light – *signal.* Given an event A, we call the set of all events B for which $s_{AB} \geq 0$ (timelike interval including the lightlike boundary) the *light cone* of A; in such a case, B is in the *absolute future* of A.

§ Velocity transformations. Let frame S' move with velocity V with respect to S (Home) and let a particle have velocity $\mathbf{v} = [v_x, v_y, v_z]$ in the S frame, and similarly $\mathbf{v}' = [v'_x, v'_y, v'_z]$ in S'. From the inverse Lorentz transformations (§8), it follows that the infinitesimal change of position is given by

$$dx = \gamma(dx' + V dt'), \qquad dy = dy', \qquad dz = dz', \qquad dt = \gamma(dt' + (V/c^2)dx').$$

By dividing the spatial change by the change in time, we obtain the velocity components $v_i = dx_i/dt$ (with $i = x, y, z$); by additionally dividing the numerator and denominator by dt', we express v_i in terms of v'_i as follows

$$\begin{aligned}
v_x &= \frac{dx}{dt} = \frac{\gamma(dx' + V dt')}{\gamma(dt' + (V/c^2)dx')} = \frac{\frac{dx'}{dt'} + V\frac{dt'}{dt'}}{\frac{dt'}{dt'} + (V/c^2)\frac{dx'}{dt'}} = \frac{v'_x + V}{1 + (V/c^2)v'_x}, \\
v_y &= \frac{dy}{dt} = \frac{dy'}{\gamma(dt' + (V/c^2)dx')} = \frac{1}{\gamma}\frac{\frac{dy'}{dt'}}{\frac{dt'}{dt'} + (V/c^2)\frac{dx'}{dt'}} = \frac{1}{\gamma}\frac{v'_y}{1 + (V/c^2)v'_x}, \\
v_z &= \frac{dz}{dt} = \frac{dz'}{\gamma(dt' + (V/c^2)dx')} = \frac{1}{\gamma}\frac{\frac{dz'}{dt'}}{\frac{dt'}{dt'} + (V/c^2)\frac{dx'}{dt'}} = \frac{1}{\gamma}\frac{v'_z}{1 + (V/c^2)v'_x},
\end{aligned}$$

thus obtaining the *inverse Einstein velocity transformations.* We may solve the system of three equations for the three variables v'_x, v'_y, v'_z (left to the reader) to obtain the direct transformations

$$v'_x = \frac{v_x - V}{1 - (V/c^2)v_x}, \qquad v'_y = \frac{1}{\gamma}\frac{v_y}{1 - (V/c^2)v_x}, \qquad v'_z = \frac{1}{\gamma}\frac{v_z}{1 - (V/c^2)v_x}. \tag{9.1}$$

It is important to note that Eq. (9.1) imply that the velocity component v_x parallel to the moving frame affects the other velocity components.

The reader may be interested in solving the following problems.[19]

9.1 Problem. Choose S such that $v_z = 0$, and describe the velocity components as $v_x = v\cos\theta$ and $v_y = v\sin\theta$ with θ an angle in the xy-plane; similarly, let $v'_x = v'\cos\vartheta'$ and $v'_y = v'\sin\vartheta'$ in the S' frame (as before). Show that $\tan\theta$ is given by

$$\tan\theta = \frac{1}{\gamma}\frac{v'\sin\vartheta'}{v'\cos\vartheta' + V}. \tag{9.2}$$

9.2 Problem. (a) Show that in the case of light (namely, $v = v' = c$), Eq. (9.2) becomes

$$\tan\theta = \frac{1}{\gamma}\frac{\sin\vartheta'}{(V/c) + \cos\vartheta'}.$$

(b) Show that the corresponding expressions for $\sin\theta$ and $\cos\theta$ are given by

$$\sin\theta = \frac{1}{\gamma}\frac{\sin\vartheta'}{1 + (V/c)\cos\theta'}, \qquad \cos\theta = \frac{(V/c) + \cos\vartheta'}{1 + (V/c)\cos\vartheta'}.$$

(c) Show that in the low speed limit $V \ll c$, the aberration angle $\Delta\theta = \vartheta' - \theta$ is given by $\Delta\theta = (V/c)\sin\vartheta'$.

§ Relativistic Doppler shift. Let an observer be located at $x = 0$ in its own frame S and a light source be some distance away travelling with $v = v_x$ in the S frame. Denote the period of the light wave by $d\tau_e$, where e denotes emission; since this is a proper time in the S frame, we obtain the emission period in coordinate time to be $d\tau_e = \sqrt{1-v^2}dt_e = \sqrt{1-v_x^2}dt_e$, or, $dt_e = d\tau_e/\sqrt{1-v_x^2}$. Now, during a period, the source moves a distance of $v_x dt_e$, hence the next wave crest will be emitted a time of $(v_x/c)dt_e$ later (for simplicity, we now work in SR units $c = 1$), and hence the observer will observe the time difference between crests to be (r – receive) $dt_r = dt_e + v_x dt_e$, which we express in terms of $d\tau_e$ as follows

$$dt_r = \frac{1+v_x}{\sqrt{1-v_x^2}}d\tau_e = \frac{\sqrt{1+v_x}\sqrt{1+v_x}}{\sqrt{1+v_x}\sqrt{1-v_x}}d\tau_e = \sqrt{\frac{1+v_x}{1-v_x}}d\tau_e.$$

Now, the wavelength λ is related to the period T by the speed of light by $\lambda = cT$; the emitted wavelength is thus $\lambda_e = cd\tau_e$ and the received wavelength $\lambda_r = cdt_r$, hence we obtain

$$\frac{\lambda_r}{\lambda_e} = \sqrt{\frac{1+v_x}{1-v_x}} \tag{9.3}$$

the *relativistic Doppler shift formula.*

[19] The solutions may be found in pp. 13–14 of The course, vol. 2, § 5.

§10. Guestimation

Fermi's approach is perhaps the one trick that separates physicists from mathematicians; all we want is a number, without much regard for relationships, asymptotics, or bifurcations and stability. Such estimates are most useful when checking whether the answer from some long and complicated computation results in the expected order of magnitude – if it is far off, it is highly likely that it contains a mistake. The following remarks elaborate on why the process works.

§ ERROR ESTIMATES. Given variables $X_1, X_2, ..., X_n$ with some corresponding errors $\sigma_{X_1}, ..., \sigma_{X_n}$, upon multiplying them the error is given by

$$\frac{\sigma_{X_1 \cdots X_n}}{X_1 \cdots X_n} = \sqrt{\sum_{i=1}^{n} \left[\frac{\sigma_{X_i}}{X_i}\right]^2}.$$

Now, if all relative errors σ_{X_i}/X_i are roughly similar, then the sum will be approximately $\sqrt{n}(\sigma_{X_1}/X_1)$, hence the relative error of a Fermi guestimation on n terms scales with $\sqrt{n}$ of the error of one variable.

In the case of $n = 9$ (already a long guestimation!), where we may estimate each variable up to an order of 2 (either half or twice as large as the real value), the relative error of the guestimation will be approximately $\sqrt{9} \cdot 2 = 6 < 10$, hence we expect to obtain our answer within a single order of magnitude.

The rest of guestimation consists entirely of examples, which the reader is expected to study.

§11. Four-momentum

§ FOUR-VECTORS. A trajectory in spacetime is simply a curve with a linearly increasing time coordinate $R(t) = (ct, x(t), y(t), z(t)) = (ct, \mathbf{r}(t))$, where $x, y, z : \mathbb{R} \to \mathbb{R}$ are continuous functions; we shall use capital letters to denote 4-vectors and letters in bold for 3-vectors. With these vectors we may define an operation, which resembles (but does not satisfy the definition of!) a norm of a vector; we do this to resemble the spacetime interval, namely, $|R(t)| = \sqrt{c^2t^2 - x^2(t) - y^2(t) - z^2(t)}$. Similarly, given any two such trajectories $R_1(t)$ and $R_2(t)$, we may define an operation resembling an inner product:

$$R_1(t) \cdot_{\text{in}} R_2(t) = c^2t^2 - x_1(t)x_2(t) - y_1(t)y_2(t) - z_1(t)z_2(t).$$

It is evident, that the norm analogue is simply the square root of the inner product of $\mathbf{r}(t)$ with itself.

§ FOUR-MOMENTUM. We obtain the 4-velocity U by taking the proper time derivative of the radius 4-vector $R(t) = (ct, \mathbf{r}(t))$; the intuition is that we want to represent the velocity in spacetime, which does not care about the frame – the meaningful quantities must be frame-independent. Thus $U = dR/d\tau$; now, by the definition of proper time, it is simply a constant multiple of the coordinate time in some fixed frame in which we measure t and $\mathbf{r}(t)$, hence $d\tau = \gamma^{-1}dt$, from which

$$U(t) = \frac{dR(t)}{d\tau} = \gamma\frac{d}{dt}(ct, \mathbf{r}(t)) = \gamma(c, \mathbf{v}(t)),$$

where $\mathbf{v}(t)$ is the 3-velocity in the frame. Now, the 4-momentum P is defined in the usual way of being the 4-velocity multiplied by the *rest mass* (mass of the object in a frame with the object at rest) m_0 of the object, namely, we obtain $P(t) = m_0U(t) = \gamma m_0(c, \mathbf{v}(t))$, where we define the first component to be the relativistic energy over a factor of the speed of light to make dimensions consistent,[20] namely, $\gamma m_0 c = E/c$, and the spatial component becomes the relativistic 3-momentum $\mathbf{p} = \gamma m_0 \mathbf{v}$, hence $P = (E/c, \mathbf{p})$ is the 4-momentum of a particle.

The magnitude of 4-momentum is computed to be

$$|P| = \gamma m_0\sqrt{c^2 - |\mathbf{v}|^2} = \gamma m_0 c\sqrt{1 - \frac{|\mathbf{v}|^2}{c^2}} = m_0 c,$$

which we call the mass of the system, while at the same time $|P|^2 = E^2/c^2 - |\mathbf{p}|^2$, thus obtaining the relation between the rest mass of the particle and its energy and momentum to be

$$m_0^2c^4 = E^2 - c^2|\mathbf{p}|^2. \tag{11.1}$$

It is clear that in the case of a particle at rest, the only energy component is the rest mass energy $E = m_0c^2$ – Einstein's famous equation.

In case of multiple particles, it is evident that the total mass of the system is in fact *not* the sum of the masses; as an example consider two balls of rest mass m moving towards one another, combining inelastically into a mass M. In the frame of the final ball M, the masses move towards $x = 0$ with the same speed v in opposite directions, thus the mass of the moving system is $Mc = |P_1 + P_2| = |\gamma m(c, -\mathbf{v}) + \gamma m(c, \mathbf{v})| = 2\gamma mc > 2mc$, thus the mass of the ball after collision becomes bigger than the sum of the rest masses of the two balls $M = 2\gamma mc > 2mc$.

§ LIGHT. We know that the mass of a photon is zero, therefore from Eq. (11.1), we obtain that $E = c|\mathbf{p}|$, namely, all energy of a photon is contained in its momentum. The dawn of

[20] Units of energy in SI units are $[\mathrm{N\,m}] = [\mathrm{m^2\,kg\,s^{-2}}] = [\mathrm{kg} \cdot (\mathrm{m\,s^{-1}})^2]$, thus $E = mc^2$ with $m = \gamma m_0$.

quantum mechanics was characterised by the brilliant French physicist Louis de Broglie[21] representing a particle by a circular standing wave, from which the relation between momentum and wavelength was obtained: $\lambda = h/p$, with the proportionality constant h being Planck's constant. Therefore, for a photon $E = cp = ch/\lambda = h\nu$ with $\nu = c/\lambda$ being the frequency of light, obtaining the formula familiar to the reader from secondary school.

* * *

This completes the special relativity course. The reader will meet the content again late in the second semester of the first year in relativistic electrodynamics (D.J. Griffiths, *Introduction to electrodynamics*). Indeed, most modern theory of physics considers phenomena far beyond the classical case, with absurdly high energies and velocities – such is without doubt the realm of relativity, and the theory must respect the postulates in Propositions 2.1 and 2.2, from which the entire theory of relativity was constructed.

I hope you enjoyed reading the notes and were able to follow the derivations; any suggestions, comments, and fixes of typos will be greatly appreciated – you may send them to me by email (my email address may be found at the bottom of the page). And with this sentimentality, of course, *good luck on the exam!*

– MGĶ.

[21]In fact, Lord (*7e hertog*) de Broglie. Awarded the Nobel prize in 1929.