Special Relativity and Nuclear Physics

(and some Particle Physics)

L3

Timo Fleig
Département de Physique
Université Paul Sabatier Toulouse

April 25, 2022

Contents

0 Introduction **7**

0.1 Generalities 7

0.1.1 Special relativity in high-energy physics: Production of the $top(t)$ quark

antimatter; mass-energy conversion 10

0.1.2 Nuclear Medicine and Special Relativity: PET Treatment

antimatter; radioactivity; mass-energy conversion 11

0.1.3 Atomic Matter and Special Relativity: Lead-Acid Battery

fundamental properties (spin); relativistic mass; magnetic interactions 13

0.1.4 Search for New Physics: Fermion Electric Dipole Moment

fundamental properties; antimatter; length contraction; magnetic interactions 15

0.1.5 Fundamental Physical Theory: Weak-Interaction Lagrangian

Nuclear physics; covariant formalism; quantum-field theory 16

0.2 Galilei Invariance 17

1 Special Theory of Relativity **23**

1.1 The Lorentz Transformation 23

1.1.1 Deduction from Axioms 23

1.1.2 Lorentz Factor γ 35
1.1.3 Addition Theorem for Velocities 37
1.1.4 Properties of the Lorentz Transformation 39
1.1.5 Minkowski Metric 42
1.1.6 Lorentz transformation in Terms of Rapidity . . 44
1.2 Four-Vectors in SpaceTime 46
1.2.1 Inverse Lorentz Transformation 46
1.2.2 Four-Vectors (Co- and Contravariant) 46
1.2.3 SpaceTime Diagrams 51
1.2.4 Space-, Light-, and Time-Like Four-Vectors . . 52
1.3 Relativistic Mechanics 55
1.3.1 Proper Time 55
1.3.2 Four-Velocity and Four-Acceleration 57
1.3.3 Relativistic Version of Newton's Equation of Motion 61
1.4 Relativistic Formulation of Classical Electrodynamics . 65
1.4.1 A Digression on Units 65
1.4.2 Continuity Equation 66
1.4.3 Maxwell's Equations 68
1.5 Relativistic Mass and Linear Momentum 71
1.5.1 Relativistic Mass 71
1.5.2 Relativistic Linear Momentum 72
1.6 Relativistic Energy 73
1.6.1 Relativistic Energy-Momentum Relation 76
1.6.2 Energy-Mass Equivalence: Mass Defect 80
1.7 Relativistic Kinematics of Particle Interactions 84
1.7.1 Non-relativistic collision processes 84
1.7.2 Relativistic collision processes 85
1.7.3 Spontaneous Two-body Decay 89
1.7.4 Pion Decay and Special Methods 92

2 (A Brief) Introduction to Elementary Particles 97

2.1 Standard Model Phenomenology 97

2.1.1 Historical Notes 97

2.1.2 Mesons 98

2.1.3 Antimatter - Dirac Equation 100

2.1.4 Neutrinos 112

2.1.5 Flavor 116

3 Introduction to Nuclear Physics 125

3.1 General Definitions 125

3.2 Strong Isospin 127

3.3 Radioactive Decay 133

3.3.1 Decay Types 133

3.3.2 Half Life 136

3.4 Nuclear Structure - Nuclear Shell Model 137

Preface

These lecture notes evolved between the years 2017 and 2022 during the course I was giving to third-year physics students at the University Paul Sabatier. Its spirit is an axiomatic (as opposed to historical) introduction to special relativity, followed by applications in particle and nuclear physics. The Covid crisis in the last two years of the lecturing period – where many courses and exercise sessions had to be given online – helped substantially in completing the LaTeX manuscript which is now fully available.

Some comments on recommended literature:

- **W. Rindler**: Relativity.
  Covers special and general relativity; good and established introductory text.
- **Jackson**: Classical Electrodynamics.
  Lorentz-covariant formulation of electrodynamics
- **D. Griffiths**: Introduction to Elementary Particles.
  Pedagogically brilliant introduction to relativistic kinetics and theoretical particle physics

Chapter 0

Introduction

0.1 Generalities

Welcome to this L3 course on "Relativity and Nuclear Physics". Let us begin with some general outline and positioning of the matters in the complete framework of modern physics.

The theory of relativity is one of the centrals pillars of modern physics (the other being quantum mechanics). The theory of **Special Relativity** was developed first (1905), and it reconciles Newton's laws of motion with electrodynamics. Newton's laws of motion are invariant under Galilean transformations, whereas classical electrodynamics was known not to be invariant under such transformations[1]. Newtonian physics can thus be rewritten in the framework of special relativity which is a first goal of this course. After this modification, Newton's laws will not be the same anymore, although they will retain an **equivalent structure**. To the contrary, Maxwell's electrodynamics will remain unchanged, but the equations will be written in a new language that is adapted to the principles of special relativity.

Later in the history of physics (1926), even quantum mechanics underwent a first round of unification with special relativity in the form of the **Dirac equation**. The course will cover this development in

[1]This fact could be demonstrated, but we will take a different route: We will develop the relativistic Einsteinian – different from the Galilean – transformation and show that Maxwell's equations are invariant under this new transformation.

the second half of the semester. The next steps were taken in the late 1940s and early 1950s when Quantum Electrodynamics and more generally Quantum Field Theory were developed, representing a complete "merger" of special relativity with quantum mechanics. These developments even encompassed two forces beyond the electromagnetic force, the **nuclear "strong" force** and the **"weak" force**. The quantum-field theoretical framework for these three forces is today known as the "Standard Model of Elementary Particles" (SM), developed in the 1960s and 1970s. It was completed in 2012 with the detection of the evasive SM "Higgs boson" at the CERN laboratories.

The (so far) last of Nature's forces, the gravitational force, was incorporated into the framework of the theory of relativity early on, in the form of **General Relativity** that today generalizes Newton's laws of gravitation. However, it remains until today one of the great unsolved problems in physics to unify General Relativity with quantum mechanics. Thus, general relativity is not a part of the SM of elementary particles. For the large majority of questions in particle physics, this is not a problem because gravity is so many orders of magnitude weaker than the other three forces of Nature. For questions involving quantum length scales and bodies with large masses – such as black holes or the very early universe – a quantum theory of general relativity seems indispensable.

This course, however, is aimed at phenomena in nuclear physics (and particle physics) where gravity is negligible. There is one aspect of general relativity, however, that will be used in the present context: The "language" of co- and contravariant four-tensors that will be developed (as far as required) in the first half of the course. This formalism is frowned upon by many students (and some teachers as well!) for the reason that it adds a level of complexity to an already difficult part of physics. However, modern fundamental physicists use this language everywhere (!), and so it is obligatory for us to learn it. The price of learning comes with immediate advantages, too, since it massively simplifies the understanding of Lorentz covariance.

A few examples from different fields of physics shall illustrate the importance of special relativity. They highlight the following relativistic phenomena and central aspects of this course:

- Mass $\leftrightarrow$ energy conversion (t production)
- Existence of antimatter (t production, PET treatment, Fermion EDM)
- Relativistic effects in bound matter (lead battery, Fermion EDM)
- Nuclear decays / radioactivity (PET treatment)
- Length contraction (Fermion EDM)

0.1.1 Special relativity in high-energy physics: Production of the *top*(t) quark

antimatter; mass-energy conversion

In 1995 the production of the so far heaviest elementary particle, the top (t) quark, succeeded at the Tevatron collider at Fermilab just outside Chicago. It was produced as and decays as shown in the image:

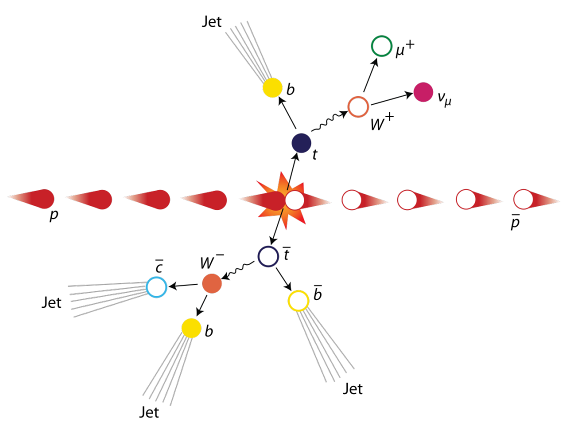


Figure 1:

A high-energy collision of a proton (p) and an antiproton ($\overline{p}$) produces a $t\bar{t}$ pair *via* the strong interaction. These have a lifetime of $\approx 10^{-25}$ [s] and rapidly decay into a bottom (b) quark and a W^+ vector boson (the $\bar{t}$ decays into an antibottom ($\bar{b}$) and a W^- vector boson). These in turn lead to further decays, the debris of which is detected at the facilities. Note that the particle electric charges are as follows (in elementary charges): $C(t) = +2/3$, $C(b) = -1/3$, $C(W^+) = +1$.

What is remarkable is that the rest mass of the t is $m(t) \approx 172 \left[\frac{\mathrm{GeV}}{\mathrm{c}^2}\right]$ whereas the sum of the rest masses of the proton and the antiproton is only $m(p) + m(\overline{p}) \approx 1.88 \left[\frac{\mathrm{GeV}}{\mathrm{c}^2}\right]$. The rest mass of a t roughly corresponds to the mass of a tungsten (W) atom!

So the mass of the incident particles constitute only about 0.5% of the mass of the produced particles. An explanation for this can only be given by one of the central theorems of special relativity which states that energy (in this case kinetic energy) can be converted into a different form of energy, in this case rest energy, representing rest mass $\times c^2$.

The production of the extremely heavy t quark is just one example of how special relativity acts in high-energy physics. Nowadays, New Physics searches, in particular for SuperSymmetric (SUSY) particles, probe into the energy range of ≈ 1000 [TeV] $= 1$ [PeV].

0.1.2 Nuclear Medicine and Special Relativity: PET Treatment

antimatter; radioactivity; mass-energy conversion

Positron Emission Tomography (PET) is used as a means to detect ("imaging") and treat cancerous cells in the human body. A "tracer" is injected which contains a radionuclide, e.g. $^{18}\mathrm{F}^*$, in a fluorine atom bound in a biologically active molecule, $^{18}\mathrm{F}^*\mathrm{BAM}$. This radioactive nucleus in a nuclear excited state arrives at the targeted position and decays under β^+ (positron) emission:

$$p^* \longrightarrow n + e^+ + \nu_e \tag{1}$$

This is the fundamental process underlying the decay of $^{18}\mathrm{F}^*$. Note that the rest mass of the proton is **smaller** than that of the neutron, but the proton is in an excited state, and it is this excitation energy that can be converted into rest energy! The positron does not survive for long because it will undergo **pair annihilation** with an electron of adjacent matter:

$$e^+ + e^- + X^+ \longrightarrow 2\gamma + X^+ \tag{2}$$

where the nucleus X^+ of the atom providing the bound electron has been included in the reaction. Pair annihilation[2] is a high-energy process with emitted photon energies $\varepsilon(\gamma)$ on the order of ≈ 500 [keV][3]. This energy is to be compared with typical atomic transition energies, on the order of $0.01 \ldots 10$ [eV], depending on the involved degrees of freedom (rovibrational, electronic). The high-energy radiation is then detected by detectors surrounding the patient.

So once again, we are confronted with the existence of antimatter. Its existence, in fact, of the positron, was one of the most spectacular *predictions* of theoretical physics, and it is a natural consequence of

[2]The amplitudes (transition probabilities) for such processes can be calculated with Feynman's formalism in the framework of QED.

[3]Such a photon energy converts via $E = h\nu$ and $\lambda = c/\nu$ into a wavelength of about 2.5 pm, constituting γ rays.

special relativity, becoming manifest in the **Dirac equation**, here written in **Lorentz covariant** form:

$$(-\imath\hbar\,\gamma^{\mu}\,\partial_{\mu} + m_0 c\,\mathbb{1}_4)\;\underline{\Psi}(x) = \underline{0} \tag{3}$$

We will walk through its derivation, its solutions, and its basic interpretation and consequences.

0.1.3 Atomic Matter and Special Relativity: Lead-Acid Battery

fundamental properties (spin); relativistic mass; magnetic interactions

As an example from the physics of ordinary atomic matter, consider the lead-acid battery. This work has been published in Phys. Rev. Lett. **106** (2011) *018301*.

The potential difference of a conventional lead-acid battery for cars is 12 [V]. The electrons produced at the negative pole give rise to an electrical current that launches the starter of the car.

The potential difference is obviously crucial for the functioning of the battery. The authors found that it is a function of the energy of electrons occupying the $6s$ state in lead, $U = U(\varepsilon(6s_{\text{Pb}}))$. This energy, in turn, is different in a non-relativistic world, i.e., if special relativity is "turned off". The main reason for this energy difference is the difference between non-relativistic and **relativistic momentum** of the electron in the rest frame of the nucleus. In perturbation theory the total energy of a one-particle atomic state can be written as

$$\varepsilon = \varepsilon_{\text{n.r.}} + \varepsilon^{(1)}_{\text{MV}} + \varepsilon^{(1)}_{\text{Dar}} \tag{4}$$

The relativistic correction "mass-velocity" and "Darwin" can be obtained from Pauli's approximation to Dirac theory, yielding

$$\begin{aligned} \varepsilon^{(1)}_{\text{MV}} &= -\frac{\alpha^2 e^2}{a_0}\,\frac{Z^4_{\text{eff}}}{8n^4}\left[\frac{4n}{\ell+\frac{1}{2}} - 3\right] \\ \varepsilon^{(1)}_{\text{Dar}} &= \frac{\alpha^2 e^2}{a_0}\,\delta_{0\ell}\,\frac{Z^4_{\text{eff}}}{2n^3} \end{aligned}$$

for single-particle states in an atom[4]. These relativistic corrections

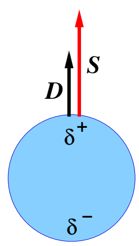
[4] Z_{eff} is the effective nuclear charge in the state, α is Sommerfeld's fine-structure constant, e is the elementary charge, a_0 is the Bohr radius, n is the state's principal quantum number, and ℓ is the orbital angular-momentum quantum number.

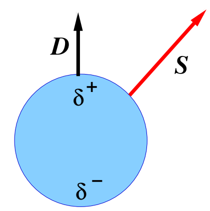
depend on the environment and are different in Pb and PbO_2. The authors were able to model the car battery without these relativistic effects, and the result was $U_{\text{non-rel}} \approx 10$ [V]. Such a tension is insufficient for launching the car's starter! So whenever you start your car, remember that this works because of special relativity.

0.1.4 Search for New Physics: Fermion Electric Dipole Moment

fundamental properties; antimatter; length contraction; magnetic interactions

It is well known that fermions have magnetic dipole moments since this is proportional to the spin of the fermion (its intrinsic angular momentum). First of all, particle spin is a relativistic phenomenon. We will see this in the context of the Dirac equation. But can fermions have an **electric** dipole moment (EDM)? If this were the case, the EDM vector could relate to the spin vector (operator) in two principal ways:

Why 

? and not 

?

We will not go into why the second scenario is impossible (it leads to an internal contradiction in Fermi-Dirac statistics). So what is the Hamiltonian for such a fundamental electric dipole in an external E field?

Classical electromagnetism:

$\varepsilon_{\text{dip}} = -\mathbf{D} \cdot \mathbf{E}$

Fermion EDM vector operator $\hat{\mathbf{d}} \propto \mathbf{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma} \end{pmatrix}$ and so[5]

$\hat{H}_{\text{EDM}} = -d_f\, \gamma^0\, \mathbf{\Sigma} \cdot \mathbf{E}$

The proportionality constant d_f is the fermion EDM.

Dirac matrix γ^0 ensures that $\left\langle \hat{H} \right\rangle$ is a Lorentz scalar (we will see in a short while what that is.)

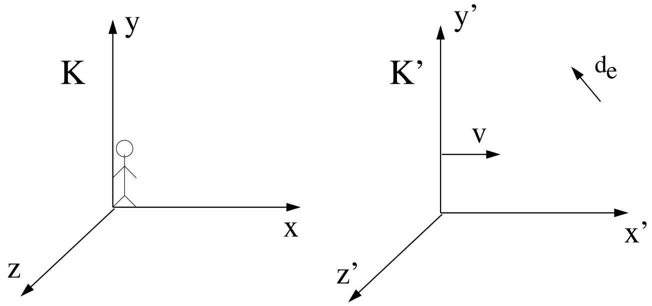
This energy $\left\langle \hat{H} \right\rangle$ is a $\mathcal{T}$-odd *pseudo*scalar.

The potential energy of a fermion EDM in an electric field ((incl.

[5] E. Salpeter, *Phys Rev* **112** (1958) *1642*

E_{ext}) for a state $\psi^{(0)}$ is thus the expectation value
$\varepsilon_{\text{EDM}} = \left\langle -d_e \gamma^0 \mathbf{\Sigma} \cdot \mathbf{E} \right\rangle_{\psi^{(0)}}$

Interpretation:[6]



Length contraction for collinear movement:
$\mathbf{d}_e(K) = \frac{\mathbf{d}_e(K')}{\gamma} = \mathbf{d}_e(K') \left(1 - \frac{\gamma}{1+\gamma} \frac{v^2}{c^2}\right)$
... and for general movement:
$\mathbf{d}_e(K) = \mathbf{d}_e(K') - \frac{\gamma}{1+\gamma} \frac{\mathbf{v}}{c} \left(\mathbf{d}_e(K') \cdot \frac{\mathbf{v}}{c}\right)$

The dipole energy in K then is
$\varepsilon_{\text{dip}} = -\mathbf{d}_e(K) \cdot \mathbf{E} = -\mathbf{d}_e(K') \cdot \left[\mathbf{E} - \frac{\gamma}{1+\gamma} \frac{\mathbf{v}}{c} \left(\frac{\mathbf{v}}{c} \cdot \mathbf{E}\right)\right]$
For small relative velocities we can approximate:
$\varepsilon_{\text{dip}} \approx -\mathbf{d}_e(K') \cdot \mathbf{E} + \frac{1}{2m_e^2 c^2} \mathbf{d}_e(K') \cdot \mathbf{p} \left(\mathbf{p} \cdot \mathbf{E}\right)$

The QM expression can be approximated similarly, yielding

$\varepsilon_{\text{EDM}} \approx -d_e \left\{ \langle \boldsymbol{\sigma} \cdot \mathbf{E} \rangle_{\Psi^L} - \frac{1}{4m^2c^2} \left[\langle \hat{\mathbf{p}} \cdot \mathbf{E}\, \boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \rangle_{\Psi^L} + \langle \mathbf{E} \cdot \hat{\mathbf{p}}\, \boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \rangle_{\Psi^L} \right] \right\}$

which corresponds to the **classical dipole energy in the observer frame**.

0.1.5 Fundamental Physical Theory: Weak-Interaction Lagrangian

Nuclear physics; covariant formalism; quantum-field theory

[6]E.D. Commins, J.D. Jackson, D.P. DeMille, *Am J Phys* **75** (2007) *532*

0.2 Galilei Invariance

At the end of the 19th century classical mechanics was a theory, the equations of motion of which were invariant to so-called **Galilei transformations**. These include rotations in three-dimensional coordinate (real) space and “boosts”. The ensemble of such transformations forms the **group of Galilei transformations**.

Let us review Newton’s axioms:

1. There exists an **inertial frame**[7] **(or system) of reference** in which the forceless motion of a particle is described by a constant velocity,
$$\mathbf{v} = \mathbf{const.} \tag{5}$$
2. In inertial frames the motion of a particle under the influence of a force is described by the **equation of motion**
$$m\mathbf{a} = \sum_i \mathbf{F}_i = \dot{\mathbf{p}} \tag{6}$$
3. For every force with which a particle (1) acts on another particle (2) there is an equal and opposite reaction force with which particle (2) acts on particle (1):
$$\mathbf{F}_{1\to 2} = -\mathbf{F}_{2\to 1} \tag{7}$$

As an illustration of the mentioned invariance, we will explicitly transform the second axiom under boost:

A coordinate transformation between two inertial frames which move at a constant velocity relative to one another is called a **boost**. We will test the invariance of the fundamental law of dynamics for one of

[7] This course will be restricted to the treatment of inertial frames. Nevertheless, in Euclidean (flat) spacetime, it is possible to treat the non-inertial frame’s acceleration as an acceleration seen in an inertial frame. It is thus possible to solve problems involving accelerated motion in the framework of special relativity.

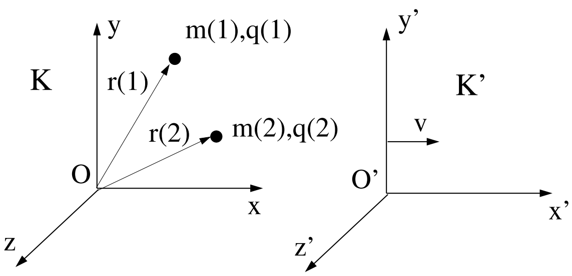


Figure 2:

Two interacting particles and two inertial frames related by a boost transformation. It is supposed that at $t = t' = 0$ we have $x = x' = 0$, *i.e.*, the origins of K and K' coincide.

the particles.

$$\begin{aligned} \mathbf{F}_{2\to 1} &= m(1)\mathbf{a}(1) \\ -\boldsymbol{\nabla}(1)\,\kappa\,\varphi_{21} &= m(1)\mathbf{a}(1) \\ -\kappa\sum_{j=1}^{3}\mathbf{e}_j\,\frac{\partial}{\partial x_j(1)}\,\varphi\left(||\mathbf{r}(1)-\mathbf{r}(2)||\right) &= m(1)\frac{d\mathbf{v}(1)}{dt} \end{aligned} \tag{8}$$

where we are considering a fundamental force that can be written as dependent on the negative gradient of some scalar potential (electrostatic, gravitational) that depends only on the distance between the particles, and κ is a constant (q_1 or m_1, respectively)[8].

The individual terms are now subjected to the Galilei transformation from inertial frame K to inertial frame K':

$$\begin{aligned} \mathbf{r}'(n) &= \mathbf{r}(n) - vt\mathbf{e}_x \\ t' &= t \end{aligned} \tag{12}$$

[8]Verification of the second line for the case of electromagnetic interaction: The electric potential at instant t and position of particle 1 due to the presence of particle 2 is

$$\varphi_{21} = \frac{1}{4\pi\varepsilon_0}\frac{q_2}{||\mathbf{r}(1)-\mathbf{r}(2)||}. \tag{9}$$

The gradient of this potential with respect to coordinates of particle 1 then is

$$\boldsymbol{\nabla}(1)\,\varphi_{21} = -\frac{q_2}{4\pi\varepsilon_0}\frac{\mathbf{r}(1)-\mathbf{r}(2)}{||\mathbf{r}(1)-\mathbf{r}(2)||^3}. \tag{10}$$

From this it follows for the electric field at position 1:

$$\mathbf{E}(\mathbf{r}(1)) = -\boldsymbol{\nabla}(1)\,\varphi_{21} = \frac{q_2}{4\pi\varepsilon_0}\frac{\mathbf{r}(1)-\mathbf{r}(2)}{||\mathbf{r}(1)-\mathbf{r}(2)||^3} \tag{11}$$

and so the force on particle 1 is correctly obtained as $\mathbf{F}_{2\to 1} = q_1\mathbf{E}(\mathbf{r}(1))$ and $\kappa = q_1$ in this case.

and we obtain in detail

- $\mathbf{a}'(1) = \frac{d\mathbf{v}'(1)}{dt'} = \frac{d^2\mathbf{r}'(1)}{dt'^2} = \frac{d^2}{dt^2}\left(\mathbf{r}(1) - vt\mathbf{e}_x\right) = \frac{d^2}{dt^2}\mathbf{r}(1) = \mathbf{a}(1)$
  Whereas velocity and momentum depend on the reference frame, acceleration does not.

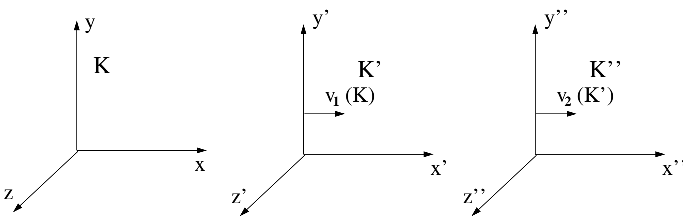
- The potential depends on distance between particles only, so we regard that distance:
  $||\mathbf{r}'(1) - \mathbf{r}'(2)|| = ||\mathbf{r}(1) - vt\mathbf{e}_x - \mathbf{r}(2) + vt\mathbf{e}_x|| = ||\mathbf{r}(1) - \mathbf{r}(2)||$. And so we can conclude that $\varphi' = \varphi$.

- For transforming the gradient we have to respect the functional relationship between the coordinates as due to the transformation, here $\mathbf{r} = \mathbf{r}' + vt\mathbf{e}_x$. Suppose the most general case of a (tensor) field f that is defined in space and time for the frame K, $f = f\big(x(x'), y(x'), z(x'), t(x')\big)$ where all the variables are generally functions of all of the coordinates of K' (x' is sufficient here). Then the total rate of change of f with respect to x' is expressed via the chain rule and we can write
  $\frac{\partial}{\partial x'(1)} = \frac{\partial}{\partial x(1)}\frac{\partial x(1)}{\partial x'(1)} + \frac{\partial}{\partial y(1)}\frac{\partial y(1)}{\partial x'(1)} + \frac{\partial}{\partial z(1)}\frac{\partial z(1)}{\partial x'(1)} + \frac{\partial}{\partial t}\frac{\partial t}{\partial x'(1)} = \frac{\partial}{\partial x(1)}$
  Since for the above Galilei transformation $\frac{\partial x(1)}{\partial x'(1)} = 1$, $\frac{\partial y(1)}{\partial x'(1)} = 0$, same for z, and $\frac{\partial t}{\partial x'(1)} = 0$ since $t = t'$ it follows that
  $\Rightarrow \boldsymbol{\nabla}_x = \boldsymbol{\nabla}_{x'}$.

which proves that $\mathbf{F}'_{2\to 1} = m(1)\mathbf{a}'(1)$ is **equivalent** (form invariant) to the untransformed law of motion. Note that particle mass is absolute in Newtonian physics, just like time is. Newtonian mechanics is said to be **Galilei (boost) invariant**.

Boosts are not the only conceivable kind of transformations between reference frames. In the case frame K' being rotated by an angle α about any particular axis relative to K, this angle plays the role of the constant velocity in the above demonstration. Newton's law is, therefore, also invariant under reference frame rotations.

From the Galilei transformation we can deduce the relative velocity between two inertial frames that move relatively to a third inertial frame in a known way. Fig. (1.1.3.1) shows the setup of the thought experiment.



Inertial frames K, K' and K" with axes aligned. Origins coincide at $t = t' = t'' = 0$. For instance, an observer may be standing on a platform at rest in K, a train (with v_1) carries a gunner who shoots off a bullet (with v_2).

Supposing that the relative velocities v_1 and v_2 are known, we wish to deduce the relative velocity v_3 between frames K and K". From the Galilei transformation in Eq. (12) using two successive Galilei boosts we have, written in matrix form which we will be useful for the treatment of the Einsteinian case,

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} 1 & -v_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} x - v_1 t \\ t \end{pmatrix} \tag{13}$$

$$\begin{pmatrix} x'' \\ t'' \end{pmatrix} = \begin{pmatrix} 1 & -v_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix} \tag{14}$$

Inserting Eq. (13) into Eq. (14) yields

$$\begin{aligned} \begin{pmatrix} x'' \\ t'' \end{pmatrix} &= \begin{pmatrix} 1 & -v_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -v_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \\ &= \begin{pmatrix} 1 & -v_1 - v_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}. \end{aligned} \tag{15}$$

Preserving the form of transformation means that the relative velocity between frames K and K" has to be $v_3 := v_1 + v_2$. This result is generalized to the **theorem of addition of velocities**:

$$\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2. \tag{16}$$

Therefore, assuming for simplicity that the involved velocities are collinear, a light beam's velocity ($v_2 = c$) emitted from an approaching train with velocity v_1 relative to a stationary observer[9] measured in the reference frame of this observer should give the result

$$v_3 = v_1 + c \tag{17}$$

But experiments[10] with sufficient accuracy carried out in the 20th century agree on the fact that the speed of light in all reference frames is $v_3 = c$, irrespective of the magnitude of v_2, i.e., the state of movement of the light source!

This finding has been elevated to become a **postulate** by A. Einstein in his seminal paper[11] from 1905. It will be invoked as the fourth and last postulate in the axiomatic construction of the special theory of relativity.

[9] Remaining in the earlier picture, imagine that the gunner is now using a laser gun.
[10] A. A. Michelson and E. W. Morley, *Am J Sci* **34** (1887) *333-345*
G. Joos, *Ann Phys* **7** (1930) *385*
[11] A. Einstein, *Ann Phys* **17** (1905) *891-921*

Chapter 1

Special Theory of Relativity

1.1 The Lorentz Transformation

1.1.1 Deduction from Axioms

Be there two *inertial frames* (see section 0.2) K with cartesian coordinates $\{x, y, z\}$ and K' with cartesian coordinates $\{x', y', z'\}$. We wish to establish a transformation

$$f : \{x, y, z, t\} \longrightarrow \{x', y', z', t'\} \tag{1.1}$$

of these coordinates including the respective time coordinate, t and t', according to

$$\begin{aligned} x' &= f_x(x, y, z, t) \\ y' &= f_y(x, y, z, t) \\ z' &= f_z(x, y, z, t) \\ t' &= f_t(x, y, z, t) \end{aligned} \tag{1.2}$$

and the inverse transformation. For simplification, we suppose that at $t = t' = 0$: $O = O'$, *i.e.*, the spatial origins of K and K' coincide at $t = t' = 0$. The two frames will be allowed to move at constant velocity v relative to each other. This situation is depicted in Fig. (1.1).

We will first establish the linearity of the transformation f.

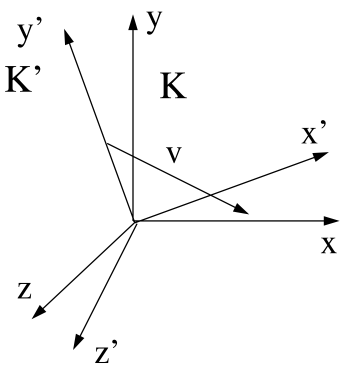


Figure 1.1:

The inertial frames at $t = t' = 0$. The origins coincide but the axes may be rotated with respect to each other. The relative velocity takes a general direction.

1.1.1.1 Homogeneity and Isotropy of Space and Time

Postulate 1: Without external influence, no point in space or time is distinguished from any other (**homogeneity of space and time**).

Postulate 2: Without external influence, no spatial direction is distinguished from any other (**isotropy of space**).

Lemma: The transformation f is linear, so $f : x \longrightarrow \eta x' + \text{const.}$

Idea of proof: Suppose the simplest form of non-linear transformation would hold[1]: $f : x \longrightarrow \zeta x'^2 + \eta x' + \text{const.}$. Then it would follow that $\zeta x'^2 + \eta x' + \text{const.} = \zeta(x' + x'_0)^2 + \text{const.'}$ by quadratic extension[2] . However, this means that $x' = -x'_0$ would be an extremum and, therefore, distinguished from other points x' (contradiction with postulate 1). The proof can be carried on to higher orders in a similar way[3].

Due to postulate 2 the inertial frames can be rotated such that $v = v_x$. So we now depart from Fig. (1.2).

Obviously, $z = 0 \Rightarrow z' = 0$ and $y = 0 \Rightarrow y' = 0$, so we can write a

[1] $\zeta, \eta, \theta \in \mathrm{I\!R}$ are real scalar constants.

[2] In this case the equivalence means that $2\zeta x'_0 = \eta$ and $\zeta x_0'^2 + \text{const.'} = \text{const.}$.

[3] All even-order polynomials have at least one local extremal value. If an odd-order polynomial does not have a local extremal value (like $f(x) = x^3$) then it has at least one point of inflection which again represents a distinguished point.

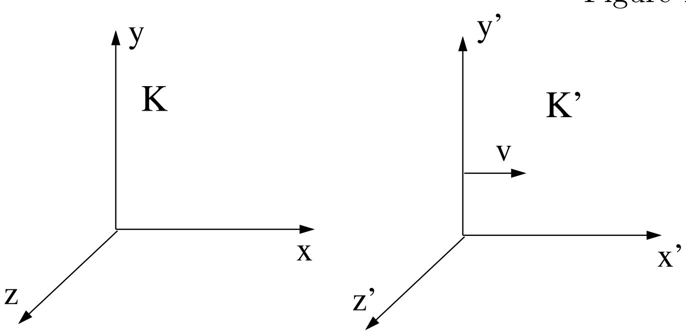


Figure 1.2: Inertial frames K and K' with axes aligned. The relative velocity can be chosen along a single axis.

transformation

$$\begin{aligned} y' &= a(v)\,y \\ z' &= \overline{a}(v)\,z \end{aligned} \tag{1.3}$$

with no constant added. Invoking postulate 2 again, we find $a(v) = \overline{a}(v)$ since the spatial coordinates y and z are on a par with respect to translation in x direction. Therefore,

$$\begin{aligned} y' &= a(v)\,y \\ z' &= a(v)\,z. \end{aligned} \tag{1.4}$$

The function $a(v)$ will be determined subsequently.

For the relation between x' and x we start out from the Galilei transformation and introduce a generalization in form of a function of velocity that is $b(v) = 1$ for the Galilei transformation

$$x' = b(v)\,(x - vt) \tag{1.5}$$

Note also that a and b are in general functions of the velocity v but not of time t since the relative motion between K and K' is constant in time.

The inverse transformation can immediately be written by analogy,

$$x = b'(v')\,(x' + v't') \tag{1.6}$$

where the + sign reflects the increase of x with time, say for the origin of K'. Without loss of generality we here understand v' as the velocity of K relative to K'.

1.1.1.2 Einstein's Principle of Relativity

Homogeneity and isotropy of space and time do not lead any further than what has been presented in section 1.1.1.1. The next task is to establish a general transformation for the time coordinate. This general transformation will become evident in Eq. (1.16).

The first of Einstein's postulates of relativity reads

Postulate 3: No physical measurement can distinguish between reference frames K and K' (**Einstein's first postulate of relativity**).
In other words: "The laws of physics are the same in all inertial frames."

We will now determine general expressions for the functions a and b. We here define the **event** which is understood as a generalization of a point in real space and an instant in time. An event has three spatial and one time coordinate.

- Consider a particle at rest at time t and position z_0 in frame K. This event in K has spatial and time coordinates as

$$\text{E:} \qquad (0, 0, z_0, t) \tag{1.7}$$

  The coordinates of that same event can immediately be written for K':

$$(0, 0, z_0', t')' = (0, 0, a(v)z_0, t')' \tag{1.8}$$

  where Eq. (1.4) has been used. Comparing Eqns. (1.7) and (1.8) it follows that $a(v) = 1$, or else an objective difference between frames K and K' would exist[4], which would contradict postulate 3.

[4]Suppose that $a(v) \neq 1$, then $z_0' \neq z_0$ and the two frames would not be equivalent. Since there is no relative movement in z direction, frame K would have to be stretched (or compressed) relative to K' and the laws of physics would not be the same in both frames.

- Postulate 3 also dictates that $v = v'$, else we would again distinguish between the two frames[5]. So we have established

$$b'(v') = b'(v). \tag{1.9}$$

In order to determine the relationship between b and b' we will now carry out an elementary length measurement of a rigid object in coordinates of both frames. However, before doing so, we need to consider a further difficulty: It will be required to relate the time variables at different positions of a reference frame to each other. In other words, we have to settle the problem of **synchronizing perfect clocks**[6]. This can be achieved through the arrangement in Fig. 1.3.

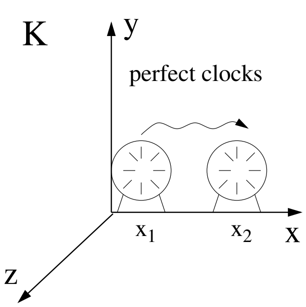


Figure 1.3:

Clocks can be synchronized by installing them in their final position and then using a signal and its duration of travel for synchronizing times on clocks 1 and 2.

It is not possible to first synchronize the clocks and then transport them to the place of measurement. The reason is that we have no guarantee that they will remain synchronized during transport (in fact, they generally don't!). So the clocks are installed in their final destinations in frame K (or K', for that matter), and **then** synchronized. This latter step can be carried out in the following way:

Be an ideal clock (1) at position x_1. A light pulse is emitted at $t(1) = 0$, so time on clock (2) is set at the time of arrival of the

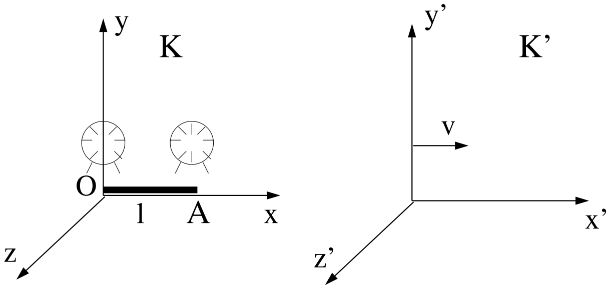
[5] v and v' are taken as positive numbers. So they designate the relative velocity between the two frames.

[6] We are not really doing this in practice. This is a so-called "Gedankenexperiment" (yes, they use a German word in English for this, it means in French "expérience à pensées").

pulse at x_2 to $t(2) := t(1) + \frac{x_2 - x_1}{c} = \frac{x_2 - x_1}{c}$. From then onward, clocks (1) and (2) are synchronized.

Now back to the elementary length measurement.

Figure 1.4:



Length measurement of the same physical object in the rest frame and in the moving frame.

– We first measure the left-hand situation in Fig. (1.4). The left tip and the right tip of the bar in coordinates of K are

$$\begin{aligned} O: &\quad (0, 0, 0, t) \\ A: &\quad (l, 0, 0, t) \end{aligned} \tag{1.10}$$

where the replacement $x = l$ is permissive since the bar rests in frame K. We carry out the same measurement in coordinates of K' **at the same time**[7] (using synchronized clocks), say at $t' = 0$. With the above discussion and Eq. (1.9), Eq. (1.6) becomes

$$\begin{aligned} x &= b'(v)\,(x' + vt') \\ x' &= \frac{x}{b'(v)} - vt' \qquad \text{and with } t' = 0 \text{ we have} \\ x' &= \frac{x}{b'(v)} \end{aligned}$$

So the coordinates of the events in K' can be written as

$$\begin{aligned} O: &\quad (0, 0, 0, 0)' \\ A: &\quad (x', 0, 0, 0)' = (l/b'(v), 0, 0, 0)' \end{aligned} \tag{1.11}$$

[7] This is equivalent to supposing that there exists an inertial frame K' in which the two events occur simultaneously. From the Galilean point of view this is a trivial assumption. We will later see that this assumption is compatible with the general structure of SpaceTime in Einsteinian relativity.

– Carrying out the same type of measurement, but now with the bar at rest in frame K', Fig.

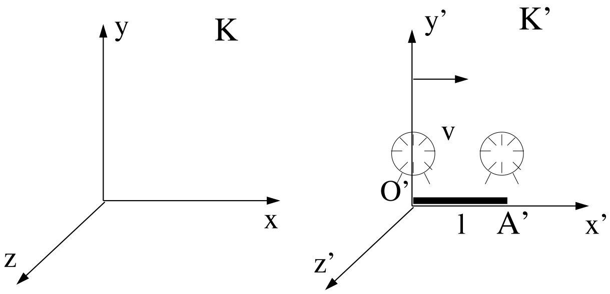


Figure 1.5:

Length measurement of the same physical object in the rest frame and in a frame in relative movement with respect to the rest frame.

(1.5), we measure in K' coordinates

$$\begin{aligned} O' : &\quad (0,0,0,t')' \\ A' : &\quad (l,0,0,t')' \end{aligned} \tag{1.12}$$

Expressed in coordinates of K at the same time, say at $t = 0$, we obtain

$$\begin{aligned} O' : &\quad (0,0,0,0) \\ A' : &\quad (x,0,0,0) = (l/b(v),0,0,0) \end{aligned} \tag{1.13}$$

where now (1.5) has been used for the equality. In detail:

$$\begin{aligned} x' &= b(v)\,(x - vt) \\ x &= \frac{x'}{b(v)} + vt \\ x &= \frac{x'}{b(v)} \end{aligned}$$

Invoking postulate 3 once again, it follows necessarily that

$$b(v) = b'(v) \tag{1.14}$$

since otherwise we would have an objective distinction between frames K and K' (inverting the measurement should not change the measured length of the bar).

The situation is symmetric: The length of the bar fixed in K' appears changed in K, and the bar fixed in K also appears changed in K' in the same manner. This is not surpising, since in both situations an observer has a bar moving away from him.

Having obtained this result, we can rewrite Eqns. (1.5) and (1.6) as

$$\begin{aligned} x' &= b(v)\,(x - vt) \\ x &= b(v)\,(x' + vt'). \end{aligned} \tag{1.15}$$

We can now summarize the findings up to this point. Eliminating x' from the second equation in Eq. (1.15) by inserting the first (with the goal of obtaining t' as a function of the unprimed variables) yields

$$\begin{aligned} x &= b(v)\,\big(b(v)\,(x - vt) + vt'\big) \\ vt' &= \frac{x}{b(v)} - b(v)\,(x - vt) \\ t' &= \frac{x}{vb(v)} + tb(v) - \frac{xb(v)}{v} \end{aligned}$$

and the transformation from coordinates of frame K to K' can be written as

$$\boxed{\begin{aligned} x' &= b(v)\,(x - vt) \\ y' &= y \\ z' &= z \\ t' &= b(v)\,\left[t + \frac{x}{v}\left(\frac{1}{b(v)^2} - 1\right)\right] \end{aligned} \quad (1.16)}$$

It is not possible to determine $b(v)$ without experimental information. Setting $b(v) = 1$ yields the Galilei transformation, which was the accepted form until measurements at very high relative velocity and very high precision became available. Then, deviations from $b(v) = 1$ could be determined[8].

1.1.1.3 The Constancy of the Speed of Light

The result of various measurements has been elevated by Einstein to become a postulate:

Postulate 4: A measurement of the speed of light in any direction in a reference frame always yields a constant value, $c \approx 2.99792458 \times 10^8 \left[\frac{m}{s}\right]^a$ (**Einstein's second postulate of relativity**).

[a]National Institute of Standards (NIST), values of physical constants

Corollary: It is sufficient to formulate Postulate 4 for one reference frame, because it follows from Postulate 3 that $c =$const. in all reference frames, since they have to be equivalent.

We are now in the position to determine the function $b(v)$ unambiguously. It follows from a simple thought experiment:

Fig. (1.6) defines the arrangement. The coordinates of the emission event are

$$\begin{aligned} E &: \ (0,0,0,0) \\ E &: \ (0,0,0,0)' \end{aligned} \tag{1.17}$$

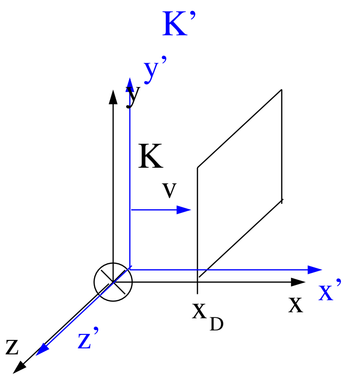
An observer at rest in K measures the detection of the pulse in x at time t. An observer at rest in K' measures the detection at x' at time t'.

Let us first see what non-relativistic transformation laws would tell us. The propagation of the detector in coordinates of K' follows

$$x'_D = x_D - vt \tag{1.18}$$

[8]Fizeau, Michelson-Morley, Joos

Figure 1.6:



A light pulse is emitted from the origin of K at time $t = 0$. At this instant the origins of K and K' coincide, so $t = t' = 0$. The detector is fixed in frame K at position x_D. K' moves with constant velocity v relative to the light source.

for any time instant t. So at the specific instant $t = t_D$ we can write this as

$$x'_D = ct_D - vt_D \tag{1.19}$$

since $x_D = ct_D$. Slightly reformulated we have

$$x'_D = (c - v)t_D \tag{1.20}$$

Since time is absolute (independent of the inertial frame) in Newtonian/Galilean physics, meaning $t_D = t'_D$, this can be written as

$$x'_D = (c - v)t'_D \tag{1.21}$$

The last expression means nothing else than that the light propagates with the velocity $c - v$ seen from the frame K' which is in accord with our classical (non-relativistic) view of physics. However, postulate 4 and the corrolary enforce that the speed of light is also c in frame K'! Therefore, for general positions and instances in time, we write

$$\begin{aligned} x &= ct \\ x' &= ct' \end{aligned} \tag{1.22}$$

Obviously, at this point there is the clear departure from the non-relativistic notion: According to the Galilei transformation we would expect $x' = (c-v)t'$. Eq. (1.22) also tells us that since after some propagation of the pulse $x_D \neq x'_D$ it follows that $t_D \neq t'_D$ which means that the detection event does **not** happen simultaneously in frames K and K'.

From all of this we get

$$\begin{aligned} D: \ (x,0,0,t) &= (ct,0,0,t) \\ D: \ (x',0,0,t')' &= (ct',0,0,t')' \end{aligned} \tag{1.23}$$

and the rest is calculation. Introducing this result into Eq. (1.15) we obtain

$$\begin{aligned} ct' &= b(v)\,(ct - vt) = b(v)\,(c-v)t \\ ct &= b(v)\,(ct' + vt') = b(v)\,(c+v)t'. \end{aligned} \tag{1.24}$$

Multiplication of the two equations eliminates the time coordinates and yields

$$c^2 tt' = b(v)^2\,(c^2 - v^2)tt'$$

$$b(v)^2 = \frac{c^2}{c^2 - v^2} \tag{1.25}$$

$$b(v) = \pm\sqrt{\frac{1}{1 - \frac{v^2}{c^2}}} \tag{1.26}$$

The sign ambiguity is resolved through an additional consideration. Taking $v = 0$ in the first line of Eq. (1.16) gives $x' = b(0)x$. But in this case $x' = x \Rightarrow b(0) = +1$, so we must use the positive sign for consistency. We can write

$$b(v) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \beta^2}} =: \gamma(v) \tag{1.27}$$

with $\beta := \frac{v}{c}$ and γ the **Lorentz factor**. Introduction of the Lorentz factor into the preliminary transformation Eq. (1.16) results in

$$t' = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left[t + \frac{x}{v}\left(1 - \frac{v^2}{c^2} - 1\right)\right]$$

and so

$$\boxed{\begin{aligned} x' &= \frac{x - vt}{\sqrt{1-\frac{v^2}{c^2}}} \\ y' &= y \\ z' &= z \\ t' &= \frac{t - x\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \end{aligned} \qquad (1.28)}$$

as the final form of the Lorentz transformation of coordinates. This type of transformation is called a "**Lorentz boost**" (rotation-free Lorentz transformation).

At this point the verification of the Lemma from subsection 1.1.1.1 is in place. In y and z the Lorentz transformation is trivially linear. The transformations in x and t

$$\begin{aligned} x' &= \gamma x - \gamma v t \\ t' &= \gamma t - \gamma x \frac{v}{c^2} \end{aligned}$$

are linear, too, since γ, v, t are constant in x (first line) and γ, v, x are constant in t (second line)[9]. Therefore, the Lorentz transformation is manifestly a linear transformation.

It is also immediately evident that for relative velocities very small compared to the speed of light, $v << c$, the Galilei transformation

[9] x and t are of course independent coordinates in frame K.

from non-relativistic physics is approximately obtained:

$$\left(\frac{x - vt}{\sqrt{1-\frac{v^2}{c^2}}}\right)_{v<<c} \approx x - vt$$
$$\left(\frac{t - x\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}\right)_{v<<c} \approx t \tag{1.29}$$

However, since in non-relativistic physics no limit for relative velocities exists, this can only be considered an approximation for special cases. In order to obtain a non-relativistic limit of the theory as such, we have to allow for arbitrary relative velocities v.

Formally, we take the speed of light to infinity in expressions involving velocity ratios[10]

$$\lim_{c\longrightarrow\infty} \frac{x - vt}{\sqrt{1-\frac{v^2}{c^2}}} = x - vt$$
$$\lim_{c\longrightarrow\infty} \frac{t - x\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} = t. \tag{1.30}$$

This **defines** the non-relativistic limit of a physical theory.

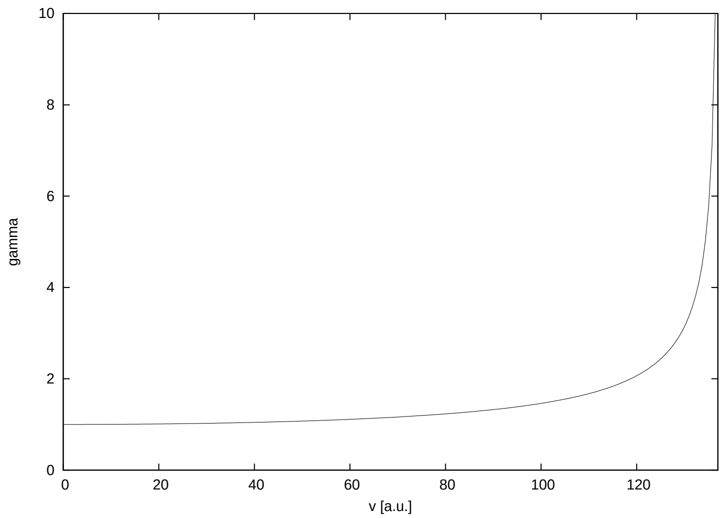
Before continuing, we will here take a look at the important Lorentz factor.

1.1.2 Lorentz Factor γ

Fig. (1.7) displays the function $\gamma(v) = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$.

[10]In electromagnetism, this limit can also be taken, although the dependency on factors $\frac{v}{c}$ is more subtle. For example, $E_m = -\boldsymbol{\mu}\cdot\mathbf{B} \propto \frac{\mathbf{v}\cdot\mathbf{v}'}{c^2}$. The factors $\frac{1}{c}$ become "visible" in the Gaussian unit system, and magnetic moment as well as magnetic field are proportional to velocities of charged particles. This means that magnetism does not exist in the non-relativistic limit of electromagnetism!

Figure 1.7:



Lorentz factor as a function of relative velocity (in atomic units, $c = \frac{1}{\alpha} \approx 137.136$ where α is Sommerfeld's fine-structure constant)

- $\lim\limits_{v \to 0} \gamma(v) = 1$
- $\lim\limits_{v \to c} \gamma(v) = +\infty$
- $\gamma(v > c) \in \mathbb{C}$

It is to be noted that for a relative velocity $v > c$ between inertial frames the Lorentz factor becomes a complex number. From Eq. (1.28) it would then be directly inferred that $x' \in \mathbb{C}$, too. But this would be in contradiction with the basic notion that observables (such as lengths) should be real numbers. We, therefore, take this finding as a first indication that relative velocities greater than the speed of light are not possible.

1.1.2.1 Corollaries

It is interesting to inspect the direct consequences of the Lorentz boost in Eq. (1.28) in terms of two aspects:

1. Time is absolute in non-relativistic physics. However, if one considers a time interval in frame K' (at fixed position x') and transforms the time interval as measured in coordinates of K, then (transformation from K' to K):

$$\Delta t = t_2 - t_1 = \gamma\, t_2' + \gamma\, x' \frac{v}{c^2} - \gamma\, t_1' - \gamma\, x' \frac{v}{c^2} = \gamma\, \Delta t' \qquad (1.31)$$

This means that generally $\Delta t' \neq \Delta t$. In other words, the passage of time depends on the relative motion between inertial frames, and the phenomenon is generally called **time dilation**.

2. We know that length is conserved under Galilei transformations. Under Lorentz boosts, let us consider a spatial interval in K', $\Delta x' = x'_2 - x'_1$, measured at fixed time t' in K'. An observer in K measures

$$\Delta x = x_2 - x_1 = \gamma\, x'_2 + \gamma\, vt' - \gamma\, x'_1 - \gamma\, vt = \gamma\, \Delta x' \tag{1.32}$$

Since for the relative velocity $v \neq 0$ we have $\gamma > 1$, it follows that the observed length (in K') appears shorter than the measured length in K. Fixed lengths in the direction of relative movement are, therefore, **not conserved** and this phenomenon is generally called **length contraction**.

Various situations can be investigated under these two aspects and interpreted in terms of length contraction and time dilation in special relativity, but the details depend crucially on the specific setup.

1.1.3 Addition Theorem for Velocities

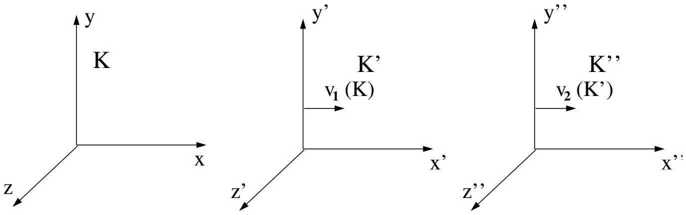
1.1.3.1 Derivation

Newtonian mechanics is invariant under Galilei transformations, and the addition theorem of velocities in the non-relativistic context is well known:

$$\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2 \tag{1.33}$$

We will now deduce the corresponding law in the framework of special relativity based on the Lorentz transformation. Fig. (1.1.3.1) shows the setup of the thought experiment.

We wish to deduce the relative velocity between frames K and K". The starting point is Eq. (1.28) with the Lorentz transformation in its



Inertial frames K, K' and K" with axes aligned. Origins coincide at $t = t' = t'' = 0$.

original form. The two Lorentz boosts can thus be written as

$$\begin{array}{cc} x' = \gamma_1(x - v_1 t) & x'' = \gamma_2(x' - v_2 t') \\ t' = \gamma_1(t - x\frac{v_1}{c^2}) & t'' = \gamma_2(t' - x'\frac{v_2}{c^2}) \\ \mathrm{K} \rightarrow \mathrm{K'} & \mathrm{K'} \rightarrow \mathrm{K''} \end{array} \tag{1.34}$$

where individual Lorentz factors $\gamma_j(v_j) = \frac{1}{\sqrt{1-\frac{v_j^2}{c^2}}}$ have been introduced. Inserting the first transformation into the second corresponds to transforming coordinates from K → K" and results in

$$\begin{aligned} x'' &= \gamma_1\gamma_2\, x\left(1 + \frac{v_1 v_2}{c^2}\right) - \gamma_1\gamma_2\, t(v_1 + v_2) \\ t'' &= \gamma_1\gamma_2\, t\left(1 + \frac{v_1 v_2}{c^2}\right) - \gamma_1\gamma_2\, \frac{x}{c^2}(v_1 + v_2) \end{aligned} \tag{1.35}$$

after convenient regrouping of terms[11]. However, the Lorentz transformation relating K → K" can also be written in a general form:

$$\begin{aligned} x'' &= \gamma_3(x - v_3 t) \\ t'' &= \gamma_3\left(t - x\frac{v_3}{c^2}\right) \end{aligned} \tag{1.36}$$

where v_3 is the relative velocity of these two frames. Comparing coefficients of x and t in Eqs. (1.35) and (1.36) results in (two times) the following conditions:

$$\gamma_3 = \gamma_1\gamma_2\left(1 + \frac{v_1 v_2}{c^2}\right) \tag{1.37}$$

$$\gamma_3\, v_3 = \gamma_1\gamma_2(v_1 + v_2) \tag{1.38}$$

[11] We could also obtain these two equations by calculating the matrix product $\tilde{\mathbf{L}}(v_2)\tilde{\mathbf{L}}(v_1)$ and then acting with the new matrix onto the SpaceTime coordinates $\begin{pmatrix} x \\ t \end{pmatrix}$.

Inserting the first into the second of these conditions directly gives

$$v_3\,\gamma_1\gamma_2\left(1+\frac{v_1v_2}{c^2}\right)=\gamma_1\gamma_2(v_1+v_2)$$

$$\boxed{v_3=\frac{v_1+v_2}{1+\frac{v_1v_2}{c^2}} \qquad (1.39)}$$

the **addition theorem for velocities** in special relativity for relative movement along aligned coordinate axes. Based on it a satisfactory explanation and interpretation of the experiments carried out by Fizeau, Michelson and Morley and others is possible.

1.1.3.2 Corollaries

A number of interesting consequences of the addition theorem for velocities shall be discussed at this point. Considering the properties of the Lorentz factor we have seen that $v_j > c$ is generally unacceptable. So we *assert* that $v_j \leq c$ is always true and investigate its consequences.

- $v_3 \leq c$.
  Proof. $\lim\limits_{v_1,v_2\to c} v_3=\frac{2c}{2}=c$. And also $\lim\limits_{v_2\to c}=\frac{v_1+c}{1+\frac{v_1}{c}}=\frac{c(v_1+c)}{v_1+c}=c$.

- $\lim\limits_{c\to\infty} v_3=v_1+v_2$
  which establishes the addition theorem for velocities in the non-relativistic limit.

1.1.4 Properties of the Lorentz Transformation

The first and foremost question concerns the invariance of physical laws under coordinate transformations. For example, Newton's second law is invariant under spatial rotations. So what does the obtained LT represent?

We may write the one-dimensional Lorentz boost given in Eq. (1.28) also as a matrix equation,

$$\begin{pmatrix} x' \\ y' \\ z' \\ t' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -v\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{v}{c^2}\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \tag{1.40}$$

where the vector space we will henceforth call **SpaceTime** has been introduced. It is the natural generalization of the usual three-dimensional coordinate space $\mathbb{R}^3$ to include the time coordinate, so Minkowski SpaceTime is defined as $\mathbb{R}^4$.

Since the y and z coordinates are left unaffected by the boost, we may for the present case simplify the matrix equation to two dimensions:

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} \gamma & -v\gamma \\ -\frac{v}{c^2}\gamma & \gamma \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} = \tilde{\mathbf{L}}\,\mathbf{x} \tag{1.41}$$

We find that

- $det(\tilde{\mathbf{L}}) = 1$
- $\mathbf{x}'^2 \neq \mathbf{x}^2$; non-conservation of the scalar product, naïvely defined as in usual $\mathbb{R}^2$
- $\tilde{\mathbf{L}}^T\,\tilde{\mathbf{L}} \neq \mathbf{1}$ and so $\tilde{\mathbf{L}}^T \neq \tilde{\mathbf{L}}^{-1}$.

This means that $\tilde{\mathbf{L}}$ is not an orthogonal matrix. So clearly $\tilde{\mathbf{L}}$ does not represent a rotation in Minkowski SpaceTime. Note also that the physical dimensions of the SpaceTime vector $\begin{pmatrix} x \\ t \end{pmatrix}$ are not homogeneous.

In order to obtain an orthogonal transformation in SpaceTime Minkowski introduced a trick: He defined one SpaceTime vector component as *imaginary*, *i.e.* we have the SpaceTime coordinates $\{ict, x\}$ instead of

$\{x, t\}$. Then the Lorentz transformation becomes

$$\begin{pmatrix} x' \\ \imath ct' \end{pmatrix} = \begin{pmatrix} \gamma & \imath\frac{v}{c}\gamma \\ -\imath\frac{v}{c}\gamma & \gamma \end{pmatrix} \begin{pmatrix} x \\ \imath ct \end{pmatrix} = \mathbf{L}\,\mathbf{x} \tag{1.42}$$

The reader can easily verify that this version of the Lorentz boost is identical to the original one above, Eq. (1.41). However, the Space-Time vector components are now physically homogeneous. In addition, the following properties of the transformation matrix with respect to Minkowski coordinates can be shown straightforwardly:

- $det(\mathbf{L}) = 1$
- $\mathbf{x}'^2 = \mathbf{x}^2$; conservation of the scalar product[12]
- $\mathbf{L}^T = \mathbf{L}^{-1}$; orthogonality
- $v \longrightarrow -v \Rightarrow \mathbf{L} \longrightarrow \mathbf{L}^{-1}$; inverse transformation property

The conservation of the scalar product can also be written as

$$\mathbf{x}'^2 \;=\; \mathbf{x}^2 \tag{1.43}$$
$$x'^2 + y'^2 + z'^2 - c^2t'^2 \;=\; x^2 + y^2 + z^2 - c^2t^2 \tag{1.44}$$

by taking all three spatial components. This finding has a direct physical interpretation. If at $t = 0$ a light pulse is emitted from the origin of inertial frame K, then its radial position at time t is

$$r = \sqrt{x^2 + y^2 + z^2} \;=\; ct \tag{1.45}$$
$$\Rightarrow x^2 + y^2 + z^2 - c^2t^2 \;=\; 0. \tag{1.46}$$

In coordinates of K', where the origins coincide at $t = t' = 0$, we infer from the constancy of the speed of light in all frames (Postulate 4):

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0 \tag{1.47}$$

[12] The scalar product in Minkowski space is here (!) defined with the usual Euclidian metric tensor, so here $\mathbb{1}_2$.

Equating the two above expressions reproduces the conservation of the scalar product under Lorentz transformation.

In fact, the Lorentz invariant $x^2+y^2+z^2-c^2t^2$ comprises a so-called **Lorentz scalar**. We will come back to a more general discussion of Lorentz scalars in a later section.

Note that the scalar product can also be written as $x^2 + y^2 + z^2 + (\imath ct)^2$ which comprises the natural form of a scalar product in a four-dimensional space.

1.1.5 Minkowski Metric

The modern standard representation of the Lorentz transformation is different from both the ones we have established in subsection 1.1.4. Let us review the scalar product of the SpaceTime vector $\begin{pmatrix} x' \\ \imath ct' \end{pmatrix}$ with itself (Eq. (1.42)):

$$\begin{aligned}
\begin{pmatrix} x' & \imath ct' \end{pmatrix} \begin{pmatrix} x' \\ \imath ct' \end{pmatrix} &= \begin{pmatrix} x & \imath ct \end{pmatrix} \begin{pmatrix} \gamma & -\imath\frac{v}{c}\gamma \\ \imath\frac{v}{c}\gamma & \gamma \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & \imath\frac{v}{c}\gamma \\ -\imath\frac{v}{c}\gamma & \gamma \end{pmatrix} \begin{pmatrix} x \\ \imath ct \end{pmatrix} \\
&= \begin{pmatrix} x & \imath ct \end{pmatrix} \begin{pmatrix} \gamma^2 - \frac{v^2}{c^2}\gamma^2 & 0 \\ 0 & -\frac{v^2}{c^2}\gamma^2 + \gamma^2 \end{pmatrix} \begin{pmatrix} x \\ \imath ct \end{pmatrix} \\
&= \begin{pmatrix} x & \imath ct \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ \imath ct \end{pmatrix} = x^2 - c^2t^2
\end{aligned} \tag{1.48}$$

where we have used the usual 3-dimensional Euclidean metric for the scalar product, a unit matrix[13].

But it is possible to write the same scalar product for **real-valued coordinate axes** by changing the form of the metric[14].

The Euclidean metric in two-dimensional flat space is defined as

$$\mathbf{g} = \begin{pmatrix} \mathbf{e}_x \cdot \mathbf{e}_x & \mathbf{e}_x \cdot \mathbf{e}_y \\ \mathbf{e}_y \cdot \mathbf{e}_x & \mathbf{e}_y \cdot \mathbf{e}_y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.49}$$

[13] in the present case 2-dimensional.

[14] Likewise, the scalar product over two vectors in a vector space when changing from orthogonal to non-orthogonal axes can be conserved by introducing an accompanying change of the metric.

Now a coordinate transformation is characterized by the corresponding Jacobian matrix $\mathbf{J}$[15]. Suppose we define the coordinate transformation as follows:

$$\begin{aligned} u' &= u \\ v' &= \imath v \end{aligned} \tag{1.53}$$

Then we get the Jacobian of the transformation as

$$\mathbf{J} = \begin{pmatrix} \frac{\partial u'}{\partial u} & \frac{\partial u'}{\partial v} \\ \frac{\partial v'}{\partial u} & \frac{\partial v'}{\partial v} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \imath \end{pmatrix} \tag{1.54}$$

The conservation of the scalar product under coordinate transformation is achieved by transforming the metric, since

$$\begin{pmatrix} u' & v' \end{pmatrix} \mathbf{g} \begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} u & v \end{pmatrix} \mathbf{J}^T \mathbf{g} \, \mathbf{J} \begin{pmatrix} u \\ v \end{pmatrix} \tag{1.55}$$

and so we can define a new metric as

$$\mathbf{g}' = \mathbf{J}^T \mathbf{g} \, \mathbf{J} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.56}$$

for the scalar product. And so we have

$$\begin{pmatrix} u' & v' \end{pmatrix} \mathbf{g} \begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} u & v \end{pmatrix} \mathbf{g}' \begin{pmatrix} u \\ v \end{pmatrix} \tag{1.57}$$

which means that the scalar product is conserved under coordinate transformation if the metric tensor is transformed accordingly.

In special relativity coordinates this means

$$\begin{pmatrix} x & \imath ct \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ \imath ct \end{pmatrix} = \begin{pmatrix} x & ct \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ ct \end{pmatrix}$$

where the new non-unity metric for the scalar product has been introduced. Moreover, as we have seen earlier, for the conservation of the scalar product $x^2 - c^2t^2 = x'^2 - c^2t'^2$ (Eq. (1.44)) the global sign is

[15] in the following way: The total differentials for the set of coordinates can be written as

$$du' = \frac{\partial u'}{\partial u} du + \frac{\partial u'}{\partial v} dv \tag{1.50}$$

$$dv' = \frac{\partial v'}{\partial u} du + \frac{\partial v'}{\partial v} dv \tag{1.51}$$

where $du' = u'_1 - u'_2$ is an infinitesimal interval along u'. $u'_2 = 0$ is just a special case of this. Arranging this in matrix form makes the Jacobian matrix appear:

$$\begin{pmatrix} du' \\ dv' \end{pmatrix} = \begin{pmatrix} \frac{\partial u'}{\partial u} & \frac{\partial u'}{\partial v} \\ \frac{\partial v'}{\partial u} & \frac{\partial v'}{\partial v} \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix} \tag{1.52}$$

The rest follows from there. See texts on metric tensors for more information.

irrelevant; what matters is only the relative sign between spatial and time coordinates. So we are free to choose the metric according to

$$\begin{pmatrix} ct & x \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} = c^2t^2 - x^2 \tag{1.58}$$

where the "minus" sign is on the spatial coordinate instead. This metric has become the modern standard[16]. We will call

$$\mathbf{g} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.59}$$

modern Minkowski metric, for the four dimensions of SpaceTime, and use it from here onward. Moreover, in the new basis of SpaceTime (where the time coordinate also becomes the first coordinate) the Lorentz transformation matrix changes. The reader can easily verify the equivalence of the two representations:

$$\begin{pmatrix} \imath ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\imath\frac{v}{c}\gamma \\ \imath\frac{v}{c}\gamma & \gamma \end{pmatrix} \begin{pmatrix} \imath ct \\ x \end{pmatrix} \quad \text{with complex axes} \tag{1.60}$$

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\frac{v}{c}\gamma \\ -\frac{v}{c}\gamma & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \quad \text{with real axes} \tag{1.61}$$

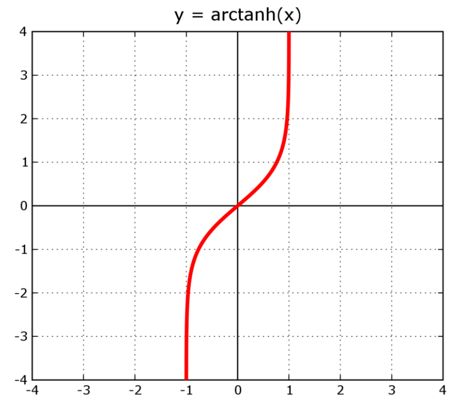
From now on we use the designation $\mathbf{\Lambda}(v) := \begin{pmatrix} \gamma & -\frac{v}{c}\gamma \\ -\frac{v}{c}\gamma & \gamma \end{pmatrix}$ for Lorentz boosts.

1.1.6 Lorentz transformation in Terms of Rapidity

There is a convenient way of expressing the Lorentz boost which is useful in the context of combined Lorentz transformations and the **Lorentz group**. If we define the "rapidity"

[16]It has been the standard for at least 50 years. A landmark text that still uses the old unit metric is Bethe and Salpeter, "Quantum Mechanics of One- and Two-Electron Atoms".

$$\Phi_x := \operatorname{arctanh}\left(\frac{v}{c}\right) \qquad (1.62)$$



then it follows that

$$\begin{aligned} \frac{v}{c} &= \tanh \Phi_x \\ \frac{v^2}{c^2} &= \frac{\cosh^2 \Phi_x - 1}{\cosh^2 \Phi_x} \\ \gamma &= \cosh \Phi_x \\ \frac{v}{c}\gamma &= \sinh \Phi_x \end{aligned} \qquad (1.63)$$

and so we obtain for the modern representation of the Lorentz transformation

$$\begin{pmatrix} \gamma & -\frac{v}{c}\gamma \\ -\frac{v}{c}\gamma & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} \cosh \Phi_x & -\sinh \Phi_x \\ -\sinh \Phi_x & \cosh \Phi_x \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \qquad (1.64)$$

It is then a straightforward exercise to prove the following identity:

$$\begin{aligned} \mathbf{\Lambda}(\Phi_{x_1})\mathbf{\Lambda}(\Phi_{x_2}) &= \mathbf{\Lambda}(\Phi_x) \\ &\text{with} \\ \Phi_{x_1} + \Phi_{x_2} &= \Phi_x \end{aligned} \qquad (1.65)$$

Eq. (1.65) shows that the double boost occurring, e.g., in the derivation of the velocity addition theorem can — using the rapidity parameter — be written conveniently as a new Lorentz transformation where the new rapidity Φ_x simply is the sum of the two original rapidities. This resembles the formal situation for the addition of velocities in Newtonian mechanics using the Galilei transformation.

1.2 Four-Vectors in SpaceTime

1.2.1 Inverse Lorentz Transformation

If we want to represent the Lorentz boost from frame K' into frame K, we need the inverse Lorentz transformation matrix. It can be derived straightforwardly from Eq. (1.61) to be

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} \gamma & \frac{v}{c}\gamma \\ \frac{v}{c}\gamma & \gamma \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix}. \tag{1.66}$$

Therefore, $\mathbf{\Lambda}^{-1}(v) = \mathbf{\Lambda}(-v)$. This result is physically intuitive, since the origins of K and K' propagate in opposite directions relative to the respective other frame. It is easily verified that $\mathbf{\Lambda}^{-1}(v)\,\mathbf{\Lambda}(v) = \mathbf{\Lambda}(v)\,\mathbf{\Lambda}^{-1}(v) = \mathbb{1}_4$ for the four-dimensional case. We are now in the position to introduce general four-vectors in Minkowski SpaceTime.

1.2.2 Four-Vectors (Co- and Contravariant)

1.2.2.1 Position four-vector

We have already given the time coordinate in SpaceTime physical dimension of length (through the multiplication with the speed of light), and so, with $x_0 = ct$, it is suggested to introduce a **position four-vector** as

$$x := \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \qquad \text{With components} \, \{x_\mu\} \tag{1.67}$$

where the x_0 is the **time-like** component and $x_k, k \in \{1, \dots, 3\}$ are the cartesian **space-like** components[17]. Then we can write the Lorentz boost also as

$$x' = \mathbf{\Lambda}_{\mathrm{K}\to\mathrm{K'}}\, x \tag{1.68}$$

[17] In the present case of the position vector, the time-like component is actually time itself and the space-like components represent space itself.

or, using component notation,

$$x'_\nu = \sum_{\mu=0}^{3} \Lambda_{\nu\mu}\, x_\mu \qquad \forall \nu \in \{0, \ldots, 3\} \tag{1.69}$$

This is the Minkowski-space equivalent of the transformation law of a vector, a three-tensor of rank 1, in real space! Here we transform a four-vector in Minkowski space.

From now on, we will use **Einstein summation convention** which is defined as a sum in Minkowski space (or coordinate space) over duplicate indices in any given term. Then,

$$x'_\nu = \Lambda_{\nu\mu}\, x_\mu. \tag{1.70}$$

From the above it immediately follows that

$$\frac{\partial x'_\nu}{\partial x_\mu} = \Lambda_{\nu\mu}. \tag{1.71}$$

Conversely, from Eq. (1.70), we have

$$\begin{aligned} \left(\Lambda^{-1}\right)_{\kappa\nu}\, x'_\nu &= \left(\Lambda^{-1}\right)_{\kappa\nu}\, \Lambda_{\nu\mu}\, x_\mu \\ &= \delta_{\kappa\mu}\, x_\mu \\ &= x_\kappa \\ \left(\Lambda^{-1}\right)_{\mu\nu}\, x'_\nu &= x_\mu. \end{aligned} \tag{1.72}$$

and so

$$\frac{\partial x_\mu}{\partial x'_\nu} = \left(\Lambda^{-1}\right)_{\mu\nu}. \tag{1.73}$$

1.2.2.2 General Contra- and Covariant Four-Vectors

Let us now take a first look at fields in special relativity. Be $\varphi(x)$ a scalar differentiable field with x the position four-vector in frame K. We regard the derivative with respect to coordinates in K', *i.e.*,

$$\frac{\partial \varphi(x)}{\partial x'_\mu} = \frac{\partial \varphi(x)}{\partial x_\kappa}\, \frac{\partial x_\kappa}{\partial x'_\mu} \tag{1.74}$$

since $\varphi\left(x(x')\right)$, where in the second equality the chain rule and Einstein summation have been used. This means that we can write for the differential operator

$$\frac{\partial}{\partial x'_\mu} = \frac{\partial}{\partial x_\kappa}\frac{\partial x_\kappa}{\partial x'_\mu} = \frac{\partial}{\partial x_\kappa}\left(\Lambda^{-1}\right)_{\kappa\mu} \tag{1.75}$$

where we have used Eq. (1.73). If we compare this result with Eq. (1.70), we see a difference: The components of the position four-vector are transformed from K to K' *via* the transformation matrix Λ, but the components of the derivative vector with respect to position transform — also from K to K' — *via* the *inverse* transformation matrix Λ^{-1}! This implies that there are, generally speaking, two types of four-vectors in relativity theory[18]:

1. The **contravariant components** of four-vectors transform from K to K' as to the above Lorentz transformation. They are, by convention, written with upper indices, a^μ.

2. The **covariant components** of four-vectors transform from K to K' as to the inverse Lorentz transformation. They are conversely written with lower indices, b_ν.

Let us now consider the scalar product of a covariant four-vector with a contravariant four-vector in Minkowski SpaceTime. For consistency with general matrix algebra, we will here consider covariant vectors as row vectors, i.e., $x'_\mu = x_\nu\left(\Lambda^{-1}\right)^\nu{}_\mu$ and always **sum over repeated upper and lower indices**. Then

$$\begin{aligned} b'_\mu a^{\mu'} &= b_\nu\left(\Lambda^{-1}\right)^\nu{}_\mu\,\Lambda^\mu{}_\kappa\,a^\kappa \\ &= b_\nu\,\delta^\nu_\kappa\,a^\kappa \qquad (1.76)\\ &= b_\nu\,a^\nu = b_\mu\,a^\mu. \qquad (1.77)\end{aligned}$$

[18]The reader may argue that this situation actually already exists in non-relativistic physics for position and gradient vectors, for example. This is true, but for a Euklidean metric there is no difference between the contra- and covariant components of a vector!

We find that the scalar product of a covariant with a contravariant vector is invariant under Lorentz transformation. $b_\nu\, a^\nu$ is therefore called a **Lorentz scalar**. It can without difficulties be shown that $a^{\mu'} b'_\mu = a^\nu b_\nu$, and so the scalar product of any contravariant with any covariant four-vector is also a Lorentz scalar.

The result in Eq. (1.77) for general four-vectors will be of utmost importance in the construction of relativistic theories.

1.2.2.3 Relationship Between Contra- and Covariant Four-Vectors

In Eq. (1.58) we had agreed that the scalar product involves the "modern metric", $\{g_{\mu\nu}\}$. It shall by definition have the property $g_{\mu\nu} = g^{\mu\nu}$. From this it directly follows that

$$(gg)_\mu^\kappa = g_{\mu\nu}\, g^{\nu\kappa} = \delta_\mu^\kappa \tag{1.78}$$

where δ_μ^κ is the usual Kronecker delta symbol[19].

Let us now see how contra- and covariant components of a four-vector are related to each other. The expression $b_\nu\, a^\nu$ is a Lorentz scalar. So is $b^\nu\, a_\nu$. Since these two Lorentz scalars are made up of the same four-vectors, they should be identical. We now suppose that

$$b^\nu\, a_\nu = b_\mu\, g^{\mu\nu}\, g_{\nu\kappa}\, a^\kappa \tag{1.79}$$

which can simply be calculated, using Eq. (1.78):

$$b_\mu\, g^{\mu\nu}\, g_{\nu\kappa}\, a^\kappa = b_\mu\, \delta_\kappa^\mu\, a^\kappa = b_\mu\, a^\mu = b_\nu\, a^\nu \tag{1.80}$$

By comparing Eqs. (1.79) and (1.80) we have the following relationships between co- and contravariant indices of a given four-vector:

$$b^\nu \;=\; b_\mu\, g^{\mu\nu} \tag{1.81}$$

$$a_\nu \;=\; g_{\nu\kappa}\, a^\kappa \tag{1.82}$$

[19]In the case of diagonal matrices we do not have to care about which index is the row and which is the column index, so we use this simplified notation.

In other words, contravariant and covariant vectors of the same type differ in their signs on the space-like components. For the case of the position four-vector we have explicitly

$$x^{\mu=0} = ct \tag{1.83}$$
$$x^{\mu=1} = x \tag{1.84}$$
$$x_{\mu=0} = ct \tag{1.85}$$
$$x_{\mu=1} = -x \tag{1.86}$$

Choosing the vector components with upper indices to correspond to non-relativistic notation is a plain matter of convention. This means that the previous formulation of the position four-vector in the usual non-covariant notation will now be changed. We choose the same four-vector to be defined by its contravariant components (upper indices), as

$$x := \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \qquad \text{With components } \{x^\mu\} \tag{1.87}$$

As a known example, consider the scalar product of the position vector with itself. Obviously,

$$\begin{aligned} x'_\mu \, x^{\mu'} = x^{\mu'} \, x'_\mu &= x^\nu \, x_\nu = x_\nu \, x^\nu \\ (ct')^2 - (x')^2 - (y')^2 - (z')^2 &= (ct)^2 - x^2 - y^2 - z^2 \end{aligned} \tag{1.88}$$

which is just identical to Eq. (1.58), confirming the logical consistency of the formalism.

The generalization of these concepts to rank-n four-tensors is straightforward and will be useful for the reformulation of electrodynamics and in the framework of general relativity, as well as in many other fields of physics.

1.2.3 SpaceTime Diagrams

The phenomenological consequences of the Lorentz tranformation can be conveniently visualized through a technique already introduced by Minkowski: SpaceTime diagrams. A reduction to two dimensions – time and one spatial dimension – is sufficient for many purposes.

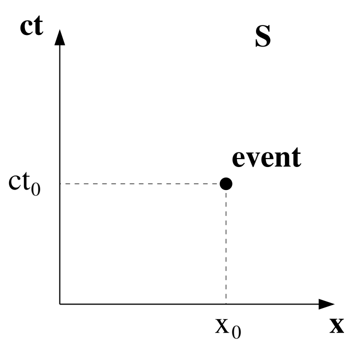


Figure 1.8:

An **event** is represented by a point in SpaceTime with coordinates $\{ct_0, x_0\}$.

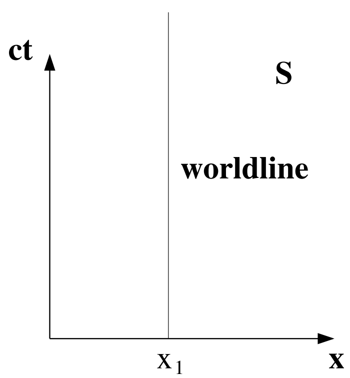


Figure 1.9:

An observer/object at rest at position x_1 is represented by a **worldline** in the SpaceTime diagram.

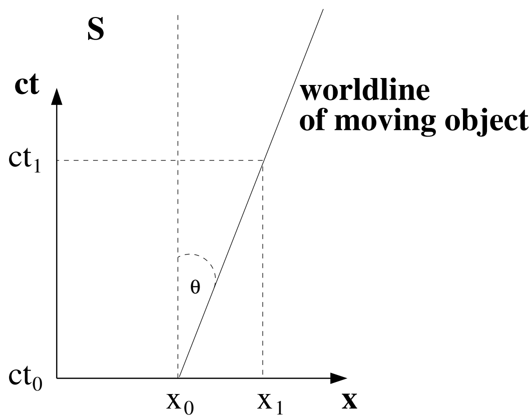
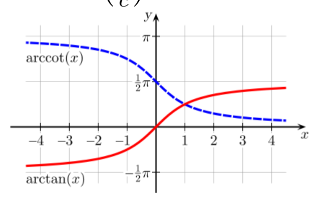


Figure 1.10:

An object moving in x-direction with velocity v in coordinates of S. This velocity follows from graphical observation: $\tan\vartheta = \frac{x_1 - x_0}{ct_1 - ct_0} = \frac{v(t_1-t_0)}{c(t_1-t_0)} = \frac{v}{c}$. So $\vartheta = \arctan\left(\frac{v}{c}\right)$.



Since $1 \geq \frac{v}{c} \geq 0$ we have $\max\left(\arctan\left(\frac{v}{c}\right)\right) = \max(\vartheta) = \frac{\pi}{4}$.

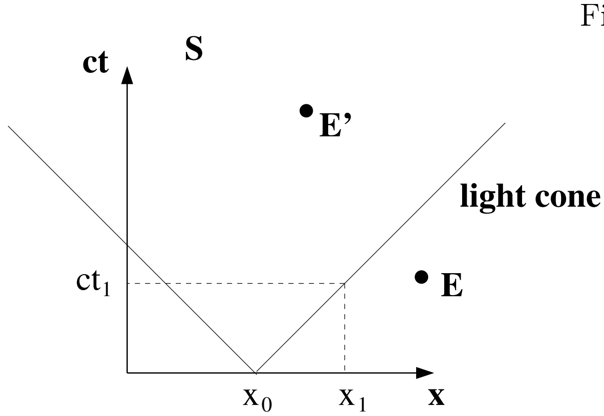


Figure 1.11:

Emission of a light pulse as event at $\{ct_0, x_0\}$ and propagation of the light pulse in frame S. Event E cannot be **causally connected** with an event at $\{ct_0, x_0\}$, because information cannot propagate faster than with c. Event E' can be **causally connected** with an event at $\{ct_0, x_0\}$ because it is inside the light cone.

As a simple illustration for the scenario with event E, imagine a distant observer appearing within $\delta t = \varepsilon t$ around t_0 at $x_E > x_1$. The observer cannot see the light pulse because he has "disappeared" before the light pulse can reach his position.

1.2.4 Space-, Light-, and Time-Like Four-Vectors

With the establishments of the preceding subsections we can revisit four-vectors in the context of SpaceTime diagrams and come to a couple of interesting conclusions.

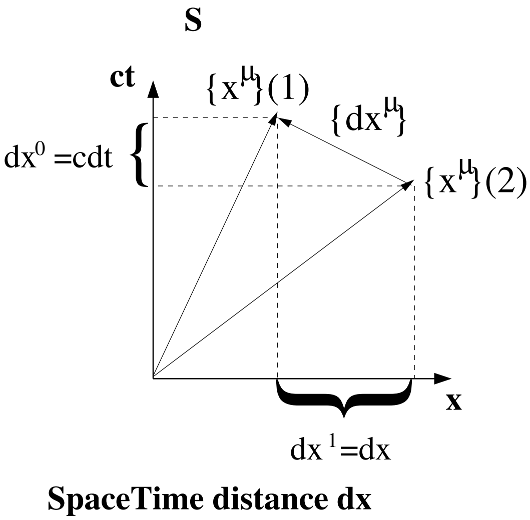


Figure 1.12: Two events (1) and (2) and their SpaceTime distance represented by vectors in a two-dimensional vector space.

From Fig. (1.12) we infer that

$$\begin{aligned} dx^0 &= x^0(1) - x^0(2) = c(t_1 - t_2) = cdt \\ dx^1 &= x^1(1) - x^1(2) = x_1 - x_2 = dx. \end{aligned}$$

[20] The scalar product of the four-vector dx with itself is then

$$\begin{aligned} D &:= (ds)^2 = dx^\mu dx_\mu = dx^\mu\, g_{\mu\nu}\, dx^\nu \\ &= \left(c(t_1 - t_2) \;\; x_1 - x_2 \right) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} c(t_1 - t_2) \\ (x_1 - x_2) \end{pmatrix} \\ &= \left(c(t_1 - t_2) \;\; x_1 - x_2 \right) \begin{pmatrix} c(t_1 - t_2) \\ -(x_1 - x_2) \end{pmatrix} \qquad (1.89) \\ &= c^2(t_1 - t_2)^2 - (x_1 - x_2)^2. \qquad (1.90) \end{aligned}$$

D is a Lorentz scalar (it is the scalar product of a contra- and a co-variant four-vector), i.e., all of our following conclusions are Lorentz invariant. We distinguish three general cases:

$D = 0$ From this it follows that $c^2(t_1 - t_2)^2 = (x_1 - x_2)^2$ which means that ds is on a light cone.

Four-vectors v with $v^\mu v_\mu = 0$ are called "**light-like**" four-vectors.

[20] The notation dx^0 here means $dx^{\mu=0}$.

$D < 0$ This would correspond to the case shown in Fig. (1.12) if we would assume the events to be the points chosen for illustration. Then, if we suppose that for the Lorentz-transformed spatial components $x_1^{k'} - x_2^{k'} = 0 \quad \forall k \in \{1, \ldots, 3\}$, then $\Rightarrow D = c^2(t_1' - t_2')^2 \geq 0$ which is in contradiction with the assumption. This means that given $D < 0$ there is no reference frame K' in which the two events occur at the same position and chronologically.

On the other hand, if $t_1' - t_2' = 0 \Rightarrow D = -(x_1' - x_2')^2 \leq 0$. Therefore, a reference frame K' exists such that the two events occur **simultaneously** but at **different positions**. D is ensuingly a "**space-like**" interval[21]. Note that in this case event 1 is outside the light cone based in event 2.

Four-vectors v with $v^\mu v_\mu < 0$ are called "**space-like**" four-vectors.

$D > 0$ In this case, if $t_1' - t_2' = 0$, *i.e.*, in a frame S' the two events occur simultaneously, then $\Rightarrow D = -(x_1' - x_2')^2 \leq 0$ contradicts the assumption and the two events cannot occur simultaneously in any K'.

However, if $x_1^{k'} - x_2^{k'} = 0 \rightarrow D = c^2(t_1' - t_2')^2 > 0$ is possible which means that there exists a frame K' in which the two events occur at the **same point in space** and **chronologically**. They can be causally connected (inside the respective light cone).

Four-vectors v with $v^\mu v_\mu > 0$ are called "**time-like**" four-vectors.

[21] to be distinguished from the space-like *components* of a four-vector which is quite a different thing

1.3 Relativistic Mechanics

The new (Einsteinian) view of space and time, or rather SpaceTime, entails that most of the physics that had been invented thus far had to be re-written in order to be in accord with the principles of special relativity. The new relativistic theories should be constructed such that their non-relativistic limit gives rise to the "old", classical version of pre-Einsteinian physical theory.

We will begin with mechanics. Newton's mechanics was known to be Galilei, but not Lorentz invariant. The evident program is thus to formulate a Lorentz invariant rendering of classical[22] mechanics. This shall be attempted by retaining the structure of the equations of motion and by replacing the ordinary three-vectors by four-vectors in SpaceTime.

1.3.1 Proper Time

In non-relativistic mechanics velocity is written as

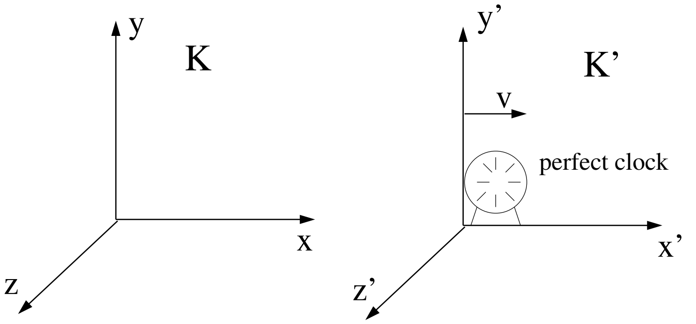
$$v_x = \frac{dx}{dt}. \tag{1.91}$$

We can guess that the relativistic version of velocity might involve the four-vector x and a four-vector v which should have the same transformation properties as x. But we know from Eq. (1.28) that time is obviously not a Lorentz scalar, $t' \neq t$ for a simple boost. In order to make v a contravariant vector, as we would expect it to be, we will have to formulate a Lorentz invariant quantity "dt".

We have seen as a result of the Michelson-Morley experiment using the Lorentz boost transformation that a time interval $\Delta t'$ in a clock's frame (time interval measured by the experimenters on Earth at a fixed

[22] I use the "international meaning" of 'classical' which is 'non-quantum' and not 'non-relativistic non-quantum'. In other words, we are going to develop relativistic classical mechanics.

position in K') relates to the time interval Δt for the same two events in an observer frame



A perfect (unaffected by any violent acceleration and undestroyable) clock in its rest frame K', moving with velocity v relative to a laboratory observer in K.

as to

$$\Delta t' = \frac{1}{\gamma(v)}\,\Delta t. \tag{1.92}$$

If we define $v(t)$ as the relative velocity in the infinitesimal time interval dt, we can generalize to time-dependent velocities and, therefore, accelerations. From now on we will call

$$d\tau := dt' = \frac{1}{\gamma(v)}\,dt \tag{1.93}$$

the **proper time** differential in the rest frame of the clock. Let's take a closer look at the properties of $d\tau$.

For this we calculate the scalar product of the four-distance ds in S with itself[23]:

$$\begin{aligned}
(ds)^2 &= dx^\mu\,dx_\mu = dx^\mu\,g_{\mu\nu}\,dx^\nu = (cdt)^2 - dx^2 - dy^2 - dz^2 \\
&= \left[c^2 - \frac{dx^2}{dt^2} - \frac{dy^2}{dt^2} - \frac{dz^2}{dt^2}\right] dt^2 \\
&= \left(c^2 - v^2\right)\,dt^2 = c^2\left(1 - \frac{v^2}{c^2}\right)\,dt^2 \\
&= c^2\,\frac{1}{\gamma(v)^2}\,dt^2 = c^2 d\tau^2
\end{aligned} \tag{1.94}$$

[23] Getting from the second to the third line can be seen by taking the scalar product of $\mathbf{v} = \frac{dx}{dt}\mathbf{e}_x + \frac{dy}{dt}\mathbf{e}_y + \frac{dz}{dt}\mathbf{e}_z +$ with itself. The resulting velocity is the speed of "something" in the coordinates of frame K, so this might be a clock at rest in K', and that is the reason why we can take this velocity as the one used in the Lorentz factor γ!

where in the last equality we have used Eq. (1.93).

However, we know that $ds^2 = ds'^2$ is a Lorentz scalar, see Eqs. (1.90) and (1.77). And since the speed of light, c, is the same in all reference frames, it follows that $d\tau^2$, and therefore also $d\tau$, have to be Lorentz scalars as well.

The physical interpretation of this finding is that the proper time interval (in the clock's frame) has to be the same for all observers. We can now proceed to building relativistic mechanics based on the proper time differential, $d\tau$.

1.3.2 Four-Velocity and Four-Acceleration

1.3.2.1 Four-Velocity

The obvious generalization of Eq. (1.91) in terms of four-vector quantities and the proper time differential is

$$\left\{ \frac{dx^\mu}{d\tau} \right\} =: \{u^\mu\} \tag{1.95}$$

Now, by construction, the components u^μ transform like the components of a contravariant four-vector, x^μ, because $d\tau$ is Lorentz invariant. However, Eq. (1.95) is a sort of "mixed" expression, since τ refers to a clock's rest frame whereas x^μ is a coordinate of an arbitrary frame [24].

We derive the components of velocity in a general frame K as follows:

$$u^\mu = \frac{dx^\mu(t)}{d\tau} = \frac{dx^\mu(t)}{dt}\,\frac{dt}{d\tau} = \gamma(v)\,\frac{dx^\mu(t)}{dt} \tag{1.96}$$

where the chain rule has been used with t regarded as a function of τ.

Note that now we are considering the situation with a general three velocity $\mathbf{v} = \sum_{j=1}^{3} v_j \mathbf{e}_j$ in coordinates of K. The γ factor is then more precisely $\gamma(v := ||\mathbf{v}||)$.

[24] We may choose it to be the laboratory frame.

Summary for the velocity four-vector[25]:

$$\boxed{\left\{ \begin{array}{c} u^0 \\ u^k \end{array} \right\} \equiv \left\{ \begin{array}{c} \gamma c \\ \gamma v_k \end{array} \right\} \qquad \forall k \in \{1, \ldots, 3\} \quad (1.97)}$$

It is immediately obvious that in the non-relativistic limit the space-like components of velocity turn into the usual velocities with respect to reference frame K. The time-like component does not exist as component of a four vector in non-relativistic theory.

Let's have a look at the scalar product of the velocity four-vector with itself. We know from Eq. (1.77) that this product $u^2 = u^\mu\, u_\mu = u_\mu\, u^\mu$ has to be a Lorentz scalar. The direct calculation gives

$$u^\mu\, u_\mu = u^\mu\, g_{\mu\nu}\, u^\nu = \gamma(||\mathbf{v}||)^2 (c^2 - v_x^2 - v_y^2 - v_z^2) = \frac{c^2}{c^2 - \mathbf{v}^2}(c^2 - \mathbf{v}^2) = c^2. \tag{1.98}$$

It is confirmed that u^2 is a Lorentz scalar. Furthermore, since $c^2 > 0$ it follows that u is a time-like four-vector.

u^0 is the time-like and u^k a space-like component of the velocity four-vector. Its slope in SpaceTime is, therefore, $\frac{\gamma c}{\gamma v_x} = \frac{c}{v_x}$, taking only one spatial direction for simplicity. However, from the discussion around Fig. (1.10) it is evident that this slope corresponds to the slope of the worldline of a moving particle in SpaceTime which is also $\frac{c(t_1 - t_0)}{v_x(t_1 - t_0)} = \frac{c}{v_x}$. We conclude that the velocity four-vector is tangent to the worldline of a moving particle in SpaceTime which explains why the four-vector u is time like.

[25] Bear in mind that we here have the contravariant component of a four-vector on the left-hand side and the component of usual velocity in non-relativistic notation on the right-hand side.

1.3.2.2 Four-Acceleration

Following the same principles, it is a straightforward exercise to formulate four-acceleration.

$$\begin{aligned}\{b^\mu\} &= \left\{\frac{du^\mu(t)}{d\tau}\right\} \\ b^\mu &= \frac{du^\mu(t)}{d\tau} = \frac{du^\mu(t)}{dt}\,\frac{dt}{d\tau} = \gamma(v)\,\frac{du^\mu(t)}{dt}\end{aligned} \tag{1.99}$$

Using the results for four-velocity we can calculate the individual components of four-acceleration.

$$\begin{aligned}b^0 &= \gamma(v(t))\,\frac{c\,d\gamma(v(t))}{dt} = \gamma(v(t))\,c\,\frac{d\gamma(v)}{dv}\,\frac{dv}{dt} \quad \text{with} \quad v = ||\mathbf{v}|| \\ &\frac{d||\mathbf{v}||}{dt} = \frac{1}{||\mathbf{v}||}\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right) \\ &\frac{d\gamma(v)}{dv} = \gamma^3\frac{||\mathbf{v}||}{c^2} \\ \Rightarrow b^0 &= \frac{\gamma^4}{c}\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right)\end{aligned}$$

Similarly,

$$\begin{aligned}b^k &= \gamma(v)\,\frac{du^k}{dt} \\ &= \gamma(v)\,\frac{d}{dt}\big(\gamma(v(t))\,v_k(t)\big) \\ &= \gamma(v)\,\left\{\left[\frac{||\mathbf{v}||}{c^2}\gamma^3\,v_k(t)\,\frac{1}{||\mathbf{v}||}\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right)\right] + \gamma a_k\right\} \\ &= \gamma(v)\,\left[\frac{\gamma^3}{c^2}\,v_k\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right) + \gamma a_k\right] \\ b^k &= \frac{\gamma^4}{c^2}\,v_k\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right) + \gamma^2 a_k.\end{aligned}$$

Summary for the acceleration four-vector:

$$\boxed{\{b^\mu\} = \left\{ \begin{array}{c} \frac{\gamma^4}{c}\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right) \\ \frac{\gamma^4}{c^2}\,v_k\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right) + \gamma^2 a_k \end{array} \right\} \qquad \forall k \in \{1,\ldots,3\} \qquad (1.100)}$$

In the non-relativistic limit, we simply obtain:

$$\begin{aligned} \lim_{c\to\infty} b^0 &= 0 \\ \lim_{c\to\infty} b^k &\equiv a_k \end{aligned}$$

It can be shown that $b^\mu b_\mu < 0$ which means that four-acceleration is a space-like four-vector. However, it isn't really necessary to prove this since we can derive the result indirectly:

Using Eq. (1.98) we can also determine the Minkowski scalar product between four-velocity and four-acceleration. First it is noted that

$$\begin{aligned} \frac{d}{d\tau}\left(u^\mu u_\mu\right) &= \left(\frac{d}{d\tau}\,u^\mu\right) u_\mu + u^\mu \left(\frac{d}{d\tau}\,u_\mu\right) \\ &= \left(\frac{d}{d\tau}\,u^\mu\right) u_\mu + u^\mu \left(\frac{d}{d\tau}\,g_{\mu\nu}\,u^\nu\right) \\ &= \left(\frac{d}{d\tau}\,u^\mu\right) u_\mu + u^\nu\,g_{\nu\mu} \left(\frac{d}{d\tau}\,u^\mu\right) \\ &= \left(\frac{d}{d\tau}\,u^\mu\right) u_\mu + u_\mu \left(\frac{d}{d\tau}\,u^\mu\right) \\ &= 2b^\mu u_\mu \end{aligned}$$

$$b^\mu u_\mu = \frac{1}{2}\frac{d}{d\tau}\left(u^\mu u_\mu\right) = \frac{1}{2}\frac{d}{d\tau}c^2 = \frac{\gamma}{2}\,\frac{d}{dt}\,c^2 = 0 \qquad (1.101)$$

This means that in Minkowski SpaceTime four-velocity and four-acceleration are always orthogonal four-vectors.

1.3.3 Relativistic Version of Newton's Equation of Motion

We can now proceed to formulating the fundamental law of dynamics in the domain of special relativity. For this, a further definition is required.

We do not know at this point, how mass behaves under Lorentz transformation, but this will be derived. So we **define** that m_0 be the **mass of a particle in its rest frame**, which is then by definition a Lorentz scalar. This means that rest mass of a particle never changes, just like proper time is invariant of the state of movement of a relative observer.

In view of Newton's equation of motion from non-relativistic theory,

$$m\,\mathbf{a} = \sum_i \mathbf{F}_i = \mathbf{F} \tag{1.102}$$

where there is only one "type" of mass of a particle which never changes, irrespective of any state of motion and force is a three-vector, we **define**

$$\boxed{m_0\, b^\mu = K^\mu \qquad (1.103)}$$

to be the μth SpaceTime component of the **relativistic fundamental law of dynamics**. $m_0 b^\mu$ necessarily transforms like a contravariant four-vector since m_0 is a Lorentz scalar. Likewise, K^μ is a component of a contravariant four-vector which verfies the homogeneity of the equation in that respect. $\{K^\mu\}$ is called **Minkowski's force four-vector**.

Obviously, the guiding principles for this definition are the replacement of three- by four-vectors and the conservation of the form of the equation of motion in the relativistic domain. Nevertheless, the consequences of this formulation have to conform with experimental tests.

For the analysis of the new equation we resort to relations developed earlier. In subsection 1.3.1 it has been shown that for the proper time differential $d\tau = \frac{1}{\gamma(v)}\, dt$. Furthermore, using Eq. (1.99), Eq. (1.103) yields

$$m_0 \frac{du^\mu}{d\tau} = m_0 \gamma(v) \frac{du^\mu}{dt} = K^\mu \tag{1.104}$$

1.3.3.1 Space-Like Part of the Equation of Motion

Using Eq. (1.97) the **space-like part** of Eq. (1.104) becomes

$$\begin{aligned} m_0\, \gamma(v) \frac{d}{dt} \left(\gamma(v)\, v_k\right) &= K^k \\ \frac{d}{dt} \left[\gamma(v)\, m_0\, v_k\right] &= \frac{1}{\gamma(v)} K^k \end{aligned} \tag{1.105}$$

This equation is reminiscent of the temporal change of the quantity of movement, linear momentum, in non-relativistic mechanics which is related to the force acting on the particle:

$$\frac{d\mathbf{p}}{dt} = \mathbf{F} \tag{1.106}$$

The immediate implication is that we can postulate how relativistic linear momentum should be formulated. However, before doing so, let us complete the discussion of Minkowski's force four-vector.

First, in accord with this finding, it is postulated that

$$K^k :\equiv \gamma(v)\, F_k \tag{1.107}$$

for the relation between four- and three-force components. Again, components K^k and F_k become equivalent in the non-relativistic limit of relativistic theory (also notation-wise, considering our convention).

1.3.3.2 Time-Like Part of the Equation of Motion

Now it is time to derive the time-like component of the Minkowski force vector. For this, we multiply the relativistic fundamental law of dynamics, Eq. (1.103), by $g_{\mu\nu}\, u^\nu$ (including summation according to Einstein convention) and obtain

$$\begin{aligned} m_0\, b^\mu &= K^\mu \\ m_0\, g_{\mu\nu}\, u^\nu\, b^\mu &= g_{\mu\nu}\, u^\nu\, K^\mu \\ m_0\, u^\nu\, g_{\nu\mu}\, b^\mu &= K^\mu\, g_{\mu\nu}\, u^\nu \\ m_0\, u_\mu\, b^\mu &= K^0\, u^0 - \sum_{k=1}^{3} K^k\, u^k \end{aligned}$$

We can now readily use earlier results. For one thing, $u_\mu\, b^\mu = 0$ according to Eq. (1.101). For the other, the definition in Eq. (1.107) and the form of the velocity four-vector in Eq. (1.97) entail the identity $\sum_{k=1}^{3} K^k\, u^k = \gamma(v)\mathbf{F} \cdot \gamma(v)\mathbf{v}$, and so we arrive at

$$K^0\, u^0 - \gamma^2\, \mathbf{F} \cdot \mathbf{v} = 0 \tag{1.108}$$

$$K^0\, \gamma\, c = \gamma^2\, \mathbf{F} \cdot \mathbf{v}$$

$$K^0 = \frac{\gamma}{c}\, \mathbf{F} \cdot \mathbf{v} \tag{1.109}$$

In summary, Minkowski's force four-vector is thus written as

$$\boxed{\{K^\mu\} = \left\{ \begin{array}{c} \frac{\gamma}{c}\, \mathbf{F} \cdot \mathbf{v} \\ \gamma F_k \end{array} \right\} \quad \forall k \in \{1, \ldots, 3\} \quad (1.110)}$$

In the non-relativistic limit we find:

$$\begin{aligned} \lim_{c\to\infty} K^0 &= 0 \\ \lim_{c\to\infty} K^k &= F_k. \end{aligned}$$

With the above equations (1.100) and (1.110) Newton's equation of motion can be rewritten in relativistic form. The full relativistic mechanical equation of motion is thus given as

$$\boxed{\begin{array}{rcl} m_0\, b^\mu &=& K^\mu \; (1.111) \\ \hline m_0 \dfrac{\gamma^4}{c} \left(\mathbf{v} \cdot \dfrac{d\mathbf{v}}{dt} \right) &=& \dfrac{\gamma}{c} \mathbf{F} \cdot \mathbf{v} \\ m_0 \left[\dfrac{\gamma^4}{c^2} v_k \left(\mathbf{v} \cdot \dfrac{d\mathbf{v}}{dt} \right) + \gamma^2 a_k \right] &=& \gamma F_k \end{array}}$$

where the first line is the covariant notation with contravariant components of four-vectors, the second line is the time-like component and the third line are the space-like components ($\forall k$) of the equation of motion.

It is quite evident that this equation describes physics different from Newton's law. Focussing on the space-like part and neglecting the first term on the l.h.s. leads to

$$m_0 \gamma a_k \;=\; F_k.$$

This approximate equation differs from the non-relativistic law of motion through the presence of the γ factor. Moreover, the first contribution scales differently with γ and cannot compensate for this change as simple acceleration problems (like a particle in an external electric field) demonstrate.

1.4 Relativistic Formulation of Classical Electrodynamics

Let us briefly look at the reformulation of classical electrodynamics in terms of the new quantities for relativistic theory, four-tensors and Lorentz scalars.

1.4.1 A Digression on Units

A few comments concerning the choice of units for the following sections are in place. We replace the standard S.I. units in the following by Gaussian-based S.I. units which corresponds to making the replacements for the electric and magnetic field as well as the charge density

$$\begin{aligned} \mathbf{E} &\to \frac{\mathbf{E}}{\sqrt{4\pi\varepsilon_0}} \\ \mathbf{B} &\to \mathbf{B}\sqrt{\frac{\mu_0}{4\pi}} \\ \rho &\to \rho\sqrt{4\pi\varepsilon_0} \end{aligned} \tag{1.112}$$

Then, for instance, Faraday's law of induction

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad (S.I.) \tag{1.113}$$

becomes

$$\begin{aligned} \frac{1}{\sqrt{4\pi\varepsilon_0}} \boldsymbol{\nabla} \times \mathbf{E} &= -\sqrt{\frac{\mu_0}{4\pi}} \frac{\partial \mathbf{B}}{\partial t} \\ \boldsymbol{\nabla} \times \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} \qquad (\text{Gaussian } S.I.) \end{aligned} \tag{1.114}$$

with $\sqrt{\mu_0\varepsilon_0} = \frac{1}{c}$. We furthermore see that in the Gaussian system the electric and magnetic fields have the **same** units.

Since $q \propto \rho$, Coulomb's law becomes

$$\mathbf{F}_{12} = \frac{q_1 q_2}{r_{12}^2} \mathbf{e}_{12} \tag{1.115}$$

We will also be interested in Ampère's law:

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{1.116}$$

With the replacements defined in Eq. (1.112) Ampère's law is written as

$$\begin{aligned}
\sqrt{\frac{\mu_0}{4\pi}}\, \boldsymbol{\nabla} \times \mathbf{B} &= \mu_0 \sqrt{4\pi\varepsilon_0}\, \mathbf{J} + \frac{\mu_0 \varepsilon_0}{\sqrt{4\pi\varepsilon_0}} \frac{\partial \mathbf{E}}{\partial t} \\
\sqrt{\mu_0 \varepsilon_0}\, \boldsymbol{\nabla} \times \mathbf{B} &= 4\pi \mu_0 \varepsilon_0 \, \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \\
\boldsymbol{\nabla} \times \mathbf{B} &= \frac{1}{c} \left(4\pi \mathbf{J} + \frac{\partial \mathbf{E}}{\partial t} \right) \qquad (1.117)
\end{aligned}$$

1.4.2 Continuity Equation

This is a simple – but important – example to begin with. The continuity equation of electrodynamics is derived from applying the theorem of Gauss and Ostrodradsky[26] to the relation between charged current density flux through a surface and change of total charge in a volume delimited by that surface:

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} = -\boldsymbol{\nabla} \cdot \mathbf{J}(\mathbf{r}, t) \tag{1.118}$$

Note that the continuity equation takes the same form in the Gaussian unit system.

Before proceeding, we introduce a new quantity, the four-vector of charged current density

$$\boxed{\{J^\mu\} = \left\{ \begin{array}{c} J^0 = c\rho \\ J^k \equiv J_k \end{array} \right\}^{22} \quad \forall k \in \{1, \ldots, 3\}. \qquad (1.119)}$$

[26] Which can be proven for **any** vector field on purely geometric grounds.

The above form of the four-vector seems to be an *ad hoc* assumption at this stage. However, considering that classical current density is charge density times velocity, we see that Eq. (1.119) has the same structure as the velocity four vector in Eq. (1.97). Note that the physical dimension of the time-like component of this four-vector is $\left[\frac{Q}{L^3} \times \frac{L}{T}\right]$, *i.e.*, charge density times velocity, which is the same as $\dim[\vec{J}]$.

We now reformulate Eq. (1.118), using $\partial_0 = \frac{\partial}{\partial x^0} = \frac{1}{c}\frac{\partial}{\partial t}$:

$$\begin{aligned} \frac{1}{c}\frac{\partial}{\partial t} c\rho + \boldsymbol{\nabla} \cdot \mathbf{J} &= 0 \\ \frac{\partial}{\partial x^0} c\rho + \sum_{j=1}^{3} \frac{\partial}{\partial x^j} J^j &= 0 \\ \partial_0 J^0 + \sum_{j=1}^{3} \partial_j J^j &= 0 \\ \partial_\mu J^\mu &= 0 \end{aligned} \tag{1.120}$$

Eq. (1.120) is manifestly written in covariant form, with either side of the equation a Lorentz scalar. In other words, the continuity equation of electrodynamics is Lorentz invariant, even though the individual terms (∂_μ and J^μ) transform as four-vectors. This also means that charge conservation is independent of the inertial frame in which it is regarded.

It is an instructive exercise to show that the continuity equation of electrodynamics is not Galilei invariant. This can be achieved, for example, by using the Galilei boost transformation in Eqs. (12) and the elucidations from section 0.2.

[22] Note that the same symbol "J" is used for denoting four-vectors and three-vectors, so the identity $J^k \equiv J_k$ means the equivalence between the contravariant four-vector component J^k and J_k in non-relativistic notation, NOT the equivalence with the covariant components of J.

1.4.3 Maxwell's Equations

We begin by defining a four-vector of the electromagnetic potential:

$$\boxed{\{A^\mu\} = \left\{ \begin{array}{c} A^0 \\ A^k \end{array} \right\} \equiv \left\{ \begin{array}{c} V \\ A_k \end{array} \right\} \qquad \forall k \in \{1, \dots, 3\} \qquad (1.121)}$$

where V is the usual scalar potential and $\mathbf{A}$ is the vector potential (where again the same symbol is used in covariant and non-relativistic notation).

Further, we require what is called the **tensor of the electromagnetic field**

$$F^{\mu\nu} := \partial^\mu A^\nu - \partial^\nu A^\mu \tag{1.122}$$

F is here defined as a rank-2 four-tensor with two contravariant indices. Covariant (or mixed) components of this tensor can be derived by using the metric tensor. Evidently, the left- and the right-hand side of Eq. (1.122) transform in the same manner under Lorentz transformation, but of course F is not a Lorentz scalar.

The elements of the field tensor shall be determined for two examples.

$$\begin{aligned} F^{01} &= \partial^0 A^1 - \partial^1 A^0 = \partial_0 A^1 + \partial_1 A^0 \\ &= \frac{\partial}{\partial x^0} A^1 + \frac{\partial}{\partial x^1} A^0 = \frac{\partial}{\partial x} V + \frac{1}{c}\frac{\partial}{\partial t} A_x \\ &= -\left(-\boldsymbol{\nabla} V - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \right)_x \\ &= -E_x \end{aligned}$$

using a combination of the structure equations of electrodynamics in the last step[27]. As a second example consider

$$\begin{aligned}
F^{12} &= \partial^1 A^2 - \partial^2 A^1 = \frac{\partial}{\partial x_1} A^2 - \frac{\partial}{\partial x_2} A^1 \\
&= -\frac{\partial}{\partial x^1} A^2 + \frac{\partial}{\partial x^2} A^1 = -\frac{\partial}{\partial x} A_y + \frac{\partial}{\partial y} A_x \\
&= -(\boldsymbol{\nabla} \times \mathbf{A})_z \\
&= -B_z
\end{aligned}$$

and so on for the remaining elements. Obviously, $F^{\mu\mu} = 0$, and so the field tensor takes on the form

$$\{F^{\mu\nu}\} \equiv \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{1.124}$$

We now postulate that the **inhomogeneous Maxwell equations (Gauss's and Ampère's law)** can be written in a very elegant and compact form:

$$\boxed{\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} J^\nu \qquad (1.125)}$$

This shall be verified by two examples. First, setting $\nu = 0$:

$$\begin{aligned}
\partial_0 F^{00} + \partial_1 F^{10} + \partial_2 F^{20} + \partial_3 F^{30} &= \frac{4\pi}{c} J^0 \\
0 + \frac{\partial}{\partial x} E_x + \frac{\partial}{\partial y} E_y + \frac{\partial}{\partial z} E_z &= \frac{4\pi}{c} c\rho \\
\boldsymbol{\nabla} \cdot \mathbf{E} &= 4\pi\rho
\end{aligned} \tag{1.126}$$

[27] Writing the electric field as

$$\begin{aligned}
\mathbf{E} &= -\boldsymbol{\nabla} V - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \\
\boldsymbol{\nabla} \times \mathbf{E} &= -\boldsymbol{\nabla} \times \boldsymbol{\nabla} V - \frac{1}{c}\frac{\partial}{\partial t} \boldsymbol{\nabla} \times \mathbf{A} \\
\boldsymbol{\nabla} \times \mathbf{E} &= -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}
\end{aligned} \tag{1.123}$$

leads to the Maxwell-Faraday equation.

which is nothing else than Gauss's law in differential form[28]. Now we set $\nu = 1$:

$$\begin{aligned} \partial_0 F^{01} + \partial_1 F^{11} + \partial_2 F^{21} + \partial_3 F^{31} &= \frac{4\pi}{c} J^1 \\ \frac{1}{c}\frac{\partial}{\partial t}(-E_x) + 0 + \frac{\partial}{\partial y} B_z + \frac{\partial}{\partial z}(-B_y) &= \frac{4\pi}{c} J_x \qquad (1.127) \end{aligned}$$

$$\left(\boldsymbol{\nabla} \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} \right)_x = \frac{4\pi}{c} J_x \qquad (1.128)$$

which is one cartesian component of Ampère's law.

The power in the formulation of Eq. (1.125) lies in the fact that it is written in a homogeneous way in terms of four-vectors ($\partial_\mu F^{\mu\nu}$ transforms like a contravariant four-vector) which makes it Lorentz covariant, *i.e.*, form invariant with respect to Lorentz transformations. This is by no means obvious in the original formulation of Maxwell's equations.

[28] It is easily checked that the S.I. form of Gauss's law $\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}$ becomes Eq. (1.126) upon making the replacements from Eqs. (1.112).

1.5 Relativistic Mass and Linear Momentum

We have seen in the preceding section that the equation of motion of classical mechanics has been generalized to Minkowski space and is formulated in terms of four-vectors. In the following subsections we want to investigate the consequences of this generalization. We begin from a space-like component of the relativistic equation of motion (1.103) and rewrite it:

$$\begin{aligned} m_0\, b^k &= K^k \\ m_0\, \frac{du^k}{d\tau} &= \gamma\, F_k \\ m_0\, \frac{du^k}{dt} &= F_k \\ \frac{d\,(m_0\,\gamma\, v_k)}{dt} &= F_k \end{aligned}$$

where the definition of four-acceleration (1.99), the proper time differential (1.93), and the obtained expression for four-velocity (1.97) have been used. Form equivalence with Eq. (1.106) and dimensional analysis suggest to define **relativistic linear momentum** as

$$p^k \equiv m_0\,\gamma\, v_k. \tag{1.129}$$

So we have established the space-like components of the linear-momentum four-vector. Before completing the four-vector it is instructive to inspect Eq. (1.129) more closely.

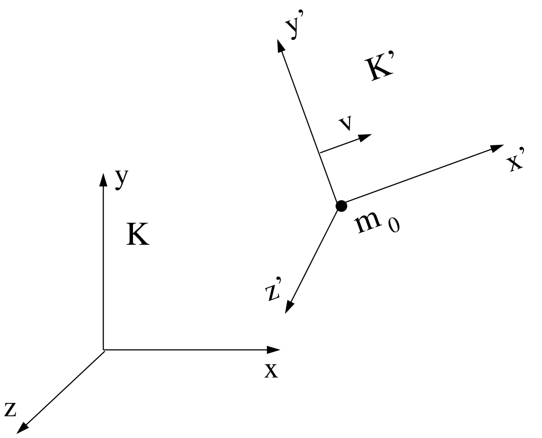
1.5.1 Relativistic Mass

In the framework of classical Newtonian mechanics $\mathbf{p} = m\,\mathbf{v}$ where m is the inertial mass of a given particle or body. Since in Eq. (1.129) v_k is the velocity of the particle in frame K and p^k the associated momentum, the implication is that

$$m := \gamma\, m_0 \tag{1.130}$$

should be regarded as the particle's **relativistic inertial mass** in frame K, instead of simply the the rest mass m_0 of the particle. This is a profound difference and means that, since $\gamma = \gamma(v)$ is a function of the velocity of the particle in frame K, so is its mass. The finding is illustrated in Fig. (1.13). Note that rest mass is a Lorentz scalar, i.e.,

Figure 1.13:



A massive particle with rest mass m_0 in frame K' moves with velocity $\mathbf{v}$ relative to frame K. Its relativistic mass in coordinates of frame K is $\gamma(v)\, m_0$.

it does not depend on any state of movement. However, in coordinates of frame K, the particle "behaves" as if it had an increased mass, its relativistic or dynamic mass (or observed mass)[29]. It can be anticipated that also the expression for the energy of the particle in K should differ from that in frame K', but that is yet to be substantiated.

1.5.2 Relativistic Linear Momentum

With these conclusions in mind, the relativistic generalization of linear momentum is straightforward. Since the space-like components are proportional to the dynamic particle mass and its velocity in frame K, the time-like component results by analogy and using the expression for the velocity four-vector in Eq. (1.97):

$$p^0 = m_0\, \gamma\, c = m_0\, u^0 \tag{1.131}$$

[29]The concept of relativistic mass is not a fundamental requirement. In fact, many authors argue against its introduction since it is sufficient to consider relativistic momentum modified by the γ factor. However, dynamic mass is a useful way of thinking about various situations for instance in the physics of an atom.

Summary for the linear momentum four-vector:

$$\{p^\mu\} = \left\{ \begin{array}{c} p^0 \\ p^k \end{array} \right\} = \left\{ \begin{array}{c} m_0\, u^0 \\ m_0\, u^k \end{array} \right\} \equiv \left\{ \begin{array}{c} mc \\ mv_k \end{array} \right\} \quad \forall k \in \{1, \ldots, 3\} \quad (1.132)$$

As a check for consistency, we take the proper-time derivative of linear momentum,

$$\frac{d}{d\tau}\, p^\mu = m_0\, \frac{d}{d\tau}\, u^\mu = m_0\, b^\mu = K^\mu \quad (1.133)$$

where the relativistic equation of motion (1.103) has been used. We thus obtain a law analogous in form to its non-relativistic counterpart. We can also verify Lorentz covariance on this last equation. Since $\frac{d}{d\tau}$ is a Lorentz scalar, the l.h.s. (left-hand side) transforms like a contravariant four vector, and so does the r.h.s.

1.6 Relativistic Energy

We now start from the time-like component of the relativistic fundamental law of dynamics (1.103) and obtain

$$\begin{aligned} m_0\, b^0 &= K^0 \\ m_0\, \frac{du^0}{d\tau} &= \frac{\gamma}{c}\, \mathbf{F} \cdot \mathbf{v} \\ m_0\, \gamma\, \frac{du^0}{dt} &= \frac{\gamma}{c}\, \mathbf{F} \cdot \mathbf{v} \\ \frac{d}{dt}\left(m_0\, \gamma(v)\, c^2\right) &= \mathbf{F} \cdot \mathbf{v} \end{aligned} \quad (1.134)$$

where the expression for the proper time differential (1.93) and the time-like component of four-velocity (1.97) have been used.

Integrating Eq. (1.134) over time results in

$$\int \frac{d}{dt}\left(m_0\, \gamma(v)\, c^2\right) dt = \int \mathbf{F} \cdot \mathbf{v} dt \quad (1.135)$$

$$m_0\, \gamma(v)\, c^2 = \int \mathbf{F} \cdot \mathbf{dx} \quad (1.136)$$

using $\mathbf{v}dt = \mathbf{dx}$.

Now, since work is $W = \int \mathbf{F} \cdot \mathbf{dx}$ the quantity on the left-hand side must correspond to **relativistic energy**,

$$E := m_0\, \gamma\, c^2. \tag{1.137}$$

It still has to be clarified which kind of energy is represented by E. Using the expression for relativistic inertial mass, Eq. (1.130), the relativistic energy can also be written in its (publicly) famous form:

$$\boxed{E = m\, c^2 \qquad (1.138)}$$

The physical meaning of this equation is that every energy corresponds to a mass and every mass corresponds to an energy, with the Lorentz scalar c^2 being the conversion factor.

The most startling consequence of this expression becomes evident when the special case for a particle at rest with respect to frame K is considered. Then, $||\mathbf{v}|| = 0$ and so $\gamma(||\mathbf{v}||) = 1$ and therefore $m = m_0$. In that case,

$$E = E_0 := m_0\, c^2 \tag{1.139}$$

and we find that the rest mass of a body corresponds to an energy! We call E_0 the **rest energy** of a particle of rest mass m_0. This implies that energy should be convertible into rest mass (energy) and vice versa[30].

It is very important to analyze the expression Eq. (1.137) before taking any further steps. m_0 and c are Lorentz scalars, but the Lorentz factor γ is a function of velocity. We Taylor expand the Lorentz factor

[30] When formulating this equation, Einstein regarded it true from pure formal aesthetics, but considered believing in its veracity in practice as an "act of faith". It had to be confirmed by experiment, which happened in the decades to come.

about $v_0 = 0$, resulting in

$$\gamma(v) = \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{d^n\gamma(v)}{dv^n} \right|_{v=v_0} (v - v_0)^n \tag{1.140}$$

$$= 1 + \frac{1}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \mathcal{O}\left(\left(\frac{v}{c}\right)^6\right) \tag{1.141}$$

We again see that in the non-relativistic limit, $c \to \infty$, the Lorentz factor becomes 1. The above representation as a Taylor expansion is often useful in order to represent leading relativistic effects by truncating the expansion at some appropriate order.

Using the expansion Eq. (1.141) in Eq. (1.137) one obtains

$$E = m_0\, c^2 + \frac{1}{2} m_0 v^2 + \frac{3}{8} m_0 v^2 \frac{v^2}{c^2} + \mathcal{O}\left(\left(\frac{v}{c}\right)^4\right) \tag{1.142}$$

In this form relativistic energy can be straightforwardly analyzed:

$\frac{1}{2} m_0 v^2$ Beginning with the known term, this represents the kinetic energy of a body with inertial mass $m = m_0$ as in non-relativistic classical mechanics.

$m_0\, c^2$ As a consequence, this term is of relativistic origin. It is evidently a Lorentz scalar, and it relates to an energy of the body independent of kinematics. It is, therefore, called the **rest energy** of the particle.

$\frac{3}{8} m_0 v^2 \frac{v^2}{c^2}$ Since this contribution vanishes in the non-relativistic limit, it is also of relativistic origin and represents the leading **relativistic correction to the particle's kinetic energy**.

$\mathcal{O}\left(\left(\frac{v}{c}\right)^4\right)$ Consequently, all following terms are also relativistic corrections to the particle's kinetic energy.

1.6.1 Relativistic Energy-Momentum Relation

We will now establish an equivalent to the energy-momentum relation from non-relativistic mechanics which reads

$$T = \frac{\left(\mathbf{p}^N\right)^2}{2m_0} \quad \text{with } \mathbf{p}^N = m_0\,\mathbf{v}. \tag{1.143}$$

In the relativistic régime we realize that we can identify one relationship between relativistic momentum and relativistic energy right away. From Eqs. (1.132) and (1.138) it follows that the time-like component of the relativistic momentum four-vector is

$$p^0 \equiv mc = \frac{E}{c} \tag{1.144}$$

which means that relativistic energy appears on the time-like component of the linear momentum four-vector, such that we can recast it in another (equivalent form):

$$\{p^\mu\} = \left\{ \begin{array}{c} \frac{E}{c} \\ p^k \end{array} \right\} \quad \forall k \in \{1, \ldots, 3\} \tag{1.145}$$

The derivation of the relativistic energy-momentum relation is then just a formal exercise. We start from two equivalent forms of the momentum Lorentz scalar $p^2 = p^\mu p_\mu = p^0 p_0 + \sum\limits_k p^k p_k$:

$$\begin{aligned} p^2 &= \frac{E^2}{c^2} - \mathbf{p}^2 \\ p^2 &= m_0^2\,\gamma^2\,c^2 - m_0^2\,\gamma^2\,\mathbf{v}^2 \end{aligned} \tag{1.146}$$

where the second relation follows from Eqs. (1.129) and (1.132). The first relation can be rewritten as

$$E^2 = c^2\left(p^2 + \mathbf{p}^2\right) \tag{1.147}$$

and inserting the second relation into it yields

$$\begin{aligned} E^2 &= c^2 \left(m_0^2 \gamma^2 c^2 - m_0^2 \gamma^2 \mathbf{v}^2 + \mathbf{p}^2 \right) \\ &= \mathbf{p}^2 c^2 + m_0^2 c^2 \left(\gamma^2 c^2 - \gamma^2 \mathbf{v}^2 \right) \\ &= \mathbf{p}^2 c^2 + m_0^2 c^2 \left(\frac{c^2}{1 - \frac{\mathbf{v}^2}{c^2}} - \frac{\mathbf{v}^2}{1 - \frac{\mathbf{v}^2}{c^2}} \right) \\ &= \mathbf{p}^2 c^2 + m_0^2 c^4 \left(\frac{c^2}{c^2 - \mathbf{v}^2} - \frac{\mathbf{v}^2}{c^2 - \mathbf{v}^2} \right) \end{aligned}$$

and so we obtain

$$E^2 = \mathbf{p}^2 c^2 + m_0^2 c^4. \tag{1.148}$$

Taking the positive square root gives

$$\boxed{E = \sqrt{\mathbf{p}^2 c^2 + m_0^2 c^4} \quad (1.149)}$$

which is known as the **relativistic energy-momentum relation**. Note that the first term under the square root is the square of linear *three*-momentum, not to be confused[31] with the scalar product of four-momentum in Eq. (1.146).

At the time of its first appearance, there was no dispute about taking into account the positive square root only, although formally the negative square root could also be permissible. After all, the notion of *negative energy* of a free particle seems queer. This point became a remarkable twist in the history of physics and will be picked up again later on in the context of relativistic quantum mechanics.

A number of interesting conclusions can be drawn from Eq. (1.149) when considering various possible cases, for which it is equally valid. The first distinction concerns the body's rest mass, m_0.

[31] Some texts use unclear notation on this point.

1.6.1.1 Particles with zero rest mass

As a first observation, we note that this case is particular to relativistic theory. The notion of a particle with zero mass makes no sense in non-relativistic mechanics. Due to the intimate relationship between energy and mass in Eq. (1.138), however, we must take this possibility seriously here.

Omitting the rest-mass term from Eq. (1.149) yields

$$E = ||\mathbf{p}||\, c. \tag{1.150}$$

However, we know that relativistic three-momentum is

$$\mathbf{p} = m_0\, \gamma(\mathbf{v})\, \mathbf{v}$$

so the energy of the "particle" seems to be zero, except if the particle is allowed to travel at the speed of light in frame K, in which case the γ factor tends to infinity! The problem obviously remains not fully resolved in classical relativistic mechanics, but a first glance at **quantum mechanics** in this context reveals an interesting connection:

Inserting de Broglie's relation ($p = \frac{h}{\lambda}$) and Planck-Einstein's relation ($E = h\nu$) which are valid for de Broglie matter waves, where h is Planck's constant, into Eq. (1.150) we obtain

$$\begin{aligned} h\nu &= \frac{h}{\lambda}\, c \\ \nu &= \frac{c}{\lambda} \end{aligned}$$

which is the well-known relationship between frequency and wavelength for waves propagating according to Maxwell's equations[32].

We conclude that the theory of **massless particles** is **necessarily a relativistic quantum theory**. We will make a first step toward this

[32] Note that trying to make the same argument based on non-relativistic momentum $p = mv$ does not lead to a consistent theory. In that case, since $\lambda = \frac{h}{mc}$, supposing propagation at the speed of light, the wavelength for a particle whose mass tends to zero becomes infinite which is in contradiction with observation. In other words, non-relativistic quantum mechanics "works" for massive particles at lower velocities.

theory in later sections (for massive particles), but a conclusive answer for massless particles will have to await the introduction of relativistic quantum field theory.

1.6.1.2 Massive Particles

For massive particles we can discuss two limiting cases:

$v << c$. From Eq. (1.142) in this approximation it follows that

$$E \approx m_0\, c^2 + \frac{\left(\mathbf{p}^N\right)^2}{2m_0} \qquad \text{for } m_0 \neq 0 \tag{1.151}$$

$v \approx c$. According to Eq. (1.132) for the linear relativistic three-momentum we can write

$$\mathbf{p}^2\, c^2 = m_0^2 \gamma^2\, \mathbf{v}^2 c^2 = m_0^2 c^4\, \frac{\mathbf{v}^2}{c^2 - \mathbf{v}^2} \tag{1.152}$$

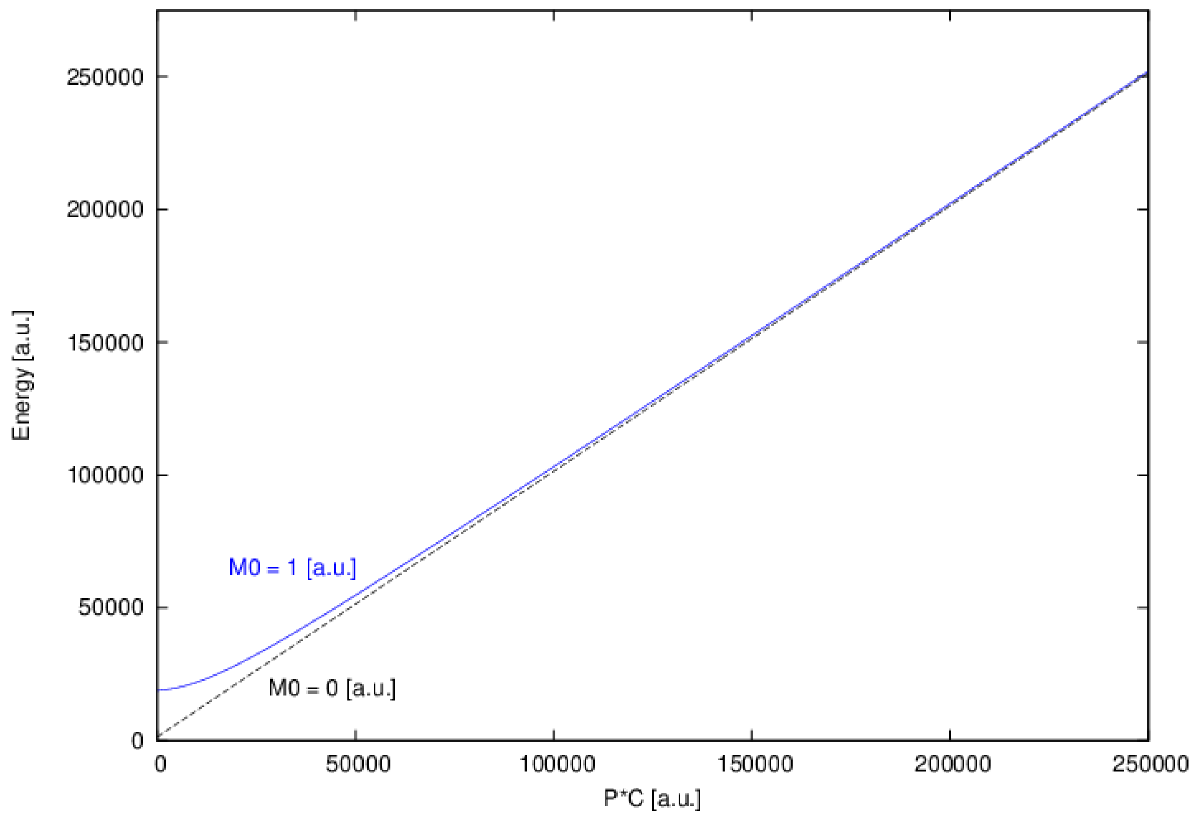
Now since $\frac{\mathbf{v}^2}{c^2-\mathbf{v}^2} >> 1$ with the applied condition it follows that here $\mathbf{p}^2\, c^2 >> m_0^2 c^4$ and it can be approximated

$$E \approx ||\mathbf{p}||\, c \qquad \text{for } m_0 \neq 0 \tag{1.153}$$

in this limit. As a consequence, for very large relative velocities the rest energy becomes negligible as a contribution to the total relativistic energy.

Summary. Relativistic energy as a function of linear momentum for vanishing (e.g. for the photon) and non-vanishing rest mass is depicted in Fig. (1.14).

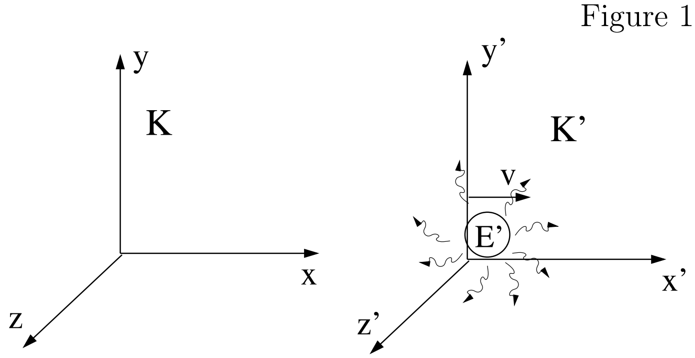
Figure 1.14:



1.6.2 Energy-Mass Equivalence: Mass Defect

In order to deepen the understanding of Eqs. (1.138) and (1.149) we will consider the following thought experiment. Be there a system at rest with respect to a frame K' that moves with velocity v relative to a laboratory frame K, Fig. (1.15).

Figure 1.15:



During a short time span $\Delta t'$ the system emits radiation of energy E' in K' symmetrically such that its total momentum in K' does not change, *i.e.*, it remains at rest in K'.

Such an emission process may, for example, occur in the formation of an atomic nucleus from its constituent nucleons (protons and neutrons). The energy is released in form of ejected particles (neutrons, α particles etc.) including kinetic energy, and radiation.

Since the energy pulse is emitted in a spherically symmetric manner its total 3-momentum in K' is zero[33]. Then the momentum four-vector for the released energy pulse takes on the following form in K and K', respectively:

Momentum four-vector in K: Momentum four-vector in K':

$$\{p^\mu\} = \left(\frac{E_{\text{rad}}}{c}, p^1 \equiv p, 0, 0\right) \qquad \{p^{\mu\prime}\} = \left(\frac{E'_{\text{rad}}}{c}, 0, 0, 0\right)$$

Of course, the momentum four-vector $\{p^{\mu\prime}\}$ has to be related with $\{p^\mu\}$ through a Lorentz transformation. In the present case we want to transform from K' to K, so we use $\mathbf{\Lambda}^{-1}(v) = \mathbf{\Lambda}(-v)$ with respect to Eq. (1.66). The transformation of momentum for the boost then reads

$$\begin{aligned} \mathbf{\Lambda}(-v) \begin{pmatrix} p^{0\prime}_{\text{rad}} \\ p^{1\prime}_{\text{rad}} \end{pmatrix} &= \begin{pmatrix} p^0_{\text{rad}} \\ p^1_{\text{rad}} \end{pmatrix} \\ \begin{pmatrix} \gamma & \frac{v}{c}\gamma \\ \frac{v}{c}\gamma & \gamma \end{pmatrix} \begin{pmatrix} \frac{E'_{\text{rad}}}{c} \\ 0 \end{pmatrix} &= \begin{pmatrix} \gamma \, \frac{E'_{\text{rad}}}{c} \\ \frac{v}{c}\,\gamma\,\frac{E'_{\text{rad}}}{c} \end{pmatrix} = \begin{pmatrix} p^0_{\text{rad}} \\ p^1_{\text{rad}} \end{pmatrix} \end{aligned} \tag{1.154}$$

The resulting four-vector has to be identical with the original formulation of the momentum four-vector in frame K. Let us inspect the relevant space-like component of three-momentum. We find

$$p^1_{\text{rad}} = p_{\text{rad}} = \frac{v}{c}\,\gamma\,\frac{E'_{\text{rad}}}{c} \tag{1.155}$$

which is the momentum (in K) corresponding to the energy of the radiation pulse in K'. However, the system did not change its momentum in K', due to the assumed symmetrical emission. This implies that its

[33] For instance, photons have momentum $p = \frac{h}{\lambda}$.

velocity relative to the laboratory has also not changed. What are the consequences?

At this point, it is imperative to exploit principles of symmetry. The system of our thought experiment is **isolated**. Therefore, it is **invariant to a spatial translation** and so its **total momentum in a given frame is conserved**.

The basic form of momentum in accord with Eq. (1.132) for the system is

$$p_{\rm sys} = m_{\rm sys}\,\gamma(v_{\rm sys})\,v_{\rm sys} \tag{1.156}$$

for the component of interest, where $m_{\rm sys}$ here is the rest mass of the system. This quantity has to be conserved due to symmetry. But Eq. (1.155) forces us to consider the momentum of the radiation pulse that is non-zero in the balance of momentum in K. And since relative velocity does not change due to the pulse, the only way to compensate is by a loss of rest mass of the system due to the emission of radiation! Formally, momentum conservation in K is thus written as

$$\begin{aligned} p_{{\rm sys}_{\rm before}} &= p_{{\rm sys}_{\rm after}} + p_{\rm rad} \\ m_{{\rm sys}_{\rm before}}\,\gamma\,v &= m_{{\rm sys}_{\rm after}}\,\gamma\,v + \frac{v}{c}\,\gamma\,\frac{E'_{\rm rad}}{c} \end{aligned} \tag{1.157}$$

from which it follows that

$$m_{{\rm sys}_{\rm before}} = m_{{\rm sys}_{\rm after}} + \frac{E'_{\rm rad}}{c^2}. \tag{1.158}$$

The energy of the radiation pulse divided by the square of the speed of light is the **rest mass lost** by the system due to the emission of the radiation pulse of energy E'. The system has, therefore, suffered a **mass defect**, defined as

$$\frac{E'_{\rm rad}}{c^2} = m_{{\rm sys}_{\rm before}} - m_{{\rm sys}_{\rm after}} \tag{1.159}$$

Note that the mass defect is Lorentz invariant because $E'_{\rm rad}$ refers to the rest frame of the system.

Table 1.1: Nuclear mass, mass defect and ratio for some sample nuclides

Nuclide	Mass $[u]$	Mass defect $[u]$	$\frac{\text{Mass defect } [u]}{\text{Mass } [u]}$
^{4}He	4.002603	0.029279	0.0073
^{12}C	12.000000	0.095646	0.0080
^{59}Fe	58.934874	0.540247	0.0092
^{225}Ra	225.023611	1.803782	0.0080

$$u = 931.49410242(28)\,[\tfrac{\text{MeV}}{c^2}]$$

1.7 Relativistic Kinematics of Particle Interactions

With the developments of the previous sections it is possible to study collisions of bodies in the classical (non-quantum) regime. We can even take a look at particle decays, although we here surpass the classical notion of a particle[34]. A number of general assumptions are made that lead to important simplifications:

1. **External forces have no influence on the collision process.** This implies the conservation of total energy and total momentum (before and after the process) which will be rewritten in relativistic form.

2. We here do **not consider the details of the collision process at very short range**. For example, the collision of two neutrons could be studied at the level of the constituent quarks which would require a deeper understanding of the bound state of the neutron.

Let us first set the stage by re-iterating the principles of such collisions in non-relativistic theory. Generally, primes (') denote properties after the process, no primes denote properties before the process.

1.7.1 Non-relativistic collision processes

1. $\sum_j m_j = \sum_j m'_j$
   Total mass of all intervening particles (j) is conserved. Bodies may break up in the process (or stick together), but the sum of the inertial masses always remains the same.

2. $\sum_j \mathbf{p}_j = \sum_j \mathbf{p}'_j$
   All components of total momentum are conserved. This immedi-

[34] Particles generally have finite lifetimes and the decay of a particle is a complicated quantum process where the probability of decay per unit of time plays an important role.

ately follows from the fact that there are no external forces, by assumption (consider Noether's theorem).

3. Kinetic energy T may or may not be conserved.
   Of course, **total** energy is conserved, but a process may be such that kinetic energy of incident bodies is converted to some form of internal energy (such as vibrational energy, alas, heat). The relevant distinctions are made in the following.

- **Elastic collisions.** These are characterized by a conservation of total kinetic energy
  $\sum\limits_j T_j = \sum\limits_j T'_j$
  There is no conversion of kinetic energy into internal energy or vice versa.

- **Inelastic collisions.** Here total kinetic energy is not conserved. We distinguish between two cases:

  1. *"Sticky"* collisions, in which kinetic energy decreases:
     $\sum\limits_j T_j > \sum\limits_j T'_j$
  2. *"Explosive"* collisions, in which kinetic energy increases:
     $\sum\limits_j T_j < \sum\limits_j T'_j$

  If there are no internal degrees of freedom available – such as is the case for elementary particles – then the non-relativistic physics picture breaks down.

1.7.2 Relativistic collision processes

As before, external forces are irrelevant, and the conservation laws can now be written in four-vector form. For this, we use the momentum four-vector from Eq. (1.145) before and after the collision:

$$\begin{pmatrix} \sum\limits_j \frac{E_j}{c} \\ \sum\limits_j \mathbf{p}_j \end{pmatrix} = \begin{pmatrix} \sum\limits_j \frac{E'_j}{c} \\ \sum\limits_j \mathbf{p}'_j \end{pmatrix} \tag{1.160}$$

Note that here $E_j = E_{0j} + T_j$ is the total relativistic energy, the sum of rest energy and relativistic kinetic energy of the particle.

Also as before, kinetic energy may or may not be conserved, depending on whether a conversion into or from internal energy is characteristic of the collision. The corresponding distinctions are in this context made as follows:

- **Elastic collisions.** These are characterized by a conservation of total rest and kinetic energy
  $\sum\limits_j E_{0j} = \sum\limits_j E_{0j}'$
  $\sum\limits_j T_j = \sum\limits_j T_j'$
  There is no conversion of kinetic energy into internal (rest) energy or vice versa.

- **Inelastic collisions.** Here neither total kinetic energy nor internal (rest) energy are conserved:

  1. *"Sticky"* collisions, in which the decrease of total kinetic energy is accompanied by an increase of total rest energy:
     $\sum\limits_j T_j > \sum\limits_j T_j'$
     $\sum\limits_j E_{0j} < \sum\limits_j E_{0j}'$
     This is the typical scenario in collider physics where very heavy particles are created from lighter particles using their great incident kinetic energies.

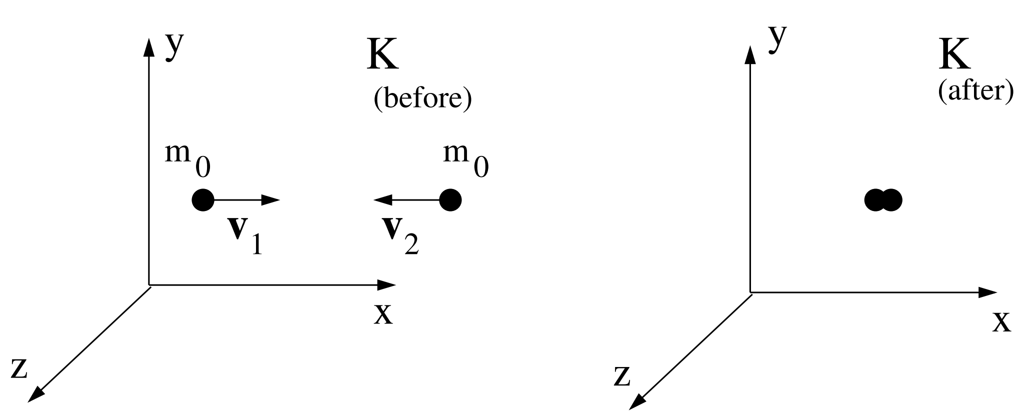
2. *"Explosive"* collisions, where it is the other way around:
$\sum_j T_j < \sum_j T'_j$
$\sum_j E_{0j} > \sum_j E_{0j}'$

The discussion of several exemplifying cases is useful for two reasons: First, examples make the peculiarities of relativistic kinematics very clear. Second, the treatment of concrete cases reveals the sometimes particular techniques of calculation.

1.7.2.1 "Sticky" Two-body Frontal Collision

The first example is very simple. We imagine two particles (or pieces of clay) of equal rest mass m_0 at high velocity, $v = \frac{3}{5}\,c$, in frontal collision under the sole assumption that the two bodies form a single body after collision (extreme sticky collision). We wish to determine what the rest

Figure 1.16:



mass, M, of the resulting body will be.

We start out by giving the 3-momentum for the two particles in K (the lower index is here a particle index)

$$\begin{aligned}\mathbf{p}_1 &= m_0\,\gamma_1\,\mathbf{v}_1\\ \mathbf{p}_2 &= m_0\,\gamma_2\,\mathbf{v}_2\end{aligned}$$

Since by assumption $\mathbf{v}_2 = -\mathbf{v}_1$ it follows that $\gamma_1 = \gamma_2$ and, therefore,

$$\mathbf{p}_2 = -\mathbf{p}_1 \tag{1.161}$$

Momentum conservation now dictates that

$$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}_M$$

From this and Eq. (1.161) it follows that

$$\mathbf{p}_M = \mathbf{0} \Rightarrow \gamma_M = 1 \Rightarrow M = M_0$$

If the relativistic mass of the resulting body equals its rest mass in frame K, then its relativistic energy is

$$E_M = M_0\,c^2$$

We now invoke the conservation of energy, i.e., the relativistic energies before and after the collision have to be identical:

$$E_1 + E_2 = M_0\,c^2 \tag{1.162}$$

Since $\gamma_1 = \gamma_2 := \gamma$ we have

$$E_1 = m_0\,\gamma\,c^2 = E_2$$

The γ factor for the incident particles is calculated as $\gamma = \frac{1}{\sqrt{1-\frac{9}{25}\frac{c^2}{c^2}}} = \frac{1}{\sqrt{\frac{16}{25}}} = \frac{5}{4}$ and so the total energy becomes

$$E_1 + E_2 = 2m_0\,\gamma\,c^2 = \frac{5}{2}\,m_0\,c^2$$

Therefore, with Eq. (1.162),

$$\begin{aligned} M_0\, c^2 &= \frac{5}{2}\, m_0\, c^2 \\ \Leftrightarrow M_0 &= \frac{5}{2}\, m_0 \end{aligned} \tag{1.163}$$

The essential observation is that the final rest mass is $\frac{5}{2}\, m_0$ which **surpasses** the initial rest mass which is $\frac{4}{2}\, m_0$. In other words, since the kinetic energy of the created particle is zero, the **complete kinetic energy of the incident particles has been converted into rest energy** of the resulting particle[35].

1.7.3 Spontaneous Two-body Decay

Suppose now that the initial particle has a finite lifetime[36]. So this is an example of a relativistic explosive "collision". Under the assumption that it decays into two particles of equal rest mass, what will the velocities of these be in K? Similar to the above example, we start out from momentum conservation:

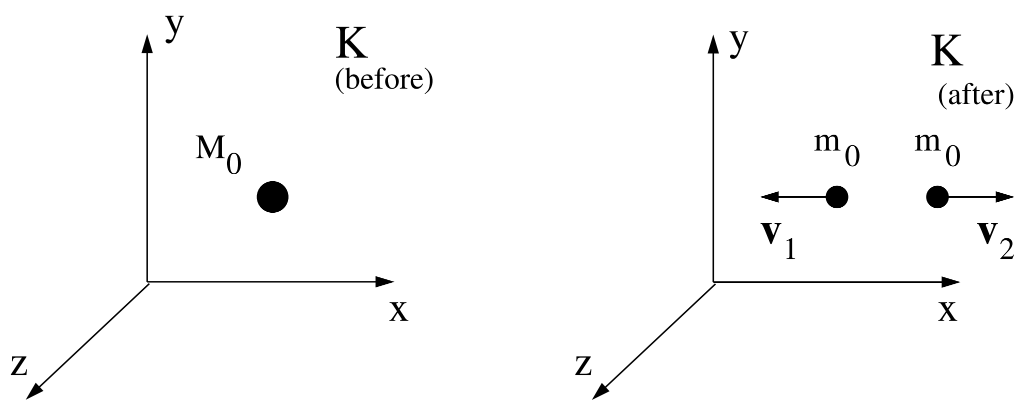
$$\begin{aligned} \mathbf{p}_M &= \mathbf{0} = \mathbf{p}_1 + \mathbf{p}_2 \\ \Rightarrow \mathbf{p}_1 &= -\mathbf{p}_2 \end{aligned} \tag{1.164}$$

and the created particles necessarily are ejected "back to back". Since

[35] Had we taken two lumps of clay in the non-relativistic picture, then this possibility would not be all that astonishing since we could imagine that the incident kinetic energy has been converted into internal energy, say heat, represented by the vibrational energy of the clay molecules. Relativity allows, however, that a body is created that has no obvious substructure (or at least not one we know of today). This is astonishing: Particles orders of magnitude heavier than the sum of the incident particles can be created, and this happens copiously at accelerator facilities.

[36] Clearly, this is a departure from the classical worldview. We here enter a semi-quantum theory in which excited states may decay into more stable states and thus have finite lifetimes.

Figure 1.17:



their rest masses are assumed to be equal,

$$\begin{aligned} \gamma_1\,\mathbf{v}_1 &= -\gamma_2\,\mathbf{v}_2 \\ \Rightarrow \frac{1}{\gamma_2}\,||\mathbf{v}_1|| &= \frac{1}{\gamma_1}\,||\mathbf{v}_2|| \\ \Rightarrow \left(c^2 - v_2^2\right) v_1^2 &= \left(c^2 - v_1^2\right) v_2^2 \\ \Leftrightarrow v_1^2 &= v_2^2 \\ \Rightarrow \gamma_1 &= \gamma_2 \end{aligned} \tag{1.165}$$

The velocities and gamma factors of the two resulting particles are identical, not surprisingly. Due to energy conservation

$$\begin{aligned} M_0 c^2 &= 2m_0\,\gamma\,c^2 \\ \Leftrightarrow \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} &= \frac{M_0}{2m_0} \\ \Rightarrow \frac{c^2}{c^2-v^2} &= \frac{M_0^2}{(2m_0)^2} \\ \Leftrightarrow c^2 - v^2 &= \frac{(2m_0)^2\,c^2}{M_0^2} \\ \Leftrightarrow v &= c\sqrt{1-\left(\frac{2m_0}{M_0}\right)^2} \end{aligned} \tag{1.166}$$

where only the positive root is taken since we understand $v = ||\mathbf{v}||$. Thus, we find the condition

$$M_0 \geq 2m_0 \tag{1.167}$$

since otherwise velocity becomes imaginary. This result implies that there is a **threshold energy** E_T, a minimum energy of the incident particle, defined as

$$E_T = M_0\,c^2 := 2m_0c^2 \tag{1.168}$$

for the two-body decay into two particles of identical rest mass to take place. A surplus of energy in the form of initial kinetic energy is not an obstacle.

This result is of general importance as it helps exclude possible decay processes in particle physics based on the rest masses of the involved particles alone. A good example is the deuteron d with $m_d = 1875.6\,[\frac{\mathrm{MeV}}{c^2}]$. The presumed decay process $d \longrightarrow p+n$ is kinematically impossible since $m_p + m_n = 1877.9\,[\frac{\mathrm{MeV}}{c^2}]$ which makes the deuteron a stable particle.

1.7.4 Pion Decay and Special Methods

As a final example which also introduces a clever technique for solving problems of this kind, we consider the decay of the π-meson (or simply pion) π^- ($\equiv d\overline{u}, C = -1, m_{\pi^-} = 139.57\,[\frac{\mathrm{MeV}}{c^2}]$) into a muon μ^- ($C = -1, m_{\mu^-} = 105.66\,[\frac{\mathrm{MeV}}{c^2}]$) and a muon-antineutrino[37] $\overline{\nu}_\mu$ ($C = 0, m_{\overline{\nu}_\mu} \approx 0$), i.e.,

$$\pi^- \longrightarrow \mu^- + \overline{\nu}_\mu \tag{1.169}$$

Neutrinos and their antiparticles have — for all practical purposes — zero rest mass and thus propagate with the speed of light. The particular question of the **velocity of the resulting muon** is interesting here. Moreover, we want to express this velocity solely in terms of the known parameters of the problem which are the rest masses and the speed of light.

First, a useful little theorem can be derived. Three-momentum and energy of a given particle in the lab frame are

$$\begin{aligned} \mathbf{p} &= m_0\,\gamma\,\mathbf{v} \\ E &= m_0\,\gamma\,c^2 \end{aligned}$$

and so by division of these two identities

$$\begin{aligned} \frac{\mathbf{p}}{E} &= \frac{\mathbf{v}}{c^2} \\ \Leftrightarrow \mathbf{v} &= \frac{\mathbf{p}c^2}{E} \end{aligned} \tag{1.170}$$

In other words, if the momentum and energy of a particle are known, so is its velocity. We write this expression specifically as a norm for the muon quantities

$$||\mathbf{v}_\mu|| = \frac{||\mathbf{p}_\mu||c^2}{E_\mu} \tag{1.171}$$

and evaluate it.

[37] It has to be an *anti*-neutrino due to conservation laws that will be discussed further down the road.

An elegant trick to obtain the required momentum and energy is to work from the Minkowski-space scalar product four-conservation. Four-conservation is for the present case written in terms of the relevant momentum four-vectors:

$$\begin{aligned} p_\pi = \begin{pmatrix} p_\pi^0 = \frac{E_\pi}{c} \\ \mathbf{p}_\pi \end{pmatrix} &= p_\mu + p_{\overline{\nu}} = \begin{pmatrix} p_\mu^0 + p_{\overline{\nu}}^0 = \frac{E_\mu}{c} + \frac{E_{\overline{\nu}}}{c} \\ \mathbf{p}_\mu + \mathbf{p}_{\overline{\nu}} \end{pmatrix} \\ \Leftrightarrow p_{\overline{\nu}} &= p_\pi - p_\mu \end{aligned} \tag{1.172}$$

Based on the last identity we calculate the scalar product of the anti-neutrino four-momentum with itself:

$$p_{\overline{\nu}}^2 = p_\pi^2 + p_\mu^2 - 2p_\pi \cdot p_\mu \tag{1.173}$$

The various terms in Eq. (1.173) are now calculated one by one.

1. Since the problem is treated in the rest frame of the pion its three-momentum is zero and so, using the relativistic energy-momentum relation Eq. (1.149), $E_\pi = \sqrt{\mathbf{p}_\pi^2 c^2 + m_\pi^2 c^4} = m_\pi c^2 = E_{0_\pi}$. Now we do the four-scalar product

$$p_\pi^2 = \frac{E_{0_\pi}^2}{c^2} = \frac{m_\pi^2 c^4}{c^2} = m_\pi^2 c^2 \tag{1.174}$$

2. The created muon cannot be expected to have zero momentum in K. The corresponding calculation for the next term is, therefore,

$$p_\mu^2 = \frac{E_\mu^2}{c^2} - ||\mathbf{p}_\mu||^2 = \frac{||\mathbf{p}_\mu||^2\, c^2 + m_\mu^2 c^4}{c^2} - ||\mathbf{p}_\mu||^2 = m_\mu^2 c^2 \tag{1.175}$$

3. With the above findings the Minkowski scalar product between muon and pion four-momenta is easily obtained as

$$p_\pi \cdot p_\mu = \frac{E_\pi}{c}\,\frac{E_\mu}{c} - \mathbf{p}_\pi \cdot \mathbf{p}_\mu = \frac{m_\pi c^2}{c}\,\frac{E_\mu}{c} - 0 = m_\pi\, E_\mu \tag{1.176}$$

   By theory (see Eq. (1.77)) $p_\pi \cdot p_\mu = (p_\pi)^\lambda\, (p_\mu)_\lambda$ is a Lorentz scalar, as will become evident in the following.

4. Finally, the left-hand side of Eq. (1.173) gives, using the energy-momentum relation for massless particles Eq. (1.150),

$$p_{\overline{\nu}}^2 = \frac{E_{\overline{\nu}}^2}{c^2} - ||\mathbf{p}_{\overline{\nu}}||^2 = \frac{||\mathbf{p}_{\overline{\nu}}||^2\, c^2}{c^2} - ||\mathbf{p}_{\overline{\nu}}||^2 = 0 \tag{1.177}$$

With the results from calculations $1 \ldots 4$ Eq. (1.173) becomes

$$\begin{aligned} 0 &= m_\pi^2 c^2 + m_\mu^2 c^2 - 2m_\pi\, E_\mu \\ \Leftrightarrow E_\mu &= \frac{m_\pi^2 + m_\mu^2}{2m_\pi}\, c^2 \end{aligned} \tag{1.178}$$

and so we have obtained an expression for the first ingredient required to evaluate Eq. (1.171). A similar calculation leads to the momentum of the muon:

$$\begin{aligned} p_\mu &= p_\pi - p_{\overline{\nu}} \\ \Rightarrow p_\mu^2 &= p_\pi^2 + p_{\overline{\nu}}^2 - 2p_\pi \cdot p_{\overline{\nu}} \\ \Leftrightarrow m_\mu^2 c^2 &= m_\pi^2 c^2 + 0 - 2\,\frac{E_\pi}{c}\,\frac{E_{\overline{\nu}}}{c} - \mathbf{p}_\pi \cdot \mathbf{p}_{\overline{\nu}} \end{aligned} \tag{1.179}$$

$\mathbf{p}_\pi = \mathbf{0}$ and thus $E_\pi = m_\pi c^2$. Furthermore, since the anti-neutrino is massless, $E_{\overline{\nu}} = ||\mathbf{p}_{\overline{\nu}}||\, c = ||\mathbf{p}_\mu||\, c$. The latter identity follows from momentum conservation

$$\begin{aligned} \mathbf{p}_\pi &= \mathbf{0} = \mathbf{p}_\mu + \mathbf{p}_{\overline{\nu}} \\ \Leftrightarrow \mathbf{p}_\mu &= -\mathbf{p}_{\overline{\nu}} \\ \Rightarrow ||\mathbf{p}_\mu|| &= ||\mathbf{p}_{\overline{\nu}}|| \end{aligned} \tag{1.180}$$

Putting all of this together yields for Eq. (1.179)

$$\begin{aligned} m_\mu^2 c^2 &= m_\pi^2 c^2 - 2\, m_\pi\, ||\mathbf{p}_\mu|| c \\ \Leftrightarrow ||\mathbf{p}_\mu|| &= \frac{-m_\mu^2 + m_\pi^2}{2\, m_\pi}\, c \end{aligned} \tag{1.181}$$

The velocity of the created muon results from the combination of Eqs. (1.171), (1.178), and (1.181) to be

$$||\mathbf{v}_\mu|| = \frac{\frac{-m_\mu^2+m_\pi^2}{2\,m_\pi}\,c^3}{\frac{m_\pi^2+m_\mu^2}{2m_\pi}\,c^2} = \frac{m_\pi^2 - m_\mu^2}{m_\pi^2 + m_\mu^2}\,c \tag{1.182}$$

which depends only on the rest masses of the involved particles.

Using these known masses, $v_\mu \approx 0.271\,c$.

Chapter 2

(A Brief) Introduction to Elementary Particles

2.1 Standard Model Phenomenology

I shall here give a very brief overview of some important developments and discoveries in general physics and elementary particle physics relevant to the introduction to nuclear physics that will follow suite. Some of the material for the following section is taken from D. Griffiths, "Introduction to Elementary Particles".

2.1.1 Historical Notes

In 1932, the world of elementary particles was simple: The only known particles were the heavy proton (p) and neutron (n), the light electron (e), and the carrier of electromagnetic interactions, the photon (γ). What followed in the next three decades was a period of intense discovery of formerly unknown particles. The world of physics was swamped with discoveries in such a manner that Willis Lamb ('Lamb shift') joked in 1955: "For some time, a Nobel Prize was given out for the discovery of a new particle. Nowadays, someone who finds a new particle should be fined 10.000 $."

Until the 1950s, experimental particle physics was largely confined to the study of cosmic radiation and its atmospheric products. In 1952 the first modern particle accelerator went into activity, the 'Brookhaven Cosmotron' on Long Island, just outside New York. It could accelerate particles up to roughly 1 [GeV] kinetic energy (for comparison, the Large Hadron Collider currently reaches 6500 [GeV] kinetic energy per beam.)

2.1.2 Mesons

Mesons are "medium-weight" particles, hypothesized by H. Yukawa in 1934, that play a role in nuclear physics. To understand Yukawa's idea we take a look at a general parameterization of the four known forces of nature in terms of their range. Neglecting force constants, the following proportionality holds:

$$F_i \propto \frac{e^{-\frac{r}{a}}}{r^2} \tag{2.1}$$

where r is the distance between bodies and a is a range parameter that roughly takes on the following values:

i: electromagnetic, gravitational	$a \longrightarrow \infty$
i: strong	$a \approx 1$ [fermi] $= 10^{-15}$ [m]
i: weak	$a \approx 10^{-16} \ldots 10^{-17}$ [m]

So for electromagnetism and gravity the force follows a $\frac{1}{r^2}$ law. For the strong and the weak interactions it becomes very small when r becomes greater than characteristic nuclear/nucleon length scales (the proton radius is roughly $r_p \approx 0.88 \times 10^{-15}$ [m]).

Based on these experimental notions and basic quantum mechanical arguments, Yukawa proposed that there should be a particle that is exchanged between nucleons, responsible for the stability of an atomic nucleus. The nuclide ^{4}He ($2p$, $2n$), was known to be stable, despite the electromagnetic repulsion of the two protons. A "strong" nuclear force, surpassing the electromagnetic force in strength at nuclear length scale, had to be responsible for this. With the idea of the range of the interaction being inversely related to the rest mass of the force mediator, Yukawa argued that a quite heavy particle should be the mediator of the strong force.

Quantum theory demands that on a time scale Δt there is an uncertainty of measured energy according to

$$\Delta E\,\Delta t \geq \frac{\hbar}{2} \tag{2.2}$$

Since energy is related to mass, the rest mass of the mediator particle could be estimated from the time scale of its transmission. Using the proton radius r_p, a minimum value for this time scale is

$$\Delta t > \frac{r_p}{c} \approx 0.33 \times 10^{-23} \;\; [\mathrm{s}] \tag{2.3}$$

With m_π the mass of Yukawa's meson, the energy fluctuation is given as $\Delta E = m_\pi c^2$, and so

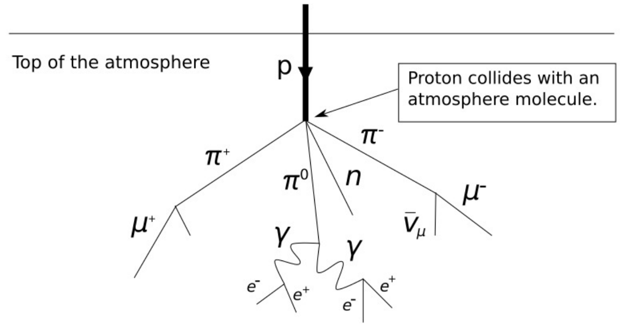
$$\begin{aligned} m_\pi c^2\,\Delta t &> \frac{\hbar}{2} \\ \Rightarrow m_\pi &> \frac{10^{-34}}{2 \times 9 \times 0.33 \times 10^{16} \times 10^{-23}} \;[\mathrm{S.I.}] \\ &\approx 0.185 \times 10^{-27} \;[\mathrm{S.I.}] \end{aligned}$$

which translates into an upper bound for the rest mass of the π meson

$$m_\pi > 104 \left[\frac{\mathrm{MeV}}{c^2}\right] \tag{2.4}$$

Today, the rest mass of the π mesons is known to be $\approx 135 \left[\frac{\mathrm{MeV}}{c^2}\right]$ which shows that Yukawa's estimate was quite good. Comparing this rest mass with the rest masses of electron and proton, $m_e \approx 0.51 \left[\frac{\mathrm{MeV}}{c^2}\right]$ and $m_p \approx 940 \left[\frac{\mathrm{MeV}}{c^2}\right]$, the term "middle weight", or meson, is explained.

The π was detected in 1947 through cosmic radiation.



2.1.3 Antimatter - Dirac Equation

The most active period concerning the detection of antimatter particles was between 1930 and 1956. What does a course on ‘relativity and nuclear physics’ have to do with antimatter? Well, the existence of antimatter particles was **predicted** by **Dirac** as a consequence of his famous equation, and it happens to be the equation of motion of nucleons (as well as of all other massive fermions, like electrons). So in my mind, a 3rd-year course on nuclear physics cannot get around the Dirac equation.

Let us, however, begin with the developments that lead to this equation. The first steps were taken by Oskar Klein and Walter Gordon.

2.1.3.1 Klein-Gordon Theory

2.1.3.1.1 Derivation of the Klein-Gordon Equation

Starting point is the time-dependent Schrödinger equation (SEQ):

$$\imath\hbar \frac{\partial}{\partial t} \Psi(\mathbf{x}, t) = \hat{H}\, \Psi(\mathbf{x}, t) \tag{2.5}$$

is a differential equation first order in time and second order in space and, therefore, cannot be Lorentz covariant. To understand this consider the one-dimensional SEQ for a free particle, slightly rewritten:

$$\hbar \left(\imath \frac{\partial}{\partial t} + \frac{\hbar}{2m} \frac{\partial^2}{\partial x^2} \right) \Psi(x, t) = 0 \tag{2.6}$$

We know that $\partial^\mu\partial_\mu$ is a Lorentz scalar, so $a\,\frac{\partial}{\partial t} + b\,\frac{\partial^2}{\partial x^2}$ with constant a, b cannot be one.

As a first step, Klein and Gordon took the time derivative of Eq. (2.5)

$$\imath\hbar\,\frac{\partial^2}{\partial t^2}\,\Psi(\mathbf{x},t) = \hat{H}\,\frac{\partial}{\partial t}\,\Psi(\mathbf{x},t). \tag{2.7}$$

Note that we are working with a time-independent Hamiltonian operator in the Schrödinger picture. After introduction of Eq. (2.5) in the form $\frac{\partial}{\partial t}\,\Psi(\mathbf{x},t) = \frac{1}{\imath\hbar}\,\hat{H}\,\Psi(\mathbf{x},t)$ into the right-hand side of Eq. (2.7) we get

$$\imath\hbar\,\frac{\partial^2}{\partial t^2}\,\Psi(\mathbf{x},t) = \frac{1}{\imath\hbar}\,\hat{H}^2\,\Psi(\mathbf{x},t). \tag{2.8}$$

This is so far not a deviation from non-relativistic quantum physics, and Eq. (2.8) is still not homogeneous in time and space derivatives when considering the non-relativistic energy-momentum relation[1].

Now, Klein and Gordon argued that Einstein's energy of a particle in relativistic mechanics is $E = \pm\sqrt{\mathbf{p}^{\,2}c^2 + m_0^2c^4}$. With the usual prescription for the momentum operator in quantum theory, $\mathbf{p} \longrightarrow -\imath\hbar\,\hat{\boldsymbol{\nabla}}$ from which $\hat{\mathbf{p}}^2 = -\hbar^2\,\boldsymbol{\nabla}^2$. Making the corresponding replacement in the Hamiltonian in Eq. (2.8) and considering rest energy simply as a multiplicative constant,

$$\begin{aligned} \left(\hbar^2\,\frac{\partial^2}{\partial t^2} - \hbar^2c^2\,\boldsymbol{\nabla}^2 + m_0^2c^4\right)\Psi(\mathbf{x},t) &= 0 \\ \left(\frac{1}{c^2}\,\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2 + \frac{m_0^2c^2}{\hbar^2}\right)\Psi(\mathbf{x},t) &= 0, \end{aligned} \tag{2.9}$$

which is equivalent to the procedure in non-relativistic quantum mechanics where the non-relativistic energy momentum relation is used at this point.

[1]One might ponder over using the square root directly in Eq. (2.5), but this results in operator roots and a mathematically quite complicated equation.

In Minkowski space with a metric tensor

$$\{g_{\mu\nu}\} = \{g^{\mu\nu}\} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2.10}$$

and real-valued coordinates, so $\{x^\mu\} := \begin{pmatrix} x_0 = ct \\ \mathbf{x} \end{pmatrix}$, the time derivative can be written like

$$\frac{1}{c}\frac{\partial}{\partial t} = \frac{1}{c}\frac{\partial x_0}{\partial t}\frac{\partial}{\partial x_0} = \frac{1}{c}\,c\,\frac{\partial}{\partial x_0} = \frac{\partial}{\partial x_0} =: \partial^0 \tag{2.11}$$

and thus Eq. (2.9) becomes

$$\begin{aligned} \left(\frac{\partial^2}{\partial x_0^2} - \sum_k \frac{\partial^2}{\partial x_k^2} + \frac{m_0^2 c^2}{\hbar^2}\right)\Psi(\mathbf{x},t) &= 0 \\ \left(\partial^\mu \partial_\mu + \frac{m_0^2 c^2}{\hbar^2}\right)\Psi(x) &= 0 \end{aligned} \tag{2.12}$$

the **Klein-Gordon equation** where in the last step the position four-vector x replaces the coordinates $\mathbf{x}, t$ in the argument of the wave function. The operator $\square = \partial^\mu \partial_\mu$ is called the **d'Alembertien**. The KG equation is manifestly Lorentz covariant, *i.e.*, it retains its form under Lorentz transformations. As can be shown straightforwardly, its solutions correspond to the correct relativistic energies of free particles of rest mass m_0. Note also that in the limit $m_0 \longrightarrow 0$ Eq. (2.12) yields the wave equation of electromagnetism.

2.1.3.1.2 Problems with Klein-Gordon Theory

It was, however, quickly realized that Klein and Gordon could not claim victory in having solved the problem of formulating the fundamental equation of relativistic quantum mechanics. The KG wave function

$\Psi_{\mathrm{KG}}(x)$ is a scalar field, but fermions have spin and a scalar wavefunction cannot describe two spin degrees of freedom, *i.e.*, spin projection "up" and spin projection "down". Perhaps particle spin could just be multiplied onto the wavefunction like in Schrödinger-Pauli theory. But this is certainly not satisfactory. The graver problem with Eq. (2.12) is, however, as follows.

In analogy to Schrödinger theory the conservation of the probability density is assured by it satisfying a continuity equation with an appropriate probability current density. For Klein-Gordon theory, this current density is

$$j^\mu := \frac{\imath\hbar}{2m_0}\left[\psi^*\left(\partial^\mu\psi\right) - \left(\partial^\mu\psi\right)^*\psi\right] \tag{2.13}$$

which shall be demonstrated.

Proof.

$$\partial_\mu j^\mu = \frac{\imath\hbar}{2m_0}\left[\left(\partial_\mu\psi^*\right)\left(\partial^\mu\psi\right) + \psi^*\left(\partial_\mu\partial^\mu\psi\right) - \left(\partial_\mu\partial^\mu\psi\right)^*\psi - \left(\partial^\mu\psi\right)^*\left(\partial_\mu\psi\right)\right] \tag{2.14}$$

The last term in Eq. (2.14) can be rewritten as

$$\begin{aligned} -\left(\partial^\mu\psi\right)^*\left(\partial_\mu\psi\right) &= -\left(\partial_\nu g^{\nu\mu}\psi\right)^*\left(\partial^\kappa g_{\kappa\mu}\psi\right) = -\left(\partial_\nu\psi\right)^* g^{\nu\mu}g_{\mu\kappa}\left(\partial^\kappa\psi\right) \\ &= -\left(\partial_\nu\psi\right)^*\delta^\nu_\kappa\left(\partial^\kappa\psi\right) = -\left(\partial_\nu\psi\right)^*\left(\partial^\nu\psi\right) \\ &= -\left(\partial_\nu\psi^*\right)\left(\partial^\nu\psi\right) \end{aligned} \tag{2.15}$$

and so cancels with the first term. Using the KG equation Eq. (2.12) in the second and third terms of Eq. (2.14) results in

$$\begin{aligned} \partial_\mu j^\mu &= \frac{\imath\hbar}{2m_0}\left[\psi^*\left(-\frac{m_0^2c^2}{\hbar^2}\psi\right) - \left(-\frac{m_0^2c^2}{\hbar^2}\psi^*\right)\psi\right] \\ &= \frac{\imath\hbar}{2m_0}\left[-\frac{m_0^2c^2}{\hbar^2} + \frac{m_0^2c^2}{\hbar^2}\right]\psi^*\psi = 0 \end{aligned} \tag{2.16}$$

which shows that the chosen probability current density in Eq. (2.13) indeed leads to a consistent KG theory. Since the probability current density four vector has the form given in Eq. (1.119) the KG probability density is the time-like component of Eq. (2.13)

$$\begin{aligned}\rho_{\mathrm{KG}} = \frac{j^0}{c} &= \frac{\imath\hbar}{2m_0 c}\left[\psi^*\left(\partial^0\psi\right) - \left(\partial^0\psi\right)^*\psi\right] \\ &= \frac{\imath\hbar}{2m_0 c}\left[\psi^*\left(\frac{1}{c}\frac{\partial}{\partial t}\psi\right) - \left(\frac{1}{c}\frac{\partial}{\partial t}\psi\right)^*\psi\right].\end{aligned} \tag{2.17}$$

The solutions of the KG equation are of plane-wave type and can be written as

$$\psi_{\mathrm{KG}}(x) = Ae^{\imath(\omega t - \mathbf{k}\cdot\mathbf{x})} + Be^{-\imath(\omega t - \mathbf{k}\cdot\mathbf{x})} \tag{2.18}$$

which when inserted into the KG equation yields the energy-momentum relation of relativistic theory. Using Eq. (2.18) in the representation of the KG density, Eq. (2.17), results in

$$\rho_{\mathrm{KG}} \propto -|A|^2 + |B|^2 \tag{2.19}$$

which is not hard to demonstrate. this means that the KG probability density could, depending on specific initial conditions which determine the coefficients A and B, become negative! This, however, is in stark contradiction with the fundamentals of quantum mechanics where probability density is always positive definite!

Historically, the KG equation was, therefore, regarded as a complete failure and discarded[2].

2.1.3.2 Dirac Theory

2.1.3.2.1 Derivation of the Dirac Equation

Dirac realized that the essential problem with the KG equation was the fact that it was a second-order differential equation. Thus, his idea was to formulate a Lorentz covariant **first-order** differential equation that treated space and time coordinates on an equal footing.

[2] As a little historical anecdote, a participant at one of the Solvay conferences in Brussels at the time is reported to have asked Dirac what he was currently working on. Dirac replied “I’m trying to construct a relativistic wave equation for the electron.” Participant: “But Klein and Gordon already solved that problem!” They had not, and Dirac knew full well.

He started out by introducing a set of general parameters $\{\gamma^\mu\}$, the properties of which had to be determined, to see if the square root of the d'Alembertien could be developed in such a way that it takes on the form of a Lorentz scalar[3]. By postulate,

$$\begin{aligned}
\Box = \partial_0\partial_0 - \sum_k \partial_k\partial_k \; &=: \; \gamma^\mu\gamma^\nu\,\partial_\mu\partial_\nu \\
&= \; \gamma^0\gamma^0\,\partial_0\partial_0 + \gamma^1\gamma^1\,\partial_1\partial_1 + \gamma^2\gamma^2\,\partial_2\partial_2 + \gamma^3\gamma^3\,\partial_3\partial_3 \\
&\quad + \left\{\gamma^0,\gamma^1\right\}\,\partial_0\partial_1 + \left\{\gamma^0,\gamma^2\right\}\,\partial_0\partial_2 + \left\{\gamma^0,\gamma^3\right\}\,\partial_0\partial_3 \\
&\quad + \left\{\gamma^1,\gamma^2\right\}\,\partial_1\partial_2 + \left\{\gamma^1,\gamma^3\right\}\,\partial_1\partial_3 + \left\{\gamma^2,\gamma^3\right\}\,\partial_2\partial_3
\end{aligned}$$

where $\{,\}$ is the anti-commutator. This would be correct if

$$\{\gamma^\mu,\gamma^\nu\} \; = \; 0 \qquad \forall\mu\neq\nu \tag{2.20}$$

$$\left\{\gamma^0,\gamma^0\right\} \; = \; 2 \tag{2.21}$$

$$\left\{\gamma^k,\gamma^k\right\} \; = \; -2 \qquad \forall k\in\{1\ldots 3\} \tag{2.22}$$

For example, since $\left\{\gamma^0,\gamma^0\right\} = 2(\gamma^0)^2 = 2 \;\Rightarrow \gamma^0 = 1$ which is in accord with the above decomposition. The quantities γ are today known as **Clifford numbers** and the equations (2.20), (2.21), (2.22) as the conditions for a **Clifford algebra**, in this case the Clifford algebra of Dirac theory.

Let's take a look at the first condition, Eq. (2.20). If γ^κ is just a complex scalar, $\gamma^\kappa \in \mathbb{C}$, then $\{\gamma^\mu,\gamma^\nu\} = 2\gamma^\mu\gamma^\nu = 0$, since scalars commute. This can only be solved by setting $\gamma^\mu = 0 \;\; \forall\mu$ (or likewise $\gamma^\nu = 0 \;\; \forall\nu$) which does not lead to a valid set of parameters. Dirac concluded that the set $\{\gamma^\mu_{n\times n}\}$ had to be $n\times n$ matrices in order to arrive at anticommutators that yield zero!

Dirac's first try was to use the set $\{\mathbf{1},\boldsymbol{\sigma}\}$ which form a basis for all complex 2×2 matrices. However, neither this choice leads to a valid

[3] Remember that $\Box = \partial^\mu\partial_\mu$ and so $\sqrt{\Box}\neq\partial^\mu$ and $\neq\partial_\mu$. $\Box$ is a Lorentz scalar, and so is $\sqrt{\Box}$, but ∂^μ and ∂_μ are not Lorentz scalars!

decomposition of the d'Alembertien[4].

Dirac found the simplest set to be

$$\gamma^0 := \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \qquad \gamma^k := \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_k \\ -\boldsymbol{\sigma}_k & \mathbf{0} \end{pmatrix} \tag{2.23}$$

where the spin-Pauli matrices have been introduced[5]. Now it is straightforward to rewrite the d'Alembertien as

$$\left(\partial_0\partial_0 - \sum_k \partial_k\partial_k\right) \mathbb{1}_4 = \gamma^\mu \gamma^\nu \partial_\mu\partial_\nu = \gamma^\mu\partial_\mu \gamma^\nu\partial_\nu = (\gamma^\mu\partial_\mu)^2 \tag{2.24}$$

and so

$$\sqrt{\Box}\, \mathbb{1}_4 = \gamma^\mu\partial_\mu \tag{2.25}$$

Note that we here have obtained $\Box$ as the product of two (formally) Lorentz scalars in a four-dimensional vector space. This is no longer the same thing as the Minkowski space scalar product $\partial_\mu\partial^\mu$ between the derivative four-vectors.

A direct consequence of this astonishing finding is that the wavefunction onto which such an operator acts cannot be a scalar function. It has to be a four-component vector-like quantity.

Dirac now formulated his famous equation. We first rewrite the KG equation as

$$\left(\hbar^2\partial^\mu\partial_\mu + m_0^2c^2\right)\Psi(x) = 0 \tag{2.26}$$

The square root of the first operator becomes in Dirac's formulation

$$\sqrt{\hbar^2\partial^\mu\partial_\mu}\, \mathbb{1}_4 = \hbar\gamma^\mu\partial_\mu \tag{2.27}$$

and so, with the four-momentum operator[6] $\hat{p}_\mu = \imath\hbar\partial_\mu$, we can formulate the **Dirac equation** for a free particle of rest mass m_0

[4]It can be shown in a general manner that only matrices with dimension multiples of 4 are possible solutions. So, for example, dimension 8 matrices can be constructed, but the resulting theory is identical to the dimension 4 theory in physical content.

[5]This set of matrices is called the "standard representation" of Dirac matrices. Other representations related through unitary transformations of the standard matrices are possible as well.

[6]Note that $\hat{p}^0 = \imath\hbar\frac{1}{c}\frac{\partial}{\partial t}$ and $p^0 = \frac{E}{c}$ which reproduces the correspondence principle from quantum mechanics for the energy operator $p^0c \longrightarrow \hat{p}^0c = \imath\hbar\frac{\partial}{\partial t}$. Likewise, $\hat{p}^k = -\imath\hbar\frac{\partial}{\partial x^k}$.

$$\boxed{(-\imath\hbar\,\gamma^\mu\,\partial_\mu + m_0 c\,\mathbb{1}_4)\ \underline{\Psi}(x) = \underline{0} \qquad (2.28)}$$

which is a matrix equation where the wavefunction becomes a **four-component spinor** that depends on the position four-vector x (which contains the time coordinate). We indeed have a first-order differential equation where space and time coordinates are treated on an equal footing, *i.e.*, in a four-vector.

Of course, we have to solve and interpret this equation and talk about what a spinor exactly is.

Using the standard representation the Dirac equation can be written out more explicitly as

$$\begin{aligned} \left\{-\imath\hbar\left[\gamma^0\partial_0+\gamma^1\partial_1+\gamma^2\partial_2+\gamma^3\partial_3\right]+m_0c\mathbb{1}_4\right\}\ \underline{\Psi}(x) &= \underline{0} \\ \left\{-\imath\hbar\left[\begin{pmatrix} \mathbb{1}_2 & 0_2 \\ 0_2 & -\mathbb{1}_2 \end{pmatrix}\frac{\partial}{\partial x_0}+\sum_{k=1}^{3}\begin{pmatrix} 0_2 & \boldsymbol{\sigma}_k \\ -\boldsymbol{\sigma}_k & 0_2 \end{pmatrix}\frac{\partial}{\partial x_k}\right]+m_0c\begin{pmatrix} \mathbb{1}_2 & 0_2 \\ 0_2 & \mathbb{1}_2 \end{pmatrix}\right\}\begin{pmatrix} \underline{\Psi}^U(x) \\ \underline{\Psi}^L(x) \end{pmatrix} &= \underline{0}. \quad (2.29) \end{aligned}$$

The second line follows from the realization that the γ matrices have a 2×2 block structure, and so the entire equation can be written in this so-called "bi-spinor" form. The associated 2-spinors are called "upper" ($\underline{\Psi}^U(x)$) and "lower" ($\underline{\Psi}^L(x)$) 2-spinors. Their significance will become clear when solutions of the free-particle Dirac equation are investigated.

Now, since in position-space representation $\mathbf{p} = -\imath\hbar\,\boldsymbol{\nabla}$ it is convenient to rewrite the term involving the spin-Pauli matrices using the scalar product between the 3-vector of the Pauli matrices and the 3-vector of momentum, $\boldsymbol{\sigma}\cdot\mathbf{p} = \boldsymbol{\sigma}_x\hat{p}_x + \boldsymbol{\sigma}_y\hat{p}_y + \boldsymbol{\sigma}_z\hat{p}_z$,[7] and the Dirac equation becomes

[7]Note, e.g., that the product of a Pauli matrix with a scalar momentum operator is well defined: $\boldsymbol{\sigma}_x\hat{p}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\hat{p}_x = \begin{pmatrix} 0 & \hat{p}_x \\ \hat{p}_x & 0 \end{pmatrix}$.

$$\begin{aligned}
\left[-\begin{pmatrix} \mathbb{1}_2 & 0_2 \\ 0_2 & -\mathbb{1}_2 \end{pmatrix} \imath\hbar\frac{\partial}{\partial t} + c\begin{pmatrix} 0_2 & \boldsymbol{\sigma}\cdot\mathbf{p} \\ -\boldsymbol{\sigma}\cdot\mathbf{p} & 0_2 \end{pmatrix} + m_0c^2\begin{pmatrix} \mathbb{1}_2 & 0_2 \\ 0_2 & \mathbb{1}_2 \end{pmatrix}\right]\begin{pmatrix} \underline{\Psi}^U(x) \\ \underline{\Psi}^L(x) \end{pmatrix} &= \underline{0} \\
\left[-\begin{pmatrix} \mathbb{1}_2 & 0_2 \\ 0_2 & \mathbb{1}_2 \end{pmatrix} \imath\hbar\frac{\partial}{\partial t} + c\begin{pmatrix} 0_2 & \boldsymbol{\sigma}\cdot\mathbf{p} \\ \boldsymbol{\sigma}\cdot\mathbf{p} & 0_2 \end{pmatrix} + m_0c^2\begin{pmatrix} \mathbb{1}_2 & 0_2 \\ 0_2 & -\mathbb{1}_2 \end{pmatrix}\right]\begin{pmatrix} \underline{\Psi}^U(x) \\ \underline{\Psi}^L(x) \end{pmatrix} &= \underline{0} \quad (2.30)
\end{aligned}$$

where the whole equation has been multiplied first by c and then from the left by $\begin{pmatrix} \mathbb{1}_2 & 0_2 \\ 0_2 & -\mathbb{1}_2 \end{pmatrix}$.

2.1.3.2.2 Dirac Equation for Stationary States

We now focus on stationary states and separate off the time-dependence in the usual way:

$$\underline{\Psi}(x) = \underline{\Psi}(\mathbf{x})\,\Psi(t) = \underline{\Psi}(\mathbf{x})\,e^{-\frac{\imath}{\hbar}Et} \tag{2.31}$$

which yields, considering that $-\imath\hbar\frac{\partial}{\partial t}\,e^{-\frac{\imath}{\hbar}Et} = -E\,e^{-\frac{\imath}{\hbar}Et}$,

$$\left[c\begin{pmatrix} 0_2 & \boldsymbol{\sigma}\cdot\mathbf{p} \\ \boldsymbol{\sigma}\cdot\mathbf{p} & 0_2 \end{pmatrix} + \begin{pmatrix} m_0c^2\mathbb{1}_2 & 0_2 \\ 0_2 & -m_0c^2\mathbb{1}_2 \end{pmatrix}\right]\underline{\Psi}(\mathbf{x}) = E\mathbb{1}_4\,\underline{\Psi}(\mathbf{x}) \tag{2.32}$$

The Dirac equation has been introduced as a relativistic covariant equation of motion for massive fermions of spin $s = \frac{1}{2}$.

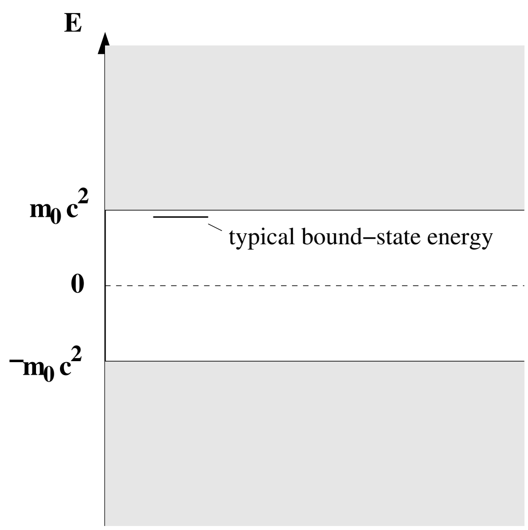
- Relativistic covariant wave equation that treats spatial and time variables on equal footing.
- Correct relativistic energy eigenvalues of the free particle, $E = \pm\sqrt{\mathbf{p}^2c^2 + m_0^2c^4}$
- Positive definite probability density, $\rho_D > 0$

2.1.3.2.3 Interpretation of the Dirac Equation

As will be shown, the energy eigenvalues of the free fermion according to Dirac theory are **found** as

$$E = \pm\sqrt{\mathbf{p}^2c^2 + m_0^2c^4}. \tag{2.33}$$

The four-dimensional spinor space allows for four solutions, two of which correspond to positive and two of which correspond to negative energy. **The negative energies that only appeared to be a vague possibility earlier are confirmed to be physical reality in Dirac theory**. They are a **mathematical consequence** of the introduction of a first-order Lorentz-covariant differential equation.



The spectrum of the free fermion according to Dirac theory. As in non-relativistic QM the possible energies are continuous (gray zones), but here for $||\mathbf{p}|| = 0$ we have the rest energy of the particle, m_0c^2, or $-m_0c^2$ for the branch of negative energies.

The so-called "Dirac gap" (impossible energies for a free particle) is defined as $E_{\text{gap}} = m_0c^2 - (-m_0c^2) = 2m_0c^2$. Dirac originally formulated the equation for an electron, so let's take $m_0 = m_e$. Then $E_{\text{gap,e}} \approx 1.02$ [MeV]. The total energy (kinetic + potential) of an electron bound in the potential of a proton is -13.61 [eV] (non-relativistic ground state), about 5 orders of magnitude smaller than $E_{\text{gap,e}}$! This is a negative energy for a bound state in non-relativistic theory. In Dirac theory we have to add the rest energy of the electron, and a resulting bound-state energy[8] is indicated in the above figure (not to scale; to scale it would have to be much closer to the rest energy bar.).

Like in non-relativistic theory of the hydrogen atom there is a discrete spectrum of bound states and a continuous spectrum of scattering states. In relativistic theory these have $E \geq m_0c^2$, **and** in addition there is a continuous spectrum of negative-energy states! Is this a problem? It is.

For if the "ground state" is no longer the true ground state, i.e., there exist states of lower energy, then the following decay process is quantum-mechanically allowed:

$$p + e \longrightarrow p + e + \gamma \tag{2.34}$$

The electron could, under emission of a photon of very high energy, transfer to a state of negative energy. This could go on and on, until the hydrogen atom has lost all its energy into radiation. Matter would no longer be stable[9].

Dirac's "Sea" and "Hole theory"

In response to the devastating property of his new equation that predicted that matter should radiate and be unstable, Dirac came up with the following solution: He **postulated** that **all states of negative energy** should be **occupied by the same type of fermions in vacuum**, and this postulate became known as **Dirac sea**. This sea should be perfectly homogeneous and have (among others) the following prop-

[8]Of course this result can be obtained by solving the Dirac equation for an electron bound to a proton, but this is a long way to go (much too long for the present course), so I only talk about the result here. In any case, the velocity of an electron in the hydrogen atom is relatively small, so relativistic corrections are small as well. The relativistic total energy – modulo rest energy – is very close to the non-relativistic total energy.

[9]At the time, Dirac's new theory drew fierce criticism from great contemporary physicists. Werner Heisenberg pounded "Dirac's theory is surely the saddest chapter of modern physics!"

erties:

$$\begin{aligned} m_{\text{sea}} &= +\infty \\ Q_{\text{sea}} &= -\infty \quad \text{(for electrons)} \\ E_{\text{sea}} &= -\infty \end{aligned}$$

Due to its homogeneity no charged particles immersed into it would experience the presence of the sea. And since we ever only measure energy differences in physical processes, never absolute energies, the total infinite energy was not a problem, either. The sea can thus be regarded as a background that has just the right properties to make matter stable.

For it was known that no two identical fermions can occupy the same microstate (Pauli exclusion principle). This means that an electron in any state of positive energy (including its rest energy) could no longer decay to a state of negative energy since all of those were already occupied.

More than this, Dirac postulated as a **direct consequence of the existence of the sea**: If say photons were produced of total energy $E_\gamma > E_{\text{gap,e}}$ the following process could occur:

$$2\gamma \longrightarrow e + \text{“hole''} \tag{2.35}$$

The radiation quanta could “kick” an electron out of the non-observable sea, i.e., excite it to a positive energy and in addition create a hole in the sea[10]. Since the hole corresponds to a “missing electron”, its properties must be – according to Dirac –

$$\begin{aligned} m_{\text{hole}} &= m_e \\ Q_{\text{hole}} &= -Q_e \\ E_{\text{hole}} &= -E_e \end{aligned}$$

[10] These radiation quanta could even be produced at very short time scales according to $\Delta E\, \Delta t \geq \frac{\hbar}{2}$, fluctuations of the vacuum that polarize it; a fundamental idea of quantum field theory was born. However, it took another 30 years for this theory to be fully developed and fleshed out.

A hole in the sea has the same inertia as a particle in vacuum, a **missing** charge corresponds to the negative particle charge, and the **missing** energy corresponds to the negative particle energy. Dirac's interpretation was that the hole in the sea had to represent a new particle of equal mass as the electron but of opposite charge (and of positive energy, since it represented the missing of negative energy).

This was one of the boldest and also one of the most spectacular predictions made in science. The particle representing the hole, the "positron", as it was called, was found in cosmic radiation by Anderson and Blackett in 1931, five years after Dirac's prediction of its existence. Since this prediction was not restricted to electrons and any other fermion could replace the electron in the argument, the finding led to the prediction of **antimatter**, i.e., that every type of particle should have a partner with identical mass, but opposite charge. Some particles should therefore be their own antiparticles, like the photon. This concept was later extended to the more general principle of "charge conjugation", $\hat{C}$.

2.1.4 Neutrinos

The basic discoveries in neutrino physics were made between 1930 and 1962. The fundamental observation concerns **nuclear** β **decay**, where β stands for an electron. At the time, the β decay of an atomic nucleus A into an atomic nucleus B

$$A \longrightarrow B + e \tag{2.36}$$

was understood in terms of the **fundamental process**

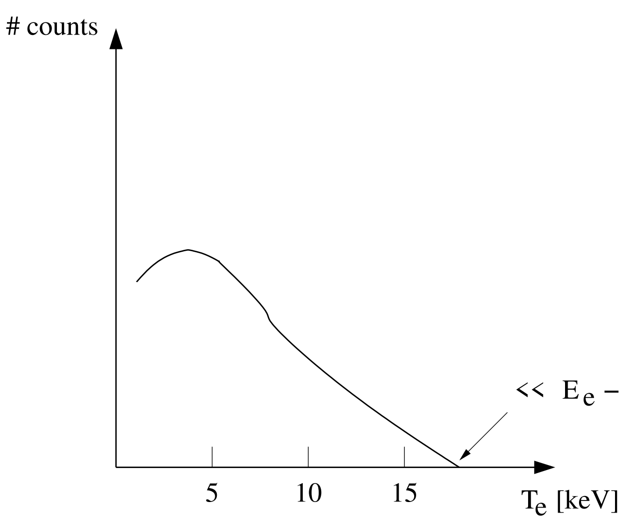
$$n \longrightarrow p + e \tag{2.37}$$

i.e., a neutron of nucleus A decays into a proton (a bound proton, so we obtain a new nucleus) and an electron. With the techniques developed earlier, we are in the position to calculate the energy of the emitted electron, see section 1.7.4. The result of the calculation is, at nuclear level,

$$E_e = \frac{m_A^2 - m_B^2 + m_e^2}{2m_A} c^2 \tag{2.38}$$

Note that this is a fixed value in terms of constants, just like the energy of the emitted muon in section 1.7.4 was. The difference is that here we have two emitted massive particles instead of just one.

Now this theoretical result can be compared to the actual observation in experiment.



Observed kinetic energy of the emitted electron in the β decay of tritium ${}^{A=3}_{Z=1}\mathrm{H} \longrightarrow {}^{3}_{2}\mathrm{He} + e$.

A here is the nucleon number and *Z* is the proton number.

In almost all events the electron's energy is **lower** than the limiting energy calculated *via* Eq. (2.38).

This meant that if conservation of energy should remain valid there is **energy unaccounted for** in the above β decay[11]. Pauli proposed that an unknown additional emitted particle with charge $Q = 0$ should account for the missing energy. Fermi figured out that this new particle must have zero rest mass, and thus the **neutrino** was born. So the

[11]When confronted with this, Niels Bohr thought that the conservation of energy should be abandoned! However, Bohr was opposed to *many* things at the time, not only to Fermi's neutrino, but also to Dirac's theory, Yukawa's meson, and even Feynman's approach to Quantum Field Theory ...

correct fundamental process can be written as

$$n \longrightarrow p + e + \overline{\nu} \tag{2.39}$$

In fact, it was only later realized that it had to be an *anti*neutrino that is produced here[12].

In the course of these discoveries a general rule was established for particle physics processes:

Crossing symmetry: If a certain reaction is observed then **crossed reactions** are also possible, where **crossed** means that a particle is placed on the other side of the reaction and conjugated into its antiparticle.

For example, a crossed reaction of the fundamental process Eq. (2.39) would be

$$p + \overline{\nu} \longrightarrow n + e^+ \tag{2.40}$$

An electron is here "crossed" into a positron. Cowan and Reynes observed this process in 1955 with solar antineutrinos, and they detected neutron and positron (e^+) formation in the reaction.

In 1953 Konopinski and Mahmoud established the conservation law $L = L'$ of **lepton number**, L, in particle reactions. This can be regarded as an equivalent to charge conservation, $Q = Q'$. A brief survey of lepton numbers:

L	particle type
0	all hadrons
$+1$	e^-, μ^-, ν
-1	e^+, μ^+, $\overline{\nu}$

Thus, lepton number changes sign when a particle is converted into its antiparticle. In fact, charge conjugation affects all quantum numbers but does not change momentum or energy. Lepton number conservation can be easily confirmed in all of the above processes.

[12]Neutrinos and their antiparticles have spin and differ in **helicity**, a concept we might talk about later. Also, conservation laws that were found later on dictated that it had to be an antineutrino of the first generation.

Further confirmations of L conservation followed. Two crossed reactions with respect to Eq. (2.39) are

$$\begin{aligned} p^+ + \overline{\nu} &\longrightarrow n + e^+ \\ \nu + n &\longrightarrow p^+ + e^- \end{aligned} \tag{2.41}$$

both of which were observed ($L = L' = -1$ in the first case and $L = L' = +1$ in the second case). On the other hand,

$$\overline{\nu} + n \longrightarrow p^+ + e^- \tag{2.42}$$

which would violate lepton number was never observed.

Unfortunately, this was not the end of the story for the leptons. The decay of the muon according to

$$\mu^- \longrightarrow e^- + \gamma \tag{2.43}$$

is kinematically allowed (the muon is heavier than the electron) and conserves lepton number, but this decay was **never observed**. It was proposed to introduce a conservation law that distinguishes between the three generations, generation 1: electron e^-, generation 2: muon μ^-, generation 3: tau τ^-, i.e., generational lepton numbers L_e, L_μ, L_τ. Then Eq. (2.43) would be forbidden since $L_\mu = 1 \neq L'_\mu = 0$ and $L_e = 0 \neq L'_e = 1$. Using a huge amount of antineutrinos produced in pion decays and testing their reactions with protons, in was in 1962 established that

$$\overline{\nu}_\mu + p^+ \longrightarrow \mu^+ + n \tag{2.44}$$

with $L_\mu = L'_\mu = -1$ and $L_e = L'_e = 0$ takes place whereas

$$\overline{\nu}_\mu + p^+ \longrightarrow e^+ + n \tag{2.45}$$

with $L_\mu = -1 \neq L'_\mu = 0$ and $L_e = 0 \neq L'_e = -1$ **never** does. The true decay channels of the muon and its antiparticle are

$$\mu^- \longrightarrow e^- + \overline{\nu}_e + \nu_\mu \tag{2.46}$$

$$\mu^+ \longrightarrow e^+ + \nu_e + \overline{\nu}_\mu \tag{2.47}$$

where L_μ and L_e are both conserved. We conclude on the first two generations of the lepton family (1962 - 1976) with a summary of their quantum numbers:

lepton	L	L_e	L_μ
e^-	1	1	0
ν_e	1	1	0
μ^-	1	0	1
ν_μ	1	0	1
antilepton	L	L_e	L_μ
e^+	-1	-1	0
$\overline{\nu}_e$	-1	-1	0
μ^+	-1	0	-1
$\overline{\nu}_\mu$	-1	0	-1

2.1.5 Flavor

2.1.5.1 Strangeness and Baryon Number

Between 1947 and 1960 more new hadrons entered the scene, and their observed behavior allowed for an extension of the conservation laws known thus far. The heavy meson K^0 (composed of a linear combination of a strange s and an antidown $\overline{d}$ quark and vice versa) and the baryon $\Lambda(uds)$ **decay under weak interaction** as follows:

$$K^0 \longrightarrow \pi^+ + \pi^- \tag{2.48}$$

$$\Lambda \longrightarrow p^+ + \pi^- \tag{2.49}$$

The K^0 has meson number $+1$, just like the π^+. So the π^- has meson number -1 since it is the antiparticle of the π^+. This means that **meson number is generally not conserved**. The same conclusion can be drawn from the decay of the Λ. On the left-hand side meson number is 0, but on the right-hand side meson number is -1. These new relatively heavy mesons were called "strange" particles (the quark decomposition became known only later!), mainly because their creation – driven by the strong interaction – is a relatively fast process, but their decay – driven by the weak interaction – is relatively slow; the difference is orders of magnitude.

On the other hand, **baryon number, B, is conserved**[13]. The conservation of B can also be verified on Eq. (2.39). Some important baryon numbers:

B	particle type
0	all leptons, all mesons
$+1$	p^+, n, Λ
-1	p^-, $\overline{n}$

The antiproton $p^- \equiv \overline{p}$ was first produced in the following inelastic collision:

$$p^+ + p^+ \longrightarrow p^+ + p^+ + p^+ + p^- \tag{2.50}$$

Note that this is a "sticky" relativistic collision. Conservation of Q (total charge), B, L_e and L_μ are easily verified.

Ongoing investigations and results and the early days of the quark model affirmed that a new quantum number could be introduced, called "strangeness" (S), that was **conserved** in processes driven by the **strong interaction**, but **not conserved** in processes driven by the

[13] B is **almost** always conserved in particle processes. It took another while to find a rare exception which is connected to charge-parity (CP) violation (the K^0 meson decays in 0.2% of events under violation of CP.). In order to formalize this we need to understand how to write the operators $\hat{C}$ (charge conjugation) and $\hat{P}$ (space inversion) in the framework of Dirac theory. This is the subject of more advanced chapters.

weak interaction. Examples:

$$\pi^-(d\overline{u}) + p^+(uud) \longrightarrow K^+(u\overline{s}) + \Sigma^-(dds) \tag{2.51}$$

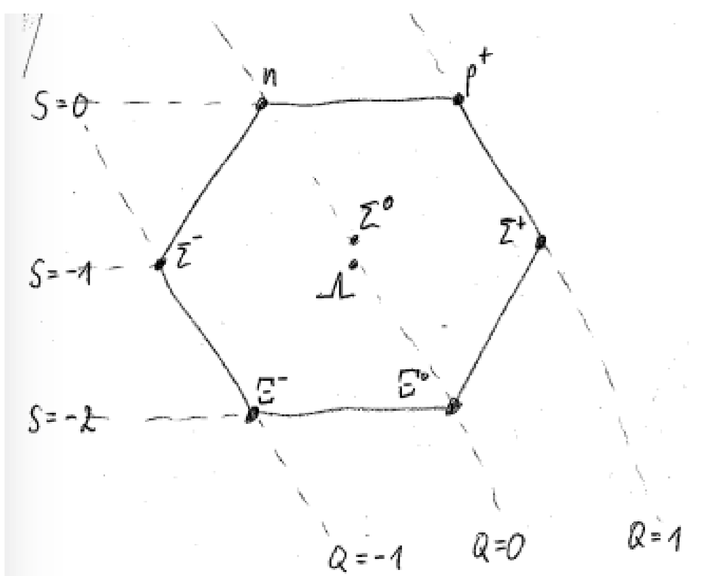
This is a strong-interaction process. We observe $B = B' = +1$, $Q = Q' = 0$, and $S = 0 + 0 = S' = 1 + (-1)$. The strange quark, s, was given $S = -1$ and its antipartner $\overline{s}$ has $S = +1$. So strangeness is conserved.

Now consider again the weak decay in Eq. (2.49). $S = -1 \neq S' = 0$ and strangeness is not conserved.

2.1.5.2 The Eightfold Way

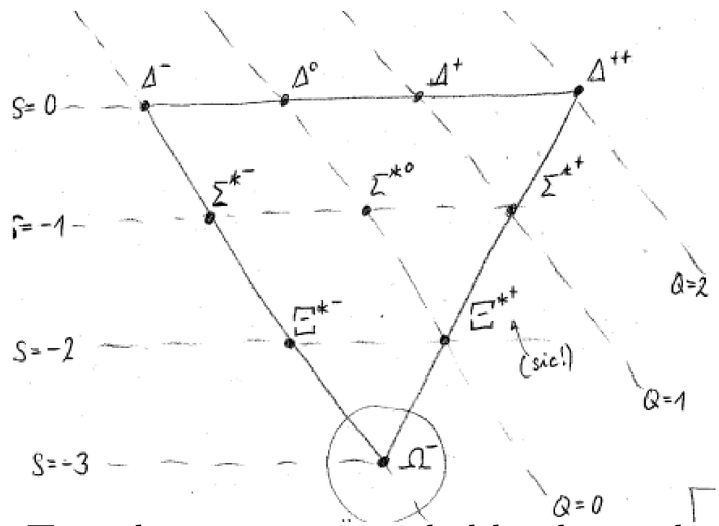
The situation having become ever more chaotic, Gell-Mann and Ne'eman in the period of 1961 - 1964 invented an ordering scheme called the "**Eightfold Way**" that not only helped understand particle phenomenology but that also made successful predictions of so far unknown particles!

Gell-Mann and Ne'eman realized that the **eight lightest baryons** could be organized **into an octet**, according to charge and strangeness.



The octet of light baryons. The quark decomposition for the "new" baryons is $\Theta^-(dss)$, $\Theta^0(uss)$, $\Sigma^0(uds)$, $\Sigma^+(uus)$.

Likewise, the ten next heavier baryons form a decuplet where isoaxes of charge and strangeness are the same as in the octet scheme.



The decuplet of heavier baryons. The quark decomposition of these baryons is $\Delta^-(ddd)$, $\Delta^0(udd)$, $\Delta^+(uud)$, $\Delta^{++}(uuu)$, $\Sigma^{*-}(dds)$, $\Sigma^{*0}(uds)$, $\Sigma^{*+}(uus)$, $\Theta^{*-}(dss)$, $\Theta^{*0}(uss)$, $\Omega^-(sss)$.

Two things are remarkable about the baryon decuplet diagram. The first is that the Ω^- forming the lower corner was not known at the time of its making. It was a prediction that was shortly afterwards confirmed!

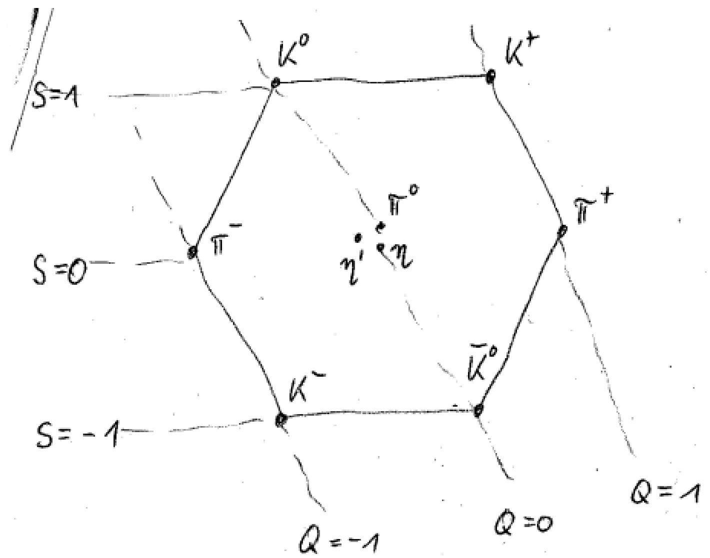
Second, some of these baryons have the same quark decomposition as the lighter baryons, for example $\Theta^{*-}(dss)$ and $\Theta^-(dss)$. The difference is that in the Θ^{*-} the three quarks are **confined** in an **excited state** which is denoted by the asterisk $(*)$. So we would expect them, according to Eq. (1.138), to have **different rest mass**. Indeed,

$$
\begin{aligned}
m_{\Theta^{*-}} &= 1533 \left[\frac{\text{MeV}}{c^2}\right]; \quad S = 3/2 \\
m_{\Theta^-} &= 1321 \left[\frac{\text{MeV}}{c^2}\right]; \quad S = 1/2
\end{aligned}
$$

In addition, they have different spin quantum numbers S. Similarly, the proton $p^+(uud)$ has rest mass $m_{p^+} = 938\ \left[\frac{\text{MeV}}{c^2}\right]$ whereas the excited $\Delta^+(uud)$ has $m_{\Delta^+} = 1232\ \left[\frac{\text{MeV}}{c^2}\right]$.

This raises an important question: When do we consider an excited-state particle as a **different particle**? We might compare the situation with atomic physics and ask whether an excited hydrogen atom, H^*, should not be regarded as a **different particle**, too, compared to

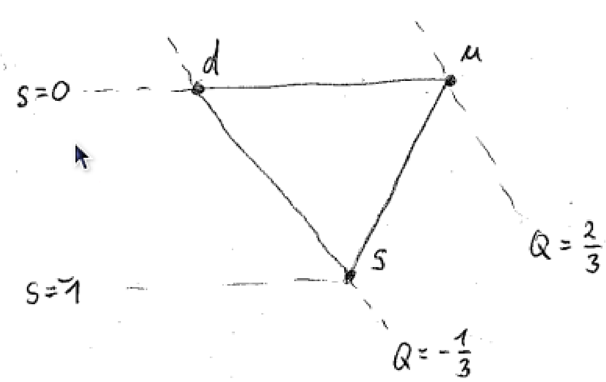
H, since it has higher rest mass than H. However, typical atomic excitation energies are on the order of [eV], and the rest energy of the proton is $E_{0,p^+} = 938$ [MeV]. This is a difference of about 9 orders of magnitude! In the above baryons, on the other hand, rest energy and excitation energy are in the same order of magnitude, $O(E_0) \approx O(E^*)$. This is why we here speak of a different particle whereas for excited atoms we do not.

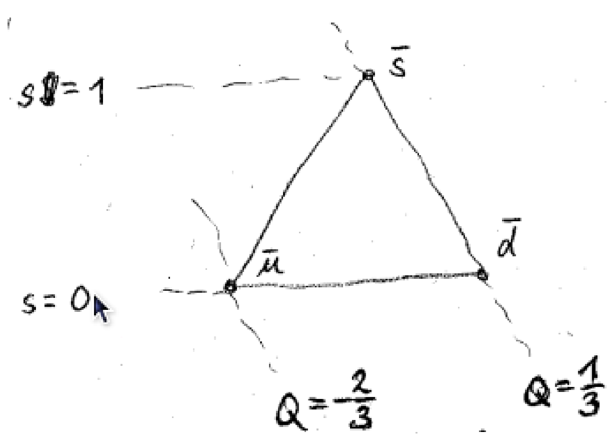


Finally, the lightest mesons are organized into a **meson nonet**. The η particles are linear combinations of $(u\overline{u})$, $(d\overline{d})$, and $(s\overline{s})$ states, the π^0 of $(u\overline{u})$ and $(d\overline{d})$ states.

2.1.5.3 Quark model and Eightfold Way

In 1964 Murray Gell-Mann and George Zweig introduced the solution to the question as of how the above Eightfold-Way diagrams **emerge** from a deeper, underlying structure.





Three quarks and their **anti**partners, organized as above into triangular diagrams, can account for the observed bound states of baryons and mesons[14]. In essence, the quark model conjectures that

1. All (anti)baryons are composed of 3 (anti)quarks.

2. All mesons are composed of one quark and one antiquark.

Quarks are **confined**[15] into bound states and are never observed individually.

2.1.5.4 Problems With the Quark model – and Some Solutions

The first problem with the quark model is that they are not detectable individually. However, scattering experiments on baryons show (similar to the early scattering experiments by Rutherford on atoms) that there are three localized mass "concentrations" in a baryon, pointing to the quark decomposition.

The second problem becomes obvious when thinking about the existence of, for example, the $\Delta^{++}(uuu)$ particle (see the baryon decuplet diagram). It has total spin $S = \frac{3}{2}$ and is composed of three identical particles, u quarks, which each have $s = \frac{1}{2}$. This means that at least two of these quarks must have identical quantum numbers, $m_s(i) = \frac{1}{2}$, which, however, is forbidden by the Pauli exclusion principle!

In 1964 Greenberg proposed a solution to this problem. First of all, particles had been given the following **"flavor"** quantum numbers:

[14]There is quite a bit more to be said here, for example why the corners of the baryon octet diagram are "missing" compared to the baryon decuplet diagram. For answering this we would have to analyze the irreducible representations of color SU(3) $\otimes$ spin SU(2).

[15]Just like us, these days ...

Quantum number	Name	Example
U	"upness"	u, $U = +1$
D	"downness"	d, $D = +1$, $U = 0$
S	"strangeness"	s, $S = -1$, $U = 0$, etc.
C	"charm"	c, $C = +1$, $U = 0$, etc.
B	"bottomness"	b, $B = -1$, $U = 0$, etc.
T	"topness"	t, $T = +1$, $U = 0$, etc.

which add up in composed particles, so $U = 0$ for the π^0 meson. In addition to flavor, Greenberg hypothesized that quarks come in "colors", i.e., "red", "green", and "blue". These are additional properties that are chosen in analogy with the property of light being colorless when all three wavelengths are mixed. Likewise, all occurring particles are "colorless". This solves the problem: If the three u quarks in the $\Delta^{++}(uuu)$ have three different colors, then no two of them can have the same total set of quantum numbers. For consistency, the mesons are colorless, too. For example, the π^0 $(u\overline{u}, d\overline{d})$ is colorless when the quark-antiquark pair has the same color. "red" + "antired" $\longrightarrow$ colorless, likewise for blue and green pairs.

This may sound very daring, but Greenberg's idea of course makes observable predictions. For example, a corrolary is that bound states consisting of two quarks (or two antiquarks) should not exist, because they would not be colorless. Likewise for bound states consisting of four quarks or four antiquarks. Indeed, no such particles have been observed[16].

The introduction of flavor and color made further predictions that were later on spectacularly confirmed, for example with the discovery of the Ψ $(c\overline{c})$ and D^+ $(c\overline{d})$ particles in 1974 and 1976.

[16]In a very recent discovery, however, tetraquarks, composed of two quark-antiquark pairs – for example the bound system $(c\overline{c}s\overline{s})$ – have been detected at CERN (arXiv:2103.01803 [hep-ex]).

2.1.5.5 Standard Model

I want to conclude this chapter with a summary of the known particles composing the Standard Model of elementary particles.

Leptons		Quarks		Force mediators	Higgs
e^-	e^+	d	$\overline{d}$	γ (photon, EM)	H
ν_e	$\overline{\nu}_e$	u	$\overline{u}$	$W^\pm$, Z^0 (vector bosons, weak)	
μ^-	μ^+	s	$\overline{s}$	8 gluons (strong)	
ν_μ	$\overline{\nu}_\mu$	c	$\overline{c}$		
τ^-	τ^+	b	$\overline{b}$		
ν_τ	$\overline{\nu}_\tau$	t	$\overline{t}$		

There are 37 particles altogether. The 8 gluons are classified according to their color transmission ($r\overline{r}$, $r\overline{g}$, $r\overline{b}$, $g\overline{g}$, $g\overline{b}$, $g\overline{r}$, $b\overline{b}$, $b\overline{r}$, $b\overline{g}$). The dimension of this set is 9, and they span two irreducible representations of color $SU(3)$, $\dim(\Gamma^1) = 1$ and $\dim(\Gamma^2) = 8$. The totally symmetric Γ^1 representation is not observed in nature, leaving 8 linear combinations of color transmissions.

Chapter 3

Introduction to Nuclear Physics

This final chapter gives an introduction to some important aspects of nuclear physics. Various sources have been used, among them Griffith's book (see above), the monograph by Povh, Rith, Scholz, Zetsche, Rodejohann "Particles and Nuclei", and the interactive website `https://www.nndc.bnl.gov/nudat3/` of the Brookhaven National Laboratory that I highly recommend[1].

3.1 General Definitions

Let us begin with some terms and definitions. A **nuclide** (french: nucléide) is understood as a **bound state** composed of A nucleons (p^+, n) among which there are Z protons (p^+). This is a non-redundant and the standard definition, although sometimes the number of protons and neutrons is given. The information can be assembled into a **symbol for the nuclide**

$$ {}^{\mathbf{A}}_{\mathbf{Z}}\mathbf{X}^{\mathbf{(Q)}}_{\mathbf{N}} $$

X denotes an element of the **periodic table of elements** (H, He, Li, ...)
N is the number of neutrons, $N = A - Z$

Sometimes, when the nuclide is understood to be an atomic nucleus

[1] The images in the manuscript have been created with the preceding version nudat2.

and electrons are present in the bound state, the total charge $Q = Z - \#e^-$ of the system can be given as well, depending on context. For example, the α particle is denoted as the nuclide ${}^4_2\mathrm{He}_2$. As an example where it is useful to add the total charges, reconsider the earlier weak decay of tritium which in terms of its fundamental process was given by Eq. (2.39). In **nuclear notation** at nucleon level this can now be written as

$$ {}^3_1\mathrm{H}_2 \longrightarrow {}^3_2\mathrm{He}^+_1 + e^- + \overline{\nu}_e \tag{3.1} $$

A neutron of tritium has decayed into a proton (which remains bound), and so A does not change. However, the proton number has changed and so we consider the product to be a nucleus of helium, not hydrogen. ${}^3_1\mathrm{H}_2$ is electrically neutral and has 1 electron. So does ${}^3_2\mathrm{He}^+_1$, but now we have 2 protons and so $Q = +1$ (denoted $+$). Note also the conservation laws, $L_e = L'_e = 1$ and $B = B' = 3$.

Further definitions:

- **Isotopes** are nuclides which have the **same proton number**, so $Z = Z'$. ${}^4_2\mathrm{He}$ and ${}^3_2\mathrm{He}$ are, therefore, isotopes. They always share the same element symbol.
- **Isotones** are nuclides which have the **same neutron number**, so $N = N' = A - Z = A' - Z'$. ${}^4_2\mathrm{He}_{N=2}$ and ${}^3_1\mathrm{H}_{N=2}$ are, therefore, isotones.
- **Isobars** are nuclides which have the **same nucleon number**, so $A = A'$. ${}^3_2\mathrm{He}$ and ${}^3_1\mathrm{H}$ are, therefore, isobars.

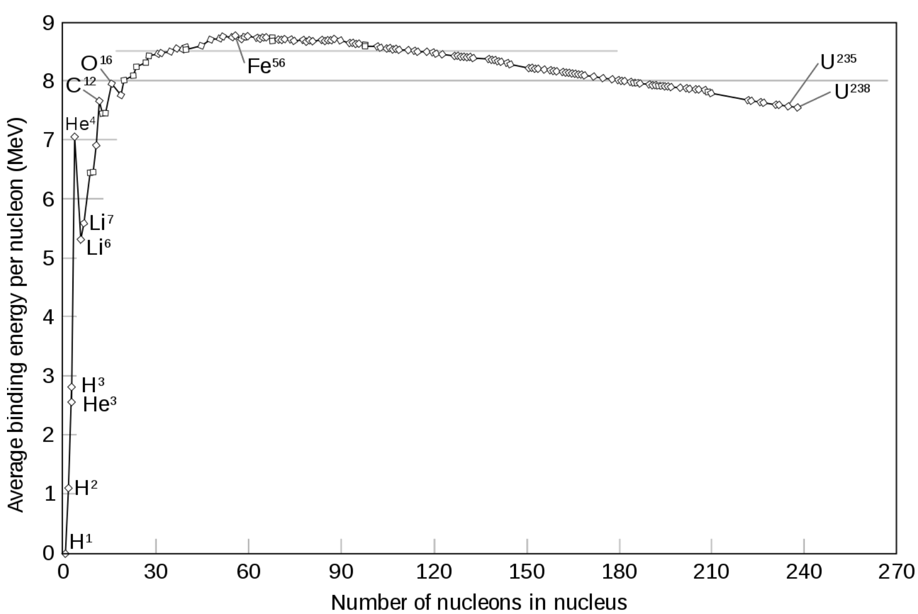
Another important quantity is the **nuclear binding energy**. It is defined as

$$ E_{\mathrm{bind}} = \left[Z m_{{}^1_1\mathrm{H}} + (A - Z) m_n - m_{{}^A_Z\mathrm{X}} \right] c^2 \tag{3.2} $$

where $m_{{}^1_1\mathrm{H}}$ is the rest mass of a hydrogen atom, m_n is the rest mass of a neutron, and $m_{{}^A_Z\mathrm{X}}$ is the rest mass of the neutral atom X under consideration[2].

Compare the nuclear binding energy Eq. (3.2) with the mass defect given in Eq. (1.159). If the mass defect is multiplied by c^2 then we can say that $m_{\text{sys}_{\text{before}}}c^2$ corresponds to the rest energy of the separated $Zm_{{}^1_1\text{H}}$ hydrogen atoms plus the rest energy of the separated $(A-Z)m_n$ neutrons, and $m_{\text{sys}_{\text{after}}}c^2$ corresponds to the rest energy of the bound nucleus composed of all these particles. In other words, the energy "lost", E'_{rad} in the formation of the nucleus **is** its binding energy.

E_{bind} is a positive quantity, $E_{\text{bind}} > 0$.

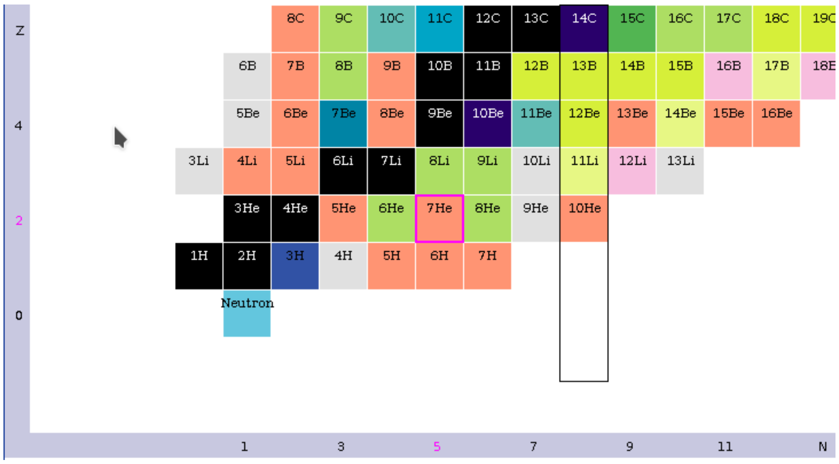


Nuclear binding energies per nucleon for common (abundant) isotopes of a given nuclide. The same trend as for the mass defect in Table 1.1 is observed.

3.2 Strong Isospin

A glance at a few of the simplest nuclides reveals an interesting fact. The following snapshot is from the aforementioned website at Brookhaven and shows the lower left corner of a (Z, N) diagram.

[2]Electron masses are generally included in this definition because neutral atoms are easier to "weigh" than ions and electrons are so light compared to nucleons.



Black nuclides are stable (very long lifetimes), colored nuclides decay in different ways (we will come back to that later). and an empty space means this nuclide has an unmeasurably short lifetime.

We see that ${}^2_1$H (p^+, n) is stable, but two protons or two neutrons do **not form a bound state**. Since there is no electrostatic repulsion between them it seemed particularly strange that two neutrons do not bind to each other. How can this be explained?

In a 1932 classic paper Werner Heisenberg came up with an explanation. It is based on the observation that the rest masses of the proton $(938.3 \left[\frac{\text{MeV}}{c^2}\right])$ and the neutron $(939.6 \left[\frac{\text{MeV}}{c^2}\right])$ are almost the same. So *via* Einstein's mass-energy equivalence, Eq. (1.138), the near-identity of their rest energies is a **near-degeneracy** of two different quantum-mechanical states.

Heisenberg proposed to write these two energies as E_{I,M_I} and $E_{I,M_I'}$ where I satisfies the algebra of an angular momentum and M_I is its projection onto the quantization axis. Now, we know the theorem that if $\left[\hat{H}, \hat{I}_k\right] = 0 \;\; \forall k \in \{1,2,3\}$ then $E_{I,M_I} = E_{I,M_I'}$. This in turn means that the Hamiltonian also commutes with a corresponding rotation, the generator of which is the operator $\hat{\mathbf{I}}$:

$$\left[\hat{H}, \hat{U}_I(\delta\varphi)\right] = 0 \tag{3.3}$$

where $\hat{U}_I(\delta\varphi) = e^{\frac{\imath}{\hbar}\delta\varphi\,\hat{\mathbf{e}}_n\cdot\hat{\mathbf{I}}}$.

The two nucleons can, therefore, be understood as two different quantum-mechanical microstates that are related by a rotation in a corresponding space. Since the number of microstates is two, Heisenberg conjectured that they form the **fundamental irreducible representation** of the Lie group $SU(2)$, called $\Gamma^{1/2}$, just like the spin of a fermion does. Since there is no additional angular momentum in the nucleon states but their degeneracy can be understood in terms of angular momentum algebra, I is called **isospin** (*like* spin), and $\hat{U}_I(\delta\varphi)$ is an infinitesimal rotation in the abstract **isospin space** which is the analog of spin space.

This has interesting consequences. First, if there is a symmetry (invariance of the Hamiltonian under isospin rotations) then there is a conservation law (*via* Noether's Theorem). The common force between an ensemble of protons and neutrons is the strong interaction, so this implies that

$$\left[\hat{H}_{\text{strong}}, \hat{U}_I(\delta\varphi)\right] = 0 \tag{3.4}$$

or in other words, the **strong interaction conserves isospin** *via* Heisenberg's equation of motion. This isospin is thus also called strong isospin.

This is a lot to swallow. But let's see if it helps us understand what is going on with bound states among protons and neutrons. Remember from the courses on symmetry that – without external fields – a spin 1/2 fermion has two degenerate states, $|m_s = 1/2\rangle$ and $|m_s = -1/2\rangle$ (that form the fundamental irreducible representation $\Gamma^{1/2}$). If we have two such particles, the spin eigenfunctions for the two-particle states represent a non-degenerate **singlet state** and a threefold-degenerate **triplet state**.

Now it is not hard to construct the isospin analog of the above finding. The general convention is to denote the proton and neutron as the following isospin states:

$$\left|p^+\right\rangle \equiv \left|I = \frac{1}{2}, M_I = \frac{1}{2}\right\rangle \tag{3.5}$$

$$\left|n\right\rangle \equiv \left|I = \frac{1}{2}, M_I = -\frac{1}{2}\right\rangle \tag{3.6}$$

By theoretical analogy, a system composed of **two** nucleons is then represented by one of the four possible isospin microstates:

$$\left|p^+\,p^+\right\rangle \equiv |I = 1, M_I = 1\rangle \tag{3.7}$$

$$\frac{1}{\sqrt{2}}\left[\left|p^+\,n\right\rangle + \left|n\,p^+\right\rangle\right] \equiv |I = 1, M_I = 0\rangle \tag{3.8}$$

$$\left|n\,n\right\rangle \equiv |I = 1, M_I = -1\rangle \tag{3.9}$$

This is the isospin triplet. Likewise,

$$\frac{1}{\sqrt{2}}\left[\left|p^+\,n\right\rangle - \left|n\,p^+\right\rangle\right] \equiv |I = 0, M_I = 0\rangle \tag{3.10}$$

which is the isospin singlet. From the conclusions on the theory of the Heisenberg Hamiltonian we know that if the interaction between the particles is repulsive (case of electrons and electromagnetism) then the spin triplet has lower energy. However, in the present case the respective interaction is **attractive** (strong interaction), and so we must conclude that here the **isospin singlet is more stable**[3]. And since the states $|p^+\,p^+\rangle$ and $|n\,n\rangle$ belong to the destabilized isospin triplet we have an explanation for the non-existence of such states!

Perhaps all of this seems like magic to you. After all, in order for the concept of strong isospin to be utterly convincing, it must manifest itself in more than just the proton-neutron states. Take a look again at the baryon octet diagram in the last chapter. The next four heavier

[3]The physical picture here is that the probability of finding two particles with the same isospin (projection) at the same point in space is zero, and so their mutual strong attraction is **reduced** on the average.

particles have rest masses

$$\begin{aligned} m_{\Sigma^+}(uus) &= 1189.4 \left[\frac{\mathrm{MeV}}{c^2}\right] \\ m_{\Sigma^0}(uds) &= 1192.5 \left[\frac{\mathrm{MeV}}{c^2}\right] \\ m_{\Sigma^-}(dds) &= 1197.3 \left[\frac{\mathrm{MeV}}{c^2}\right] \end{aligned}$$

which can be arranged into an almost fully degenerate **isospin triplet** which transforms as to the 3-dimensional *irrep* $\Gamma^{I=1}$ of isospin $SU(2)$. The fourth particle is

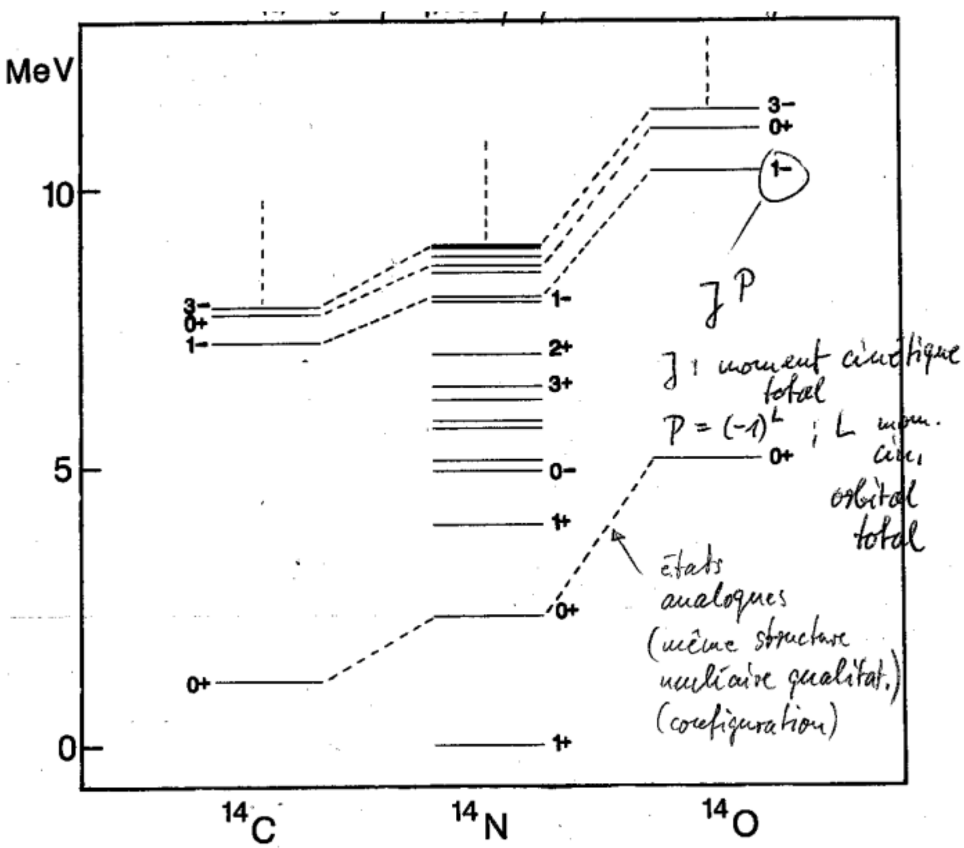
$$m_\Lambda(uds) = 1115.6 \left[\frac{\mathrm{MeV}}{c^2}\right]$$

which forms an **isospin singlet** and transforms as to the 1-dimensional *irrep* $\Gamma^{I=0}$ of isospin $SU(2)$.

Then we have an isospin doublet for Θ^0 and Θ^-, and the concept carries on for the mesons where the π particles form an isospin triplet, the η a singlet, and so on[4].

Still not convinced? Take a look at the energetically lower part of the spectrum of the following three isobaric nuclei:

[4] Isospin is a **broken symmetry**. It is very weakly broken for the lighter particles and the breaking becomes greater for the heavier particles, essentially due to mass differences between the quarks of the Standard Model. Furthermore, it is a symmetry under strong interaction but not under electromagnetic or weak interaction.



Many-nucleon states are classified as to total angular momentum J and parity P. The nuclear ground state of ${}^{14}_{7}$N is aligned with $E = 0$.

Obviously, the spectra of ${}^{14}_{6}$C and ${}^{14}_{8}$O are very similar, but they differ qualitatively from the spectrum of ${}^{14}_{7}$N where many more states are observed at low energies and the nuclear ground state is different.

Now consider the isospin of these nuclides. The projection quantum numbers can be calculated because we know the number of protons and neutrons in either case[5]:

$$\begin{aligned} M_I({}^{14}_{6}\mathrm{C}) &= 6 \times (+\frac{1}{2}) + 8 \times (-\frac{1}{2}) = -1 \\ M_I({}^{14}_{8}\mathrm{O}) &= 8 \times (+\frac{1}{2}) + 6 \times (-\frac{1}{2}) = +1 \\ M_I({}^{14}_{7}\mathrm{N}) &= 7 \times (+\frac{1}{2}) + 7 \times (-\frac{1}{2}) = 0 \end{aligned}$$

Since $M_I({}^{14}_{6}\mathrm{C}) = -1$ the lowest possible isospin quantum number for this nuclide is $I = 1$. This is because $M_I = 0$ does not exist here. So we can say that $I_{\min}({}^{14}_{6}\mathrm{C}) = 1$. Likewise, $I_{\min}({}^{14}_{8}\mathrm{O}) = 1$ but $I_{\min}({}^{14}_{7}\mathrm{N}) = 0$

[5]Remember that for coupling angular momenta $\hat{\mathbf{I}} = \sum_j \hat{\mathbf{I}}(j)$ vectorially and so $M_I = \sum_j M_I(j)$.

because in the latter case $M_I = 0$ exists.

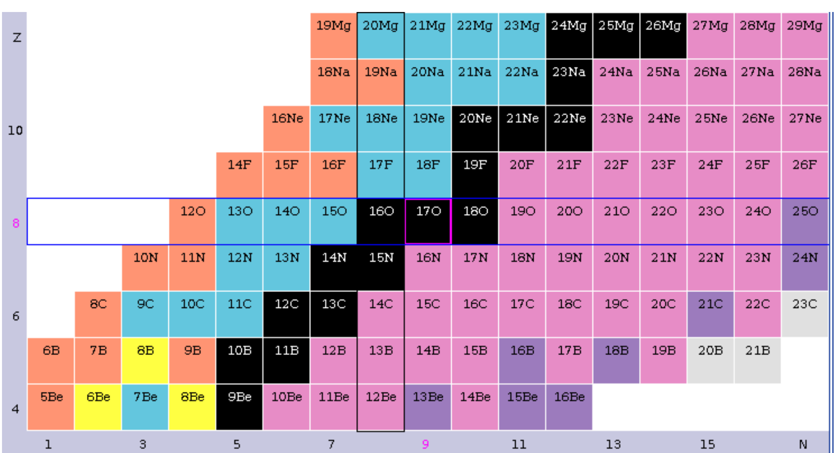
So we find that the nuclides differ qualitatively at the level of isospin. And since isospin is a symmetry of the strong interaction, the intrinsic interactions in the ${}^{14}_{7}\mathrm{N}$ nuclide differs from the other two nuclides[6].

In the bigger picture stable nuclei are therefore those where the number of protons equals the number of neutrons. For the lighter nuclei this general rule is very well fulfilled. As the nuclei become heavier, the mean electromagnetic Coulomb repulsion between the protons increases and gradually nuclei with growing neutron excess become the most stable isotopes.

3.3 Radioactive Decay

3.3.1 Decay Types

The following snapshot from the Brookhaven website is centered around a stable (black color) oxygen nuclide.



Nuclides around ${}^{17}_{8}\mathrm{O}$ display various ways of decaying, indicated by a color code.

All of the different decay modes are manifestations of **radioactive decay**. We will discuss them one by one.

■ Stable nuclide. Its **half-life**[7] is $> 10^{15}$ seconds.

[6]The exact reasons for the appearance of additional states is a matter of details.

[7]To be defined rigorously later.

As the neutron-to-proton ratio $\frac{N}{Z}$ increases (to the lower right) the nuclides become less stable. We are moving toward isospin states that are further removed from the stable isospin singlets (as discussed in the previous section). In this case the decay type is **beta (minus) decay** of one of the neutrons. In terms of the fundamental process it is denoted as

$$n \longrightarrow p^+ + e^- + \overline{\nu}_e \tag{3.11}$$

The β^- particle is synonymous with the electron. The proton typically remains bound in the nuclide.

If the neutron-to-proton ratio is increased even further, the nuclide decays by **neutron emission**. A neutron is ejected from the nuclide, written in nuclear notation:

$$^A_Z\mathrm{X}_N \longrightarrow {}^{A-1}_{Z}\mathrm{X}_{N-1} + {}^1_0\mathrm{n}_1 \tag{3.12}$$

The "new" nuclide $^{A-1}_{Z}\mathrm{X}_{N-1}$ is generally more stable than $^A_Z\mathrm{X}_N$, but it can further disintegrate *via* beta (minus) decay.

Unknown decay mode.

As the neutron-to-proton ratio $\frac{N}{Z}$ decreases (to the upper left) the nuclides decay *via* **beta (plus) decay** of one of the protons. As a fundamental process on its own, this would not happen because the rest mass of the neutron is greater than that of the proton (review the discussion in subsection 1.7.3). However, the nucleus can provide the required energy if the product nuclide has greater binding energy than the original nuclide. In nuclear notation $\beta^+(e^+)$ decay is written as

$$^A_Z\mathrm{X}_N \longrightarrow {}^{A}_{Z-1}\mathrm{Y}_{N+1} + e^+ + \nu_e \tag{3.13}$$

The underlying fundamental process is

$$p^+ \longrightarrow n + e^+ + \nu_e \tag{3.14}$$

If electrons are present (such as in an atom), a competing process can occur which is called **electron capture**:

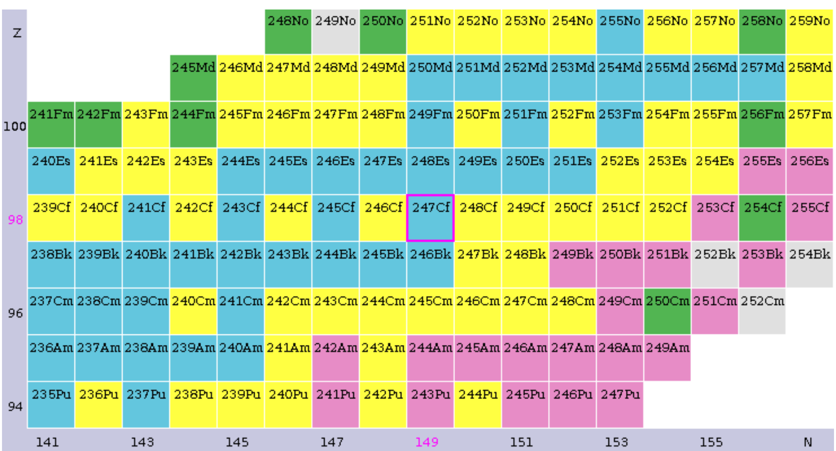
$$p^+ + e^- \longrightarrow n + \nu_e \tag{3.15}$$

Note that electron capture is a crossed reaction of the fundamental process underlying $\beta^+(e^+)$ decay.

As the neutron-to-proton ratio decreases even more, a proton is ejected from the unstable nucleus, according to

$$^{A}_{Z}\mathrm{X}_N \longrightarrow {}^{A-1}_{Z-1}\mathrm{Y}_N + {}^{1}_{1}\mathrm{H}^+_0 \tag{3.16}$$

There are two more types of nuclear decay that are only observed for heavy nuclei:



For nuclides with $Z > 50$ and a relatively small neutron-to-proton ratio **α-decay** is observed[7]:

$$^{A}_{Z}\mathrm{X}_N \longrightarrow {}^{A-4}_{Z-2}\mathrm{Y}_{N-2} + {}^{4}_{2}\mathrm{He}^{2+}_2 \tag{3.17}$$

For nuclides with $Z > 82$ **spontaneous fission** may occur:

$$^{A}_{Z}\mathrm{X}_N \longrightarrow {}^{A'}_{Z'}\mathrm{Y}_{A'-Z'} + {}^{A-A'}_{Z-Z'}\mathrm{W}_{A-Z-(A'-Z')} \tag{3.18}$$

Here's an example of spontaneous fission of Californium (Cf):

$$^{252}_{98}\mathrm{Cf}_{154} \longrightarrow {}^{140}_{54}\mathrm{Xe}_{86} + {}^{108}_{44}\mathrm{Ru}_{64} + 4\,{}^{1}_{0}\mathrm{n}_1 \tag{3.19}$$

In this case the fission into Xenon and Ruthenium is accompanied by the emission of four neutrons.

[7] This criterion is confirmed in the region where stable nuclides still exist. In the above section, however, there are no more stable structures among the nuclides of Bk (Berkelium), Es (Einsteinium), Fm (Fermium), etc.

3.3.2 Half Life

An important quantity for classifying the stability of nuclides is their **half life**, $t_{1/2}$. The half life is the instant in time when half of an ensemble N of particles (assumed to exist at $t = 0$) is still there. It can be determined analytically if the so-called **decay rate** Γ is known. The decay rate is the **probability of decay per time unit**[9].

Be $N(t)$ the number of particles at instant t and suppose that $\Gamma(\text{particle}) > 0$ is known. We define

$$\Delta N(t) := N(t_f) - N(t_i) < 0 \quad ; \quad t_f > t_i \tag{3.20}$$

the change in number of particles which is linearly[10] proportional to the number of existing particles at t, the respective time interval Δt, and the decay rate. We can write

$$\Delta N(t) = -\Gamma\, N(t)\, \Delta t \tag{3.21}$$

or in differential form

$$\begin{aligned} \frac{dN(t)}{dt} &= -\Gamma\, N(t) \\ \dot{N} &= -\Gamma\, N(t) \end{aligned}$$

The general solution of this linear homogeneous first-order differential equation is

$$N(t) = N(t=0)\, e^{\int_0^t -\Gamma\, dt'} = N(t=0)\, e^{-\Gamma\, t} \tag{3.22}$$

The number of existing particles at instant $t_{1/2}$ is, by definition,

$$N(t_{1/2}) = \frac{N(t=0)}{2} \tag{3.23}$$

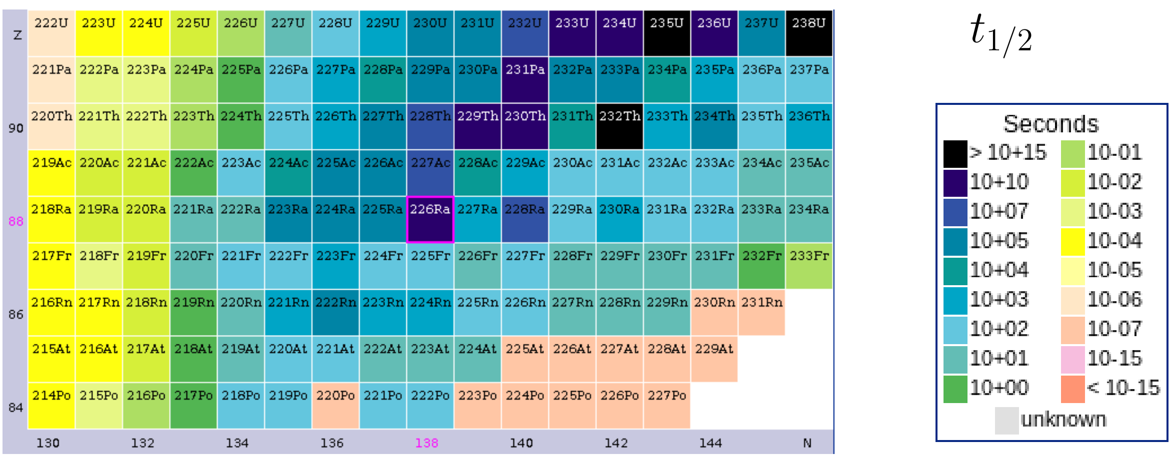
[9] In theory, this is the hard part. Γ can be calculated in the framework of Quantum Field Theory (QFT) invoking Feynman's calculus. If you are seriously interested, Griffith's book on elementary particles explains how to do it, in the final chapters. It does *not* explain the formal background of QFT, but how you get decay rates from Feynman rules. This is by itself not an easy exercise.

[10] The assumption is made that no decay affects any other.

Using this in Eq. (3.22) results in

$$\begin{aligned}
N(t_{1/2}) = \frac{N(t=0)}{2} &= N(t=0)\, e^{-\Gamma\, t_{1/2}} \\
\Leftrightarrow -\Gamma\, t_{1/2} &= \ln\left(\frac{1}{2}\right) = -\ln\left(\frac{2}{1}\right) = -\ln 2 \\
t_{1/2} &= \frac{\ln 2}{\Gamma}
\end{aligned}$$

So once the decay rate is known the half life is easy to calculate. The following chart shows a section of the nuclide table with associated half lives:



Thorium $^{232}_{90}\mathrm{Th}_{142}$ and Uranium $^{235}_{92}\mathrm{U}_{143}$ are stable nuclides with $t_{1/2} > 10^{15}$ [s]. The latter is used in *induced* nuclear fission where it absorbs a neutron to briefly form $^{236}_{92}\mathrm{U}_{144}$. This nuclide has a long half life of $\approx 10^{10}$ [s], but it is produced in an excited state which rapidly undergoes fission and releases excess energy.

3.4 Nuclear Structure - Nuclear Shell Model