5. Relativistic Mechanics

Equipped with the machinery of Chapter 4, we can finally construct a relativistic (Lorentz covariant) version of classical mechanics.

5.1. The relativistic point particle

1 | Point particle in $\mathbb{R}^{1,3}$ with trajectory $x^\mu(\tau)$:

2 | It is reasonable to define the relativistic momentum of a massive particle as follows:

\[ p^\mu := m u^\mu = m \frac{d x^\mu}{d \tau} = \left( \frac{m v c}{m v c} \right) \equiv \left( \frac{p^0}{\bar{p}} \right) \]

with (rest) mass $m$ and 3-momentum $\bar{p}$.

- $m$ is the good old (inertial) mass we would assign to the particle in classical mechanics; it is a measure of the particles resistance to changes in its state of motion. You can determine it by applying a (weak) force to the particle at rest and observing its initial acceleration: $m = F/a$. This mass is an intrinsic property of the particle and does not depend on velocity. It is sometimes called rest mass, but we will simply call it mass.

- Since the 4-velocity $u^\mu$ is a Lorentz vector, the 4-momentum is also a Lorentz vector; i.e., under a Lorentz transformation $\Lambda$ the 4-momentum transforms as $\tilde{p}^\mu = \Lambda^\mu_\nu p^\nu$.

- We will later rederive the expression for the 4-momentum as the conserved Noether charge for translations in spacetime.

3 | The spatial part of the momentum (the 3-momentum $\bar{p}$) is related to the velocity as follows:

\[ \bar{p} = m v_0 \tilde{v} = \frac{m \tilde{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (5.3) \]

Newtonian mechanics
• In special relativity the kinetic momentum is no longer proportional to the velocity. In particular for \( v \ll c \) the momentum of a massive particle diverges.

• The non-relativistic limit \( (v \ll c \Rightarrow \beta \ll 1 \Rightarrow \gamma \approx 1) \) is consistent with the Newtonian (non-relativistic!) relation \( \vec{p} = m\vec{v} \) for the kinetic momentum; the 3-momentum \( \vec{p} \) is therefore the proper relativistic version of the momentum in Newtonian mechanics.

• This explains why the above definition for the 4-momentum is reasonable – and why the mass \( m \) must be identified with the mass used in Newtonian mechanics.

• At this point it is unclear how to interpret the time-component \( p^0 \) (\( \rightarrow \) below).

\( \rightarrow \) The mass \( m \) is a Lorentz scalar: \( m^2 = p^2 / c^2 \)

• The 4-momentum is a time-like 4-vector for massive particles.

• This means that the mass \( m \) can be measured/computed in every inertial system by measuring/computing the 4-momentum \( p^\mu \) and its pseudo-norm \( p^2 \). The numerical result will always be the same, namely \( m^2 c^2 \).

5 | Equation of motion (EOM):

i | We want an EOM that …

• …is manifestly Lorentz covariant \( \rightarrow \) Lorentz tensor equation

• …reduces to Newton’s equation of motion

\[
\frac{d}{dt} \vec{p} = \vec{F} \quad \text{with} \quad \vec{p} = m\vec{v}
\]  \( \text{(5.5)} \) in the non-relativistic limit (correspondence principle).

ii | Suggestion:

\[
m b^\mu = \frac{d}{d\tau} p^\mu = K^\mu = \begin{pmatrix} K^0 \\ K \end{pmatrix} \quad \text{with} \quad \vec{K} \quad \text{4-force} \quad K^\mu.
\]  \( \text{(5.6)} \)

Because this is a equation built from Lorentz vectors, it is form-invariant (Lorentz covariant) by construction:

\[
m b^\mu = K^\mu \iff m \Lambda^\nu_\mu b^\mu = \Lambda^\nu_\mu K^\mu \quad \iff m\vec{b}^\nu = \vec{K}^\nu
\]  \( \text{(5.7)} \)

This is of course only so if the 4-force transforms like a Lorentz vector.

iii | Instantaneous rest frame (IRF) \( K_0 \):

a | At any time there is an inertial coordinate system \( K_0 \) in which the (potentially accelerated) particle is at rest \( \text{at this very moment} \) (if the particle is accelerating, it is also accelerating in this frame).

\[
m b^\mu_0 = \begin{pmatrix} 0 \\ m\vec{a}_0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vec{F}_0 \end{pmatrix} = K^\mu_0
\]  \( \text{(5.8)} \)
This follows from the correspondence principle: In the IRF the particle is in the non-relativistic, Newtonian limit. Thus its coordinate acceleration $\ddot{a}_0$ must be given by Newton’s equation of motion: $m\ddot{a}_0 = F_0$.

with

- **Proper acceleration** $\ddot{a}_0$

  The proper acceleration is the coordinate acceleration (3-acceleration) that you can measure (e.g., with an accelerometer) in the IRF $K_0$ of the particle.

  It follows immediately that the norm of the proper acceleration is a Lorentz scalar:

  $$b^2 = b^\mu b_\mu = -|\ddot{a}_0|^2 < 0$$  \hspace{1cm} (5.9)

- **Proper force** $\vec{F}_0$

  The proper force is the Newtonian force (3-force) you can measure (e.g., with a spring balance) in the IRF $K_0$ of the particle.

  We demand that this equation is Lorentz covariant, i.e., that $b^\mu$ and $K_0^\mu$ transform as contravariant Lorentz 4-vectors. We can then use a Lorentz boost to transform back into the lab frame in which the particle has coordinate velocity $\vec{v}$:

  Eq. (1.75)

  4-acceleration:

  $$b^\mu = (\Lambda_{-\vec{v}})^\mu_\nu b_0^\nu = \frac{\gamma_v}{\gamma_v^2 - (\ddot{a}_0 \cdot \vec{v})} \left( \ddot{a}_0 + \frac{\gamma_v - 1}{\gamma_v^2} (\ddot{a}_0 \cdot \vec{v}) \vec{v} \right)$$  \hspace{1cm} (5.10a)

  4-force:

  $$K^\mu = (\Lambda_{-\vec{v}})^\mu_\nu K_0^\nu = \frac{\gamma_v}{\gamma_v^2 - (\vec{F}_0 \cdot \vec{v})} \left( \vec{F}_0 + \frac{\gamma_v - 1}{\gamma_v^2} (\vec{F}_0 \cdot \vec{v}) \vec{v} \right)$$  \hspace{1cm} (5.10b)

  We will use these expressions later!

On the other hand, we can return to Eq. (5.6) and study the 4-force $K^\mu$ in more detail:

- **Spatial components of Eq. (5.6):**

  $$\frac{d\vec{p}}{d\tau} = \gamma_v(t) \frac{d\vec{p}}{dt} = \vec{K} \leftrightarrow \frac{d\vec{p}}{dt} = \frac{\vec{K}}{\gamma_v} = : \vec{F} \leftrightarrow \vec{K} = \gamma_v \vec{F}$$  \hspace{1cm} (5.11)

  with **3-force** $\vec{F}$.

  Here $\frac{d\vec{p}}{d\tau}$ denotes the change in momentum measured in coordinate time; it makes sense to identify this quantity with the relativistic analog of the Newtonian force.

- **What is the time component $K^0$ of the 4-force?**

  $$0 \overset{5.50}{=} mb^\mu u_\mu \overset{5.6}{=} K^\mu u_\mu = K^0 u^0 - \vec{K} \cdot \vec{u} \overset{4.47}{=} \gamma_v (K^0 c - \vec{K} \cdot \vec{v}) \hspace{1cm} (5.12)$$

  $\rightarrow$

  $$K^0 = \frac{\vec{K} \cdot \vec{v}}{c} \overset{5.11}{=} \frac{\gamma_v}{c} \vec{F} \cdot \vec{v}$$  \hspace{1cm} (5.13)
In summary, the 4-force in terms of the 3-force and the 3-velocity reads

\[ 4\text{-force: } K^\mu = \left( \gamma v \frac{F^\mu}{c} \right) \]  \hspace{1cm} (5.14)

**Example:**

In our discussion of electrodynamics (→ Section 5.4) we will find the following expression for the 3-force acting on a charged particle in an electromagnetic field:

\[ \vec{F} = q \vec{E} + \frac{q}{c} \vec{v} \times \vec{B} \]  \hspace{1cm} (5.15)

This is the conventional Lorentz force.

This example demonstrates that the 3-force \( \vec{F} \) is indeed the proper relativistic analog of Newtonian forces. Note, however, that it is only the component of the 4-force and thus does not transform nicely under Lorentz transformations.

**d** Spatial part of Eq. (5.6) \[ \text{Eq. (5.14)} \]

\[ \vec{\hat{F}} = \frac{d\hat{p}}{dt} = \frac{d}{dt} (m\gamma v \vec{v}) = \text{(Change in 3-momentum)} \]  \hspace{1cm} (5.16)

The Newtonian equation \( \vec{\hat{F}} = \frac{d\vec{p}}{dt} \) therefore remains valid in special relativity for the 3-force \( \vec{F} \) and the 3-momentum \( \vec{p} \). By contrast, \( \vec{\hat{p}} = m\gamma v \vec{v} \) is different from the Newtonian relation \( \vec{\hat{p}} = m\vec{v} \) between momentum and velocity.

**e** Temporal part of Eq. (5.6) \[ \text{Eq. (5.14)} \]

\[ \frac{dp^0}{dt} = \gamma v \frac{dp^0}{dt} = \gamma v \frac{\vec{F} \cdot \vec{v}}{c} \quad \Rightarrow \quad \frac{d(cp^0)}{dt} = \vec{F} \cdot \vec{v} \]  \hspace{1cm} (5.17)

\( \rightarrow \vec{F} \cdot \vec{v} \): Work performed by \( \vec{F} \) on particle

\( \rightarrow E = cp^0 \): Total energy of particle

Note that we can actually only conclude \( E = cp^0 + \text{const} \) from the differential equation above. We will later see that the constant must be set to zero because \( p^0 \) is the conserved Noether charge that derives from time translations.

The time component of the EOM Eq. (5.6) can therefore be written as:

\[ \vec{F} \cdot \vec{v} = \frac{dE}{dt} = \frac{d}{dt} (m\gamma v c^2) = \text{(Change in energy)} \]  \hspace{1cm} (5.18)

We will discuss the expression for the energy in Section 5.2 below.

**6** Above we expressed the 4-force in terms of the proper force \( \vec{F}_0 \) and in terms of the 3-force \( \vec{F} \). Equating the two expressions yields a relation between the 3-force \( \vec{F}_0 \) measured in the IRF and the 3-force \( \vec{F} \) measured in the lab frame:
Recall that the proper force is the Newtonian force you would measure with a spring scale in the IRS of the particle. In contrast to Newtonian mechanics, the force \( F \) measured from a frame in relative motion is different from \( F_0 \). In the non-relativistic limit \( v \approx 1 \) we find \( F \approx F_0 \) and this distinction becomes irrelevant (as assumed by Newtonian mechanics).

A similar comparison yields a relation between the 3-acceleration in the IRF (the proper acceleration) and the 3-acceleration in the rest frame:

This is again in sharp contrast to Newtonian mechanics where, as a consequence of absolute time, acceleration does not depend on the velocity of the reference frame. In the non-relativistic limit for \( v \approx 1 \) we find \( \ddot{a} \approx \ddot{a}_0 \), consistent with Newtonian mechanics.

Sanity check: If we integrate the equation of motion Eq. (5.16), we find:

For a finite 3-force \( |\vec{F}| < \infty \) and finite time \( T < \infty \), and non-zero mass \( m \neq 0 \), it follows for the final velocity \( \vec{v}_T \):

Thus the dynamics does not allow massive particles to reach the speed of light, no matter how strong the force or how long the acceleration! This is in direct contradiction to Newtonian mechanics and by now experimentally well-confirmed (\( \rightarrow \) below).

5.2. Momentum, Energy, and Mass

To summarize, the 4-momentum of a massive particle can be written as:

\[
p^\mu = mu^\mu = \left( \frac{p^0}{\beta} \right) = \left( \frac{E/c}{\beta} \right) = \left( \frac{\gamma v m c}{\gamma v m \beta} \right)
\]
10 | The relativistic energy of a massive particle is then (as a function of 3-velocity):

\[ E = c p^0 = \gamma_0 m c^2 = \frac{m c^2}{\sqrt{1 - \frac{v^2}{c^2}}} \] (5.24)

With

\[ m^2 c^2 = p^2 = (p^0)^2 - (\vec{p})^2 = E^2/c^2 - \vec{p}^2 \] (5.25)

we find the alternative expression as a function of 3-momentum:

\[ E = \sqrt{\vec{p}^2 c^2 + m^2 c^4} \] (5.26)

- This expression is also valid in the massless case \( m = 0 \) (→ below).
- Eq. (5.25) has actually two solutions: \( E = \pm \sqrt{\vec{p}^2 c^2 + m^2 c^4} \). In relativistic mechanics (and relativistic single-particle quantum mechanics), we can ignore the negative energy solution and consider only time-like 4-momenta \( p^\mu \) that point into the future light-cone. In quantum field theory, where interacting particles can be destroyed and produced, these negative energy solutions necessitate the introduction of ↑ antiparticles (like the positron).
- For fixed mass \( m \), Eq. (5.25) determines a 3-dimensional hypersurface in the 4-dimensional “energy-momentum space” spanned by 4-momenta \( p^\mu = (p^0, \vec{p}) \in \mathbb{R}^4 \). For \( m \neq 0 \) this hypersurface is a hyperboloid of two sheets \( E = \pm \sqrt{\vec{p}^2 c^2 + m^2 c^4} \) (for \( m = 0 \) it is a cone: \( E = \pm c |\vec{p}| \)). This hypersurface is called ♠ mass shell. If a 4-momentum satisfies the energy-momentum relation (with either sign) we say that it is “on-shell”; if not, it is “off-shell”. In quantum field theory, real particles that can be measured are always on-shell; intermediate “virtual particles” in scattering processes can be off-shell.

11 | Rest energy:

\[ p^0 = \begin{pmatrix} p_0^0 & \square \end{pmatrix} = \begin{pmatrix} E_0/c \\ 0 \end{pmatrix} \] (5.27)

For these considerations, it does not matter whether the particle is accelerating and this is an IRF, or whether the particle is in inertial motion and has a fixed rest frame. Formally, since \( p^2 = m^2 c^2 > 0 \) is a time-like Lorentz vector, there is always an inertial frame in which \( p^0 \neq 0 \) and \( \vec{p} = 0 \).

\[ E_0 = mc^2 \] (5.28)

This is Einstein’s famous principle of equivalence of (inertial) mass and (rest) energy.

- ↑ The total energy \( E \) is the time-component of a 4-vector: \( p^\mu = (E/c, \vec{p})^T \); thus it makes sense to refer to the rest energy \( E_0 \) – which is the component of this 4-vector in the rest frame \( K_0 \), i.e., the particular frame where \( \vec{p} = 0 \).
• ![By contrast, the mass is a Lorentz scalar, namely $p^2 = m^2c^2$; hence it is the same in all inertial systems and it does not make sense to refer to the rest mass $m_0$ as this term suggests that there is a “non-rest mass” (which there isn’t).

• Einstein first derived the mass-energy equivalence in his Annus Mirabilis paper Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig? [9]. In the paper, the equation is not given verbatim but encoded in the following statement:

$Gibt ein Körper die Energie L in Form von Strahlung ab, so verkleinert sich seine Masse um $L/V^2$.

Einstein concludes:

$Die Masse eines Körpers ist ein Maß für dessen Energieinhalt; […] Es ist nicht ausgeschlossen, daß bei Körperrn, deren Energie in hohem Maße veränderlich ist (z.B. bei den Radiumsalzten), eine Prüfung der Theorie gelingen wird.$

Einstein further elaborates on the relativistic energy relation and its implications in [60]. He provides self-contained step-by-step derivation in Ref. [61]. Additional insight was provided over the years with alternative derivations by various authors [62–64].

The derivation by Feigenbaum and Mermin in [64] is particularly insightful as it follows Einstein’s original derivation in [9] closely without invoking electrodynamics. They also demonstrate that the heart of relativistic mechanics is actually Eq. (5.24) (where $mc^2$ appears as a coefficient), and not Eq. (5.28) (which is conventional).

**Note 5.1: Some comments on $E_0 = mc^2$**

Eq. (5.28) is arguably the most famous equation in physics. The popularization of scientific concepts is often accompanied by simplifications and distortions. This is also the case for $E_0 = mc^2$:

- $E_0 = mc^2$ is often written as $E = mc^2$. This is either wrong or misleading (depending on the interpretation of the symbols); in any case, it is not consistent with modern conventions in relativity (→ below).

- $E_0 = mc^2$ is by no means Einstein’s most important equation. This is why it is not referred to as “Einstein equation;” this honor goes to

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu} \quad (5.29)$$

which are also known as the → Einstein field equations; these form the basis of general relativity and are empirically of much greater value than Eq. (5.28). Luckily, the Einstein field equations look daunting and are not nearly as accessible as $E_0 = mc^2$; hence they weren’t seized (and mutilated) by pop culture like $E_0 = mc^2$ was.

- How statements are phrased determines our conceptualization of the world. The often heard phrase

“$E_0 = mc^2$ says that mass can be converted into energy”

makes me think of “mass” as a sort of coal that can be lighted and then produces energy (maybe in form of light and heat or an atomic explosion). I am quite convinced that there are many who got “conceptually derailed” by statements like this, and hence think of Einstein’s revelation as modern-day equivalent of an early human realizing, perhaps by witnessing a lightning strike, that wood can be kindled to produce heat. This is completely off the mark.
\( E_0 = mc^2 \) says that rest energy and inertial mass are equivalent; not that they can be “converted” into each other. It means that the Lorentz symmetry of spacetime necessitates that our concepts of “energy” (as a quantity that can make things change in time) and “inertial mass” (as a quantity that measures how hard it is to make the state of motion of an object change in time) are like two sides of the same coin. Note that we did not arrive at the equation by studying the microscopic dynamics and interactions of matter (like we do in quantum mechanics, and especially quantum field theory); the equivalence of rest energy and mass is a consequence of the symmetries of spacetime alone. One can take \( E_0 = mc^2 \) thus as a hint at the unanswered questions “What is time?” and “What is inertia?” because energy is the generator of time translations (think of the time-evolution operator in quantum mechanics) and mass quantifies the phenomenon of inertia.

To drive the point home, here a few examples:

- An atom in an excited electronic state is heavier than the same atom in the ground state.
- A battery gets lighter when being discharged.
- A chunk of metal is heavier when it is hot.
- If you put an atomic bomb into an opaque, completely sealed “super box” that survives the explosion, the weight of the box does not change when the bomb goes off. This makes it clear that mass is not “converted” into energy.
- If the box is made out of “super glass” that lets only photons escape, the box gets lighter by \( E_{\text{phot}}/c^2 \) if the photons carry away the energy \( E_{\text{phot}} \).

For these reasons, \( E_0 = mc^2 \) is not a magical blueprint to build atomic bombs.

The equation is only relevant in this context because it provides a nice “shortcut” to compute the energies that the fission (splitting) of isotopes can yield (or cost, depending on the isotopes). Because one could measure the rest masses of isotopes rather easily (using mass spectrometry [65]) – but had almost no clue how to describe the inner workings (and therefore binding energies) of said nulei – the equation allowed for a straightforward survey of the periodic table to identify suitable isotopes that would yield energy under fission. \( E_0 = mc^2 \) is not the reason why atomic weapons work, and these weapons are not so powerful “because they convert mass into energy.” This is pure nonsense. If you discharge the battery of your phone, it also looses mass – because rest energy and mass are equivalent: \( E_0 = mc^2 \)!

And yes, this mass difference is much smaller than the mass difference accompanied by a nuclear explosion. But this is not the reason; the reason is that the strength of electromagnetic interactions – which govern chemical processes (like discharging your battery) – is dwarfed by the strength of the strong interaction (and its residual, the nuclear force) – which governs nuclear reactions.

In a nutshell:

When studying reaction processes (of any sort), the change of restmass predicted by \( E_0 = mc^2 \) is an epiphenomenon. The mass change is not causal; it cannot be, because it is a consequence of the symmetries of spacetime, and not of the inner workings of matter.

Unfortunately, the notation and interpretation of special relativity has changed since its inception. In former times it was conventional to introduce the concept of a

\[
\gamma \text{ Relativistic mass: } m_r := \gamma m = \frac{m}{\sqrt{1 - v^2/c^2}} \tag{5.30}
\]
which depends on velocity. With this definition, the relativistic relation between 3-velocity and 3-momentum reads \( \vec{p} = m_r \vec{v} \) and parallels the Newtonian relation \( \vec{p} = m \vec{v} \). The relativistic energy relation then reads \( E = m_r c^2 \).

The concept of a velocity-dependent, relativistic mass is avoided in most modern treatments of relativity (and in this script). While this is mostly a matter of concepts and semantics, there are good reasons why the concept of a velocity dependent mass is less useful than it might seem (→ below).

Here a few comments on various notations that you might encounter:

\[
E_0 = mc^2 \quad \text{Correct} \begin{equation}
E = \overline{m} c^2 \quad \text{Only makes sense if } m = m_r \text{ (which we don’t use).}
E_0 = m_0 c^2 \quad \text{Why } m_0? \text{ There is only } m!
E = \overline{m}_0 c^2 \quad \text{Energy is frame-dependent. Do you mean } E_0? \text{ Otherwise: Wrong!}
\]

For more details and explanations see Refs. [66–68].

iii | → Take home message:

There is only one mass: the rest mass \( m \) (which we call mass). Thus mass does not depend on velocity.

This convention is used by almost all modern textbooks on relativity.

Unfortunately the old conventions (using relativistic, velocity-dependent masses) are still used by school books and popular science books.

iv | Aside: Why introducing velocity depended masses leads nowhere.

If you are still inclined to think in terms of a velocity-dependent, relativistic mass \( m_r \), here is a compelling argument why this is a useless and artificial concept that needs to die:

The 3-component of the relativistic equation of motion Eq. (5.16) reads

\[
\vec{F} = \frac{d}{dt} \left( m \gamma_v \vec{v} \right) = m \gamma_v \vec{a} + m \gamma_v \frac{\vec{v} \cdot \vec{a}}{c^2} \vec{v}
\]

with two extreme cases:

\[
\vec{v} \parallel \vec{a} \quad \Rightarrow \quad \vec{F} \approx m \gamma_v^3 \vec{a}
\]

\[
\vec{v} \perp \vec{a} \quad \Rightarrow \quad \vec{F} = m \gamma_v \vec{a}
\]

If you insist on introducing a “mass” as the proportionality factor between 3-force and 3-acceleration to quantify the inertial response of an object at finite velocity, you are not only forced (\( \text{\textcircled{\#}} \)) to make this mass velocity dependent, you also need two masses:

“Longitudinal mass”: \( m_\parallel := m \gamma_v^3 \)  
“Transverse mass”: \( m_\perp := m \gamma_v \)

The above result demonstrates that the concept of a mass as a measure for inertia is not very useful in special relativity. More precisely, the result shows that the quantities \( m_\parallel \) and \( m_\perp \) are relational properties between an object and an observer (they depend on the state of motion of the observer); they are not intrinsic properties of the object itself. Only the restmass \( m \) qualifies as such an intrinsic property. The velocity dependence of \( m_\parallel \) and \( m_\perp \) is not an intrinsic feature of matter, it is a feature of spacetime.
This is why in modern textbooks there is only one mass $m$ (the rest mass) which does not depend on $v$, and one has to accept that the Newtonian relation $\vec{p} = m \vec{v}$ is no longer valid. The concepts of “longitudinal mass” and “transverse mass” (and velocity dependent mass, for that matter) are therefore no longer used in modern literature.