Relativity in Classical Physics

Main Points
Introduction
Galilean (Newtonian) Relativity
Relativity & Electromagnetism
Michelson-Morley Experiment

Introduction

The theory of relativity deals with the study of physics as described by observers that are moving relative to each other. To quantitatively describe motion, we need to specify positions and times, which is done with the use of coordinate systems. To describe the motion of an object, the observer uses a coordinate system relative to which she is at rest. A coordinate system that is used to quantitatively specify the positions and times of events is called a frame of reference (or reference frame). An event is an occurrence, which is completely specified by its position and the time at which it occurred. Clearly, the motion of an object is a succession of events (the arrival of the object at a position along its trajectory at a certain time). The value of a field (e.g., a magnetic field) at a given location and time is also an event. Thus, physics can be described in terms of events. Relativity describes the relationship between events that are observed from different reference frames. We note here that when we speak of "observing," we mean "making quantitative measurements." Thus, observers of events make measurements of quantities that characterize the event (position, time, magnetic field, etc.).

Inertial reference frames (IRF) are fundamentally important in relativity theory. In Newtonian mechanics, an IRF is one in which Newton’s first law is valid. (We shall see an improved definition later.) Thus, in an IRF, an object that is acted upon by zero net force will remain at rest or continue moving with constant velocity. Any reference frame that moves with constant velocity relative to an established IRF is itself an IRF. An IRF must be unaccelerated because in an accelerated reference frame, Newton's first law is violated. [Show diagrams of accelerated train car.]

In the most general case, one reference frame accelerates relative to another. The complete description of events under these circumstances is the domain of the general theory of relativity. General relativity turns out to be a theory of gravitation because motion in a gravitational field is indistinguishable from motion in an appropriate accelerated reference frame. In this course, we shall focus upon the special theory of relativity, which describes physics as observed from inertial reference frames.

Experimental Underpinnings of Relativity

Special relativity supersedes Newtonian mechanics in the sense that it is more general than Newtonian mechanics, which fails when speeds approach the speed of light in vacuum, c.

---

1 See Resnick (1972) for a good discussion of relativity. See also Benson's University Physics
2 See “Six Ideas” Unit R. Cubic lattice with clocks that are synchronized.
However, for speeds much less than \( c \), which correspond to our normal experience, Newton’s laws become an excellent approximation to the dynamical equations of special relativity. Thus, it is justifiable to use Newton’s laws to describe motion for speeds \( v \ll c \). All of this has been experimentally verified. The Michelson-Morley experiment (to be studied later) was among the first to clearly show a deviation from Newtonian expectations. Other experiments since then have shown the departure from Newtonian mechanics more dramatically. [Show figure of kinetic energy of electron vs. speed.] Let us begin our discussion of relativity with a look at relativity from a classical, Newtonian perspective.

**Rotated Coordinate Systems and Relativity**

Consider a point P whose position is specified by \( x-y \) coordinates \((x_P, y_P)\). Now rotate the coordinate systems about the origin so we have a new system \( x'-y' \). Obviously, the coordinates of P \((x_P', y_P')\) are different in this system than in the \( x-y \) system. How does this affect our description of the motion? The coordinates are different, but do Newton’s laws still hold? Yes! The position vector of the point P is the same in both systems. (Its magnitude and direction [relative to, for example, the "fixed" stars] are identical in both systems.) Displacements of P will also be the same. Thus, its velocity in both frames will be the same and the accelerations will also be identical. Rotation of the axes only changes the components; it does not change the physical quantities themselves. It also makes sense intuitively. If you drop a ball from a certain height, you should get the same answer for the time to hit the ground regardless of the orientation of the spatial coordinate system. This is because Newton’s laws are valid regardless of the orientation of the coordinate system as we have tacitly shown many times in doing inclined plane problems. Choosing a coordinate system or reference frame such that \( x \) is parallel to the surface of the incline simplifies the problem. But the same answer is obtained if we choose the \( x'-y' \) coordinate system, where \( x' \) is horizontal and \( y' \) is vertical. Newton’s laws do not change under rotation and/or translation (displacement) of the coordinate system.

**Galilean (Newtonian) Relativity**

Consider two inertial reference frames S and S', with S' moving at a constant velocity \( \vec{v} \) relative to S. If Newton's laws of motion are valid in S, are they valid in S'? If so, then the laws would be valid in all inertial frames and therefore all inertial frames would be equivalent.
Consider the IRFs $S$ and $S'$ as shown in the diagram. At some instant in time $t$, the position vector of the object as observed from frame $S$ is $r$ and the position vector as observed from $S'$ is $r'$. The origin $O'$ of the frame $S'$ is at a position $\vec{R} = \vec{v}t$ relative the origin $O$ of $S$. At time $t = t' = 0$, the origins coincided. Vector algebra tells us that the position vectors are related thus:

$$r' = r - \vec{R}. \quad (1.1)$$

Since $\vec{R} = \vec{v}t$, Eq. (1.1) becomes

$$r' = r - \vec{v}t. \quad (1.2)$$

Eq. (1.2) (along with the absolute-time assumption discussed below) is called the **Galilean Transformation**. Usually, it is written in component form, as we shall see later. Now let us take the derivative of Eq. (1.2) with respect to time. In classical physics, time is absolute, i.e., $t' = t$.

Thus, if two clocks are synchronized in one inertial frame, then the clocks will be synchronized in all inertial frames. Therefore, we have

$$\frac{dr'}{dt'} = \frac{d}{dt}(r - \vec{v}t), \quad (1.3)$$

i.e.,

$$\ddot{u}' = \ddot{u} - \vec{v}. \quad (1.4)$$

Eq. (1.4) is called the **Galilean Velocity Transformation**. This is equivalent to the vector addition of velocities. Let us now take the time derivative of Eq. (1.4). We find

$$\frac{d\ddot{u}'}{dt'} = \frac{d}{dt}(\ddot{u} - \vec{v}). \quad (1.5)$$

Because $\vec{v}$ is constant, we obtain

$$\dddot{u}' = \dddot{u}. \quad (1.6)$$

Thus, the accelerations are equal in both frames and therefore in all inertial frames. The acceleration is said to be **invariant** in inertial reference frames.

Let us now assume that Newton's first law is valid in inertial frame $S$, which is consistent with experiment. Thus, as observed in $S$, when no net force acts on the object, it moves with constant velocity ($\ddot{u}$). To simplify the discussion, we assume that the object is totally isolated, i.e., all forces acting on it are negligible. From Eq. (1.4), we deduce that the velocity of the object as observed (measured) in $S'$ is $\dddot{u}' = \dddot{u} - \vec{v}$, which is constant because $\vec{v}$ is constant. Therefore, Newton's first law is also valid in frame $S'$ because the object also moves with constant velocity as measured in that frame. It follows that Newton's first law is valid in all inertial reference frames.

Next, we assume that Newton's second law is valid in inertial frame $S$, as deduced from experiment. Thus, in frame $S$,

$$\vec{F}_{\text{net}} = m\dddot{u}. \quad (1.7)$$

Let us further assume that the net force is due to a position-dependent force. This is not very restrictive because commonly encountered forces such as gravity and the electric force are position-dependent. These forces really depend on the relative position of one object with respect to another with which it interacts. Thus, the forces are a function of the difference in position of the two interacting objects: $\vec{F} = \vec{F}(\vec{r}_2 - \vec{r}_1)$. We can calculate the corresponding relative position in $S'$ by using Eq. (1.2):
\[ \mathbf{r}_2 - \mathbf{r}_1 = (\mathbf{r}'_2 + \mathbf{v}_1 t) - (\mathbf{r}'_1 + \mathbf{v}_1 t) = \mathbf{r}' - \mathbf{r}'_1. \]  

Hence, the relative positions are equal; therefore, the forces must also be equal in both frames:

\[ \mathbf{F}^\prime = \mathbf{F} = \mathbf{F}' _{net} . \]  

(1.9)

The same conclusion would be obtained if the force were dependent on the relative velocity or both relative position and relative velocity. A force that is dependent on absolute position or absolute velocity would not be the same in both frames; however, experiments are consistent with the invariance of the force. (Note: friction and air resistance are really dependent on the relative velocity of two objects.)

What about the masses? Classical physics simply declares that the masses are equal:

\[ m' = m. \]  

(1.10)

Thus, mass is an invariant in classical physics. Measurements made under normal circumstances with equipment available in an undergraduate laboratory would confirm this.

Finally, we have previously seen [Eq. (1.6)] that the accelerations as measured in both frames are equal, i.e., \( \ddot{a}' = \ddot{a} \). Since the forces are equal, the masses are equal, and the accelerations are equal, it follows that Newton's second law must also be valid in \( S' \):

\[ \mathbf{F}' _{net} = m' \ddot{a}'. \]  

(1.11)

If we assume Newton's third law is valid in \( S \), then its validity in \( S' \) follows directly from the validity of the second law in both frames.

We conclude that Newton's laws are valid in all inertial reference frames and they have exactly the same form. Since the transformation of coordinates from one inertial frame to another is the Galilean transformation, we say that Newton's laws are covariant under a Galilean transformation. This is the formal expression of Galilean (or Newtonian) relativity. Since Newton's laws are exactly the same in all inertial frames, it follows that no mechanical experiment can distinguish one IRF from another. If you were in a windowless soundproof rail car, you could not tell whether it was "at rest" or moving with a constant non-zero velocity. [Group discussion: pendulum in a uniformly moving or accelerating rail-car.]

It should be noted that the classical conservation laws for momentum and energy are also covariant under a Galilean transformation, i.e., the laws are valid in all inertial frames. This follows from the fact that the conservation laws can be derived from Newton's laws, which are covariant as seen above.

**Galilean Relativity and the Speed of Light**

The other great edifice of classical physics is electromagnetism. One result of classical electromagnetism is that the speed of light in vacuum \( (c) \) is the same in all directions, independent of the
motion of the source. One quickly sees that this is inconsistent with the Galilean velocity transformation \( \vec{u}' = \vec{u} - \vec{v} \) since the speed of light would depend on the relative motion of the reference frames (see figure).

A related problem was the fact that light travels in a vacuum. Since pre-20th century physicists believed that all waves needed a medium in which to propagate, a medium called the “ether” was invented for light to travel in. The ether frame was therefore the unique frame in which the speed of light was \( c \). All frames moving relative to the ether would measure a speed of light different from \( c \) in general. The ether had distinctly unphysical properties in that it had to have infinite elasticity and inertia, no mass, and was undetectable. However, the ether concept persisted. Experiments showed that the Earth had to travel through the ether (e.g., experiments on aberration of starlight). [See Resnick; also, Anderson.] This being the case, one should then be able to detect the “ether wind,” i.e., the motion of the Earth through the ether, by showing that light beams on Earth have different speeds depending upon the velocity of the Earth through the ether, as we saw above (see figure).

**Michelson-Morley Experiment**

The most famous experiment to try to detect the motion of the Earth through the ether is the Michelson-Morley experiment. Albert Michelson designed an interferometer that could detect extremely small differences in the path lengths traveled by interfering coherent light beams. (The Michelson interferometer will be used in the Modern Physics Laboratory course to measure extremely small differences in wavelength.) In the Michelson interferometer, the half-silvered mirror (beam-splitter) partially transmits and partially reflects the beam from the source along paths 1 and 2. (See figure below.) The split beams are reflected by mirrors \( M_1 \) and \( M_2 \) and brought together at the beam-splitter where they are partially transmitted and partially reflected. The overlapping beams may then be observed using a telescope or another imaging device. If

---

3 In general, Maxwell’s equations are not covariant under a Galilean transformation. [See Tipler and Llewellyn (simple example); Griffiths; or Reitz, Milford, and Christy.]
mirrors $M_1$ and $M_2$ are nearly perpendicular, a set of nearly parallel interference fringes will be observed in the telescope. At the position of a bright fringe, the interfering beams are in phase. At the position of a dark fringe, the beams are out of phase. Changes in the path length (along the horizontal or vertical paths) result in the fringes moving across the field of vision. In the figure, $v$ is the velocity of the interferometer (and Earth) relative to the ether.

To understand the Michelson-Morley experiment and its ramifications, we will calculate the phase difference between beams 1 and 2. To do this, we need to calculate the total time beam 1 spends along path 1 and the total time beam 2 spends along path 2. The phase difference is related to the difference between these two times. We need to calculate times because the speed of light along path 1 will be different from that along path 2 according to the Galilean velocity transformation. [The fringe that is seen at a particular position depends on what phases (crest or trough or something in between) of the two wave trains arrive simultaneously at that position.]

Since the path lengths shown are in the frame of the interferometer, we need to determine the velocity of the light relative to the interferometer. We let the $S$ frame be the ether frame and $S'$ be the interferometer. Then, according to Eq. (1.4), $u' = u - v$. Since the speed of light relative to the ether is $c$, $|u| = c$. For path 1, when the beam is moving away from $M_1$, $u' = c - v$ and when the beam is toward $M_1$, $u' = c + v$. Thus, the total time to traverse path 1 (back and forth) is

$$t_1 = \frac{l_1}{c-v} + \frac{l_1}{c+v} = l_1 \left( \frac{2c}{c^2-v^2} \right) = l_1 \left( \frac{1}{c} \frac{1}{1-v^2/c^2} \right).$$

The total time to traverse path 2 can be easily calculated by recognizing that this situation is identical to that of a boat moving across a river, perpendicular to the current. The path of the light shown in the diagram above is the path traveled in the interferometer (toward mirror $M_2$). We need the speed of light along this path. Recall that $u' = u - v$. The vector diagram to the right shows how the velocities add. Using the Pythagorean theorem, the speed of the light relative to the interferometer is $u' = \sqrt{c^2-v^2}$. By constructing a velocity triangle similar to that shown, one finds that the speed of the light relative to the interferometer while traveling away from $M_2$ is the same as that while moving toward $M_2$. Thus, the total time taken to and from $M_2$ is given by

$$t_2 = 2 \frac{l_2}{\sqrt{c^2-v^2}} = \frac{2l_2}{c} \frac{1}{\sqrt{1-v^2/c^2}}.$$  \hfill (1.13)

The difference in traversal times is

$$\Delta t = t_2 - t_1 = 2 \frac{l_2}{c} \left( \frac{l_1}{\sqrt{1-v^2/c^2}} - \frac{l_1}{1-v^2/c^2} \right).$$ \hfill (1.14)

If this time difference at a particular position equals a whole number of periods of the light wave, the wave trains will be in phase and a bright fringe will be seen at that position.

Now, let us rotate the entire apparatus by $90^\circ$ so that paths 1 and 2 essentially switch places. Then the transit time difference in this case is given by
\[ \Delta t_r = t_{2r} - t_{1r} = \frac{2}{c} \left( \frac{l_2}{1 - v^2/c^2} - \frac{l_1}{\sqrt{1 - v^2/c^2}} \right). \]  

(1.15)

Thus, the rotation changes the time differences by

\[ \Delta t_r - \Delta t = \Delta t_r = t_{2r} - t_{1r} = \frac{2}{c} \left( \frac{l_2 + l_1}{1 - v^2/c^2} - \frac{l_2 + l_1}{\sqrt{1 - v^2/c^2}} \right). \]  

(1.16)

It seems reasonable to assume that the speed \( v \) of the Earth relative to the ether should be approximately equal to the orbital speed of the Earth around the Sun. The orbital speed of the Earth is approximately \( 3 \times 10^4 \) m/s and so \( v^2/c^2 \approx 10^{-8} \), which is much less than 1. We therefore use the binomial expansion up to the first order term to approximate Eq. (1.16):

\[ \Delta t_r - \Delta t \approx \frac{2(l_2 + l_1)}{c} \left( 1 + \frac{v^2}{c^2} - 1 - \frac{1}{2} \frac{v^2}{c^2} \right) = \left( \frac{l_2 + l_1}{c} \right) \frac{v^2}{c^2}. \]  

(1.17)

This corresponds to a shift in the number of fringes given by

\[ \Delta N = \frac{\Delta t_r - \Delta t}{\lambda/c} = \frac{\Delta t_r - \Delta t}{\lambda} = \left( \frac{v^2/c^2}{\lambda} \right) (l_1 + l_2). \]  

(1.18)

For a typical setup, \( l_1 = l_2 = 10 \) m and \( \lambda = 500 \) nm, which gives \( \Delta N \approx 0.4 \). This fringe shift is easily observable. The experiment was repeated many times at different times of year (to account for the difference in the Earth’s orbital velocity), and no fringe shift has ever been observed. (The precision was less than one-hundredth of a fringe.) Thus, one is forced to conclude that the ether frame does not exist, since the null result proves that the speed of light is the same in all directions in a frame that supposedly moves relative to the ether. Many more experiments support this conclusion.

Thus, Maxwell’s electromagnetic theory is valid in all inertial frames, but is inconsistent with Galilean relativity. How could this ugly discrepancy be rectified? Einstein believed deeply that all physical laws should be invariant in all inertial reference frames. This belief is the foundation of the special theory of relativity, which shows that Galilean relativity must be abandoned in order to obtain a theory that describes mechanics and electromagnetism consistently.

Discuss modern Lorentz invariance experiments. (See inside front binder cover inserts, Physics Today, July 2004, p. 40.)