On the Electrodynamics of Moving Bodies

By Albert Einstein

On the Electrodynamics of Moving Bodies

It is well known that Maxwell's electrodynamics—as usually understood at the present time—when applied to moving bodies, leads to asymmetries that do not seem to be inherent in the phenomena. Let us recall, for example, the electrodynamic interaction between a magnet and a conductor. The observable phenomenon depends here only on the relative motion of the conductor and the magnet, whereas the customary view draws a sharp distinction between the two cases in which either the one or the other of these bodies is in motion. For if the magnet is in motion and the conductor is at rest, there arises in the surroundings of the magnet an electric field endowed with a certain energy value that produces a current in the places where parts of the conductor are located. But if the magnet is at rest and the conductor is in motion, no electric field arises in the surroundings of the magnet, while in the conductor an electromotive force will arise, to which in itself there does not correspond any energy, but which, provided that the relative motion in the two cases considered is the same, gives rise to electrical currents that have the same magnitude and the same course as those produced by the electric forces in the first-mentioned case.

Examples of a similar kind, and the failure of attempts to detect a motion of the earth relative to the "light medium", lead to the conjecture that not only in mechanics, but in electrodynamics as well, the phenomena do not have any properties corresponding to the concept of absolute rest, but that in all coordinate systems in which the mechanical equations are valid, also the same electrodynamic and optical laws are valid, as has already been shown for quantities of the first order. We shall raise this conjecture (whose content will be called "the principle of relativity" hereafter) to the status of a postulate and shall introduce, in addition, the postulate, only seemingly incompatible with the former one, that in empty space light is always propagated with a definite velocity V which is independent of the state of motion of the emitting body. These two postulates suffice for arriving at a simple and consistent electrodynamics of moving bodies on the basis of Maxwell's theory for bodies at rest. The introduction of a "light ether" will prove superfluous, inasmuch as in accordance with the concept to be developed here, no "space at absolute rest" endowed with special properties will be introduced, nor will a velocity vector be assigned to a point of empty space at which electromagnetic processes are taking place.

Like every other electrodynamics, the theory to be developed is based on the kinematics of the rigid body, since assertions of each and any theory concern the relations between rigid bodies (coordinate systems), clocks, and electromagnetic processes. Insufficient regard for this circumstance is at the root of the difficulties with which the electrodynamics of moving bodies must presently grapple.

I. Kinematical Part

§1. Definition of simultaneity

Consider a coordinate system in which the Newtonian mechanical equations are valid. To distinguish it verbally from the coordinate systems that will be introduced later on, and to visualize it more precisely, we will designate this system as the "system at rest."

If a material point is at rest relative to this coordinate system, its position relative to the latter can be determined by means of rigid measuring rods using the methods of Euclidean geometry and can be expressed in Cartesian coordinates.

If we want to describe the *motion* of a material point, we give the values of its coordinates as a function of time. However, we should keep in mind that for such a mathematical description to have physical meaning, we first have to clarify what is to be understood here by "time." We have to bear in mind that all our propositions involving time are always propositions about *simultaneous events*. If, for example, I say that "the train arrives here at 7 o'clock," that means, more or less, "the pointing of the small hand of my clock to 7 and the arrival of the train are simultaneous events."[1]

It might seem that all difficulties involved in the definition of "time" could be overcome by my substituting "position of the small hand of my clock" for "time." Such a definition is indeed sufficient if time has to be defined exclusively for the place at which the clock is located; but the definition becomes insufficient as soon as series of events occurring at different locations have to be linked temporally, or—what amounts to the same—events occurring at places remote from the clock have to be evaluated temporally.

To be sure, we could content ourselves with evaluating the time of the events by stationing an observer with the clock at the coordinate origin, and having him assign the corresponding clock-hand position to each light signal that attests to an event to be evaluated and reaches him through empty space. But as we know from experience, such an assignment has the drawback that it is not independent of the position of the observer equipped with the clock. We arrive at a far more practical arrangement by the following consideration.

If there is a clock at point A of space, then an observer located at A can evaluate the time of the events in the immediate vicinity of A by finding the clock-hand positions that are simultaneous with these events. If there is also a clock at point B—we should add, "a clock of exactly the same constitution as that at A-–then the time of the events in the immediate vicinity of B can likewise be evaluated by an observer located at B. But it is not possible to compare the time of an event at A with one at B without a further stipulation; thus far we have only defined an "A-time" and a "B-time" but not a "time" common to A and B. The latter can now be determined by establishing *by definition* that

---

[1] We shall not discuss here the imprecision that is inherent in the concept of simultaneity of two events taking place at (approximately) the same location and that also must be surmounted by an abstraction.

the "time"needed for the light to travel from A to B is equal to the "time"it needs to travel from B to A. For, suppose a ray of light leaves from A toward B at "A-time"t_A, is reflected from B toward A at "B-time"t_B, and arrives back at A at "A-time"t'_A. The two clocks are synchronous by definition if

$$t_B - t_A = t'_A - t_B.$$

We assume that it is possible for this definition of synchronism to be free of contradictions, and to be so for arbitrarily many points, and that the following relations are therefore generally valid:

1. If the clock in B is synchronous with the clock in A, then the clock in A is synchronous with the clock in B.

2. If the clock in A is synchronous with the clock in B as well as with the clock in C, then the clocks in B and C are also synchronous relative to each other.

With the help of some physical (thought) experiments, we have thus laid down what is to be understood by synchronous clocks at rest that are situated at different places, and have obviously obtained thereby a definition of "synchronous"and of "time."The "time"of an event is the reading obtained simultaneously with the event from a clock at rest that is located at the place of the event and that for all time determinations is in synchrony with a specified clock at rest.

Based on experience, we also postulate that the quantity

$$\frac{2\overline{AB}}{t'_A - t_A} = V$$

is a universal constant (the velocity of light in empty space).

It is essential that we have defined time by means of clocks at rest in a system at rest; because it belongs to the system at rest, we designate the time just defined as "the time of the system at rest."

§2. On the relativity of length and times

The considerations that follow are based on the principle of relativity and the principle of the constancy of the velocity of light, two principles that we define as follows:

1. The laws governing the changes of the state of any physical system do not depend on which one of two coordinate systems in uniform translational motion relative to each other these changes of the state are referred to.

2. Each ray of light moves in the coordinate system "at rest"with the definite velocity V independent of whether this ray of light is emitted by a body at rest or a body in motion. Here,

$$velocity = \frac{light\ path}{time\ interval},$$

where "time interval"should be understood in the sense of the definition in §1.

Let there be given a rigid rod at rest; its length, measured by a measuring rod that is also at rest, shall be l. We now imagine that the axis of the rod is placed along the X-axis of the coordinate system at rest, and that the rod is then set in uniform parallel translational motion (velocity v) along the X-axis in the direction of increasing x. We now seek to determine the length of the *moving* rod, which we imagine to be obtained by the following two operations:

a The observer co-moves with the above-mentioned measuring rod and the rod to be measured, and measures the length of the rod directly, by applying the measuring rod exactly as if the rod to be measured, the observer, and the measuring rod were at rest.

b Using clocks at rest that are set up in the system at rest and are synchronous in the sense of §1, the observer determines the points of the system at rest at which the beginning and the end of the rod to be measured are found at some given time t. The distance between these two points, measured by the rod used before, which in the present case is at rest, is also a length, which can be designated as the "length of the rod."

According to the principle of relativity, the length to be found in operation (a), which we shall call "the length of the rod in the moving system,"must equal the length l of the rod at rest.

We will determine the length to be found in operation (b), which we shall call "the length of the (moving) rod in the system at rest,"on the basis of our two principles, and will find it to be different from l.

The commonly used kinematics tacitly assumes that the lengths determined by the two methods mentioned are exactly identical, or, in other words, that in the time epoch t a moving rigid body is totally replaceable, as far as geometry is concerned, by the same body when it is at rest in a particular position.

Further, we imagine that the two ends (A and B) of the rod are equipped with clocks that are synchronous with the clocks of the system at rest, i.e., whose readings always correspond to the "time of the system at rest"at the locations they happen to occupy; hence, these clocks are "synchronous in the system at rest."

We further imagine that each clock has an observer co-moving with it, and that these observers apply to the two clocks the criterion for synchronism formulated in §1. Suppose a ray of light starts out from A at time[2] t_A, is reflected from B at time t_B, and arrives back at A at time t'_A. Taking into account the principle of the constancy of the velocity of light, we find that

$$t_B - t_A = \frac{r_{AB}}{V - v}$$

---

[2]Time"here means both "time of the system at rest"and "the position of the hands of the moving clock located at the place in question.

and

$$t'_A - t_B = \frac{r_{AB}}{V + v},$$

where r_{AB} denotes the length of the moving rod, measured in the system at rest. The observers co-moving with the moving rod would thus find that the two clocks do not run synchronously while the observers in the system at rest would declare them synchronous.

Thus we see that we must not ascribe *absolute* meaning to the concept of simultaneity; instead, two events that are simultaneous when observed from some particular coordinate system can no longer be considered simultaneous when observed from a system that is moving relative to that system.

§3. Theory of transformation of coordinates and time from a system at rest to a system in uniform translational motion relative to it

Let there be given two coordinate systems in the space "at rest,"i.e., two systems of three mutually perpendicular rigid material lines issuing from one point. Let the X-axes of the two systems coincide and their Y- and Z-axes be parallel. Each system shall be supplied with a rigid measuring rod and a number of clocks, and the two measuring rods and all the clocks of the two systems should be exactly alike.

The origin of one of the two systems (k) shall now be imparted a (constant) velocity v in the direction of increasing x of the other system (K), which is at rest, and this velocity shall also be imparted to the coordinate axes, the corresponding measuring rod, and the clocks. To each time t of the system at rest K there corresponds then a definite position of the axes of the moving system, and for reasons of symmetry we may right- fully assume that the motion of k can be such that at time t ("t"always denotes a time of the system at rest) the axes of the moving system are parallel to the axes of the system at rest.

We now imagine the space to be measured both from the system at rest K by means of the measuring rod at rest and from the moving system k by means of the measuring rod moving along with it, and that the coordinates x, y, z and ξ, η, ζ are obtained in this way. Further, by means of the clocks at rest in the system at rest and using light signals in the manner described in §1, the time t of the system at rest is determined for all its points where there is a clock; likewise, the time τ of the moving system is determined for all the points of the moving system having clocks that are at rest relative to this system, applying the method of light signals described in §1 between the points containing these clocks.

To every system of values x, y, z, t that determines completely the place and time of an event in the system at rest, there corresponds a system of values ξ, η, ζ, τ that fixes this event relative to the system k, and the problem to be solved is to find the system of equations connecting these quantities.

First of all, it is clear that these equations must be *linear* because of the

properties of homogeneity that we attribute to space and time.

If we put $x' = x - vt$, then it is clear that a point at rest in the system k has a definite, time-independent system of values x', y, z belonging to it. We first determine τ as a function of x', y, z, and t. To this end, we must express in equations that τ is in fact the aggregate of the readings of the clocks at rest in the system k, which have been synchronized according to the rule given in §1

Suppose that at time τ_0 a light ray is sent from the origin of the system k along the X-axis to x' and is reflected from there at time τ_1 toward the origin, where it arrives at time τ_2; we then must have

$$\frac{1}{2}(\tau_0 + \tau_2) = \tau_1,$$

or, if we write out the arguments of the function τ and apply the principle of the constancy of the velocity of light in the system at rest,

$$\frac{1}{2}\left[\tau(0,0,0,t) + \tau\left(0,0,0,\left\{t + \frac{x'}{V-v} + \frac{x'}{V+v}\right\}\right)\right] = \tau\left(x',0,0,t + \frac{x'}{V-v}\right).$$

From this we get, if x' is chosen infinitesimally small,

$$\frac{1}{2}\left(\frac{1}{V-v} + \frac{1}{V+v}\right)\frac{\partial \tau}{\partial t} = \frac{\partial \tau}{\partial x'} + \frac{1}{V-v}\frac{\partial \tau}{\partial t},$$

or

$$\frac{\partial \tau}{\partial x'} + \frac{v}{V^2 - v^2}\frac{\partial \tau}{\partial t} = 0.$$

t should be noted that, instead of the coordinate origin, we could have chosen any other point as the starting point of the light ray, and the equation just derived therefore holds for all values of x', y, z.

Analogous reasoning—applied to the H and Z axes—yields, if we consider that light always propagates along these axes with the velocity $\sqrt{V^2 - v^2}$ when observed from the system at rest,

$$\frac{\partial \tau}{\partial y} = 0$$

$$\frac{\partial \tau}{\partial z} = 0$$

These equations yield, since τ is a *linear* function,

$$\tau = a\left(t - \frac{v}{V^2 - v^2}x'\right),$$

where a is a function $\varphi(v)$ as yet unknown, and where we assume for brevity that at the origin of k we have $t = 0$ when $\tau = 0$.

Using this result, we can easily determine the quantities ξ, η, ζ by expressing in equations that (as demanded by the principle of the constancy of the

velocity of light in conjunction with the principle of relativity) light propagates with velocity V also when measured in the moving system. For a light ray emitted at time $\tau = 0$ in the direction of increasing ξ, we will have

$$\xi = V\tau, \tag{1}$$

or

$$\xi = aV\left(t - \frac{v}{V^2 - v^2}x'\right).$$

But as measured in the system at rest, the light ray propagates with velocity $V - v$ relative to the origin of k, so that

$$\frac{x'}{V - v} = t.$$

Substituting this value of t in the equation for ξ, we obtain

$$\xi = a\frac{V^2}{V^2 - v^2}x'.$$

Analogously, by considering light rays moving along the two other axes, we get

$$\eta = V\tau = aV\left(t - \frac{v}{V^2 - v^2}x'\right),$$

where

$$\frac{y}{\sqrt{V^2 - v^2}} = t; \quad x' = 0;$$

hence

$$\eta = \frac{V}{\sqrt{V^2 - v^2}}y$$

and

$$\zeta = \frac{V}{\sqrt{V^2 - v^2}}z.$$

If we substitute for x' its value, we obtain

$$\tau = \varphi(v)\beta\left(t - \frac{v}{\sqrt{V^2}}x\right),$$
$$\xi = \varphi(v)\beta(x - vt),$$
$$\eta = \varphi(v)y,$$
$$\zeta = \varphi(v)z,$$

where

$$\beta = \frac{1}{\sqrt{1 - \left(\frac{v}{V}\right)^2}}$$

and φ is a function of v that is as yet unknown. If no assumptions are made regarding the initial position of the moving system and the zero point of τ, then an additive constant must be attached to the right-hand sides of these equations.

Now we have to prove that every light ray measured in the moving system propagates with the velocity V, if it does so, as we have assumed, in the system at rest; for we have not yet provided the proof that the principle of the constancy of the velocity of light is compatible with the relativity principle.

Suppose that at time $t = \tau = 0$ a spherical wave is emitted from the coordinate origin, which is at that time common to the two systems, and that this wave propagates in the system K with the velocity V. Hence, if (x, y, z) is a point just reached by this wave, we will have

$$x^2 + y^2 + z^2 = V^2 t^2.$$

We transform these equations using our transformation equations, and, after a simple calculation, obtain

$$\xi^2 + \eta^2 + \zeta^2 = V^2 \tau^2.$$

Thus, the wave under consideration is a spherical wave of propagation velocity V also when it is observed in the moving system. This proves that our two fundamental principles are compatible.

The transformation equations we have derived also contain an unknown function φ of v, which we now wish to determine.

To this end we introduce a third coordinate system K', which relative to the system k is in parallel-translational motion parallel to the axis Ξ such that its origin moves along the Ξ-axis with velocity $-v$. Let all three coordinate origins coincide at time $t = 0$, and let the time t' of the system K' be zero at $t = x = y = z = 0$. We denote the coordinates measured in the system K' by x', y', z' and, by twofold application of our transformation equations, we get

$$
\begin{aligned}
t' &= \varphi(-v)\beta(-v)\left\{\tau + \frac{v}{V^2}\xi\right\} &&= \varphi(v)\varphi(-v)t, \\
x' &= \varphi(-v)\beta(-v)\{\xi + v\tau\} &&= \varphi(v)\varphi(-v)x, \\
y' &= \varphi(-v)\eta &&= \varphi(v)\varphi(-v)y, \\
z' &= \varphi(-v)\zeta &&= \varphi(v)\varphi(-v)z.
\end{aligned}
$$

Since the relations between x', y', z' and x, y, z do not contain the time t, the systems K and K' are at rest relative to each other, and it is clear that the transformation from K to K' must be the identity transformation. Hence,

$$\varphi(v)\varphi(-v) = 1.$$

Let us now explore the meaning of $\varphi(v)$. We shall focus on that portion of the H-axis of the system k that lies between $\xi = 0$, $\eta = 0$, $\zeta = 0$, and $\xi = 0$, $\eta = l$, $\zeta = 0$. This portion of the H-axis is a rod that moves perpendicular to its axis

with a velocity v relative to the system K and whose ends possess in K the coordinates

$$x_1 = vt, \quad y_1 = \frac{l}{\varphi(0)}, \quad z_1 = 0$$

and

$$x_2 = vt, \quad y_2 = 0, \qquad z_2 = 0$$

The length of the rod, measured in K, is thus $l/p(v)$; this establishes the meaning of the function φ. For reasons of symmetry it is obvious that the length of a rod measured in the system at rest and moving perpendicular to its own axis can depend only on its velocity and not on the direction and sense of its motion. Thus, the length of the moving rod measured in the system at rest does not change when v is replaced by $-v$. From this we arrive at

$$\frac{l}{\varphi(v)} = \frac{l}{\varphi(-v)},$$

or

$$\varphi(v) = \varphi(-v)$$

It follows from this relation and the one found before that $\varphi(v)$ must equal 1, so that the transformation equations obtained become

$$\tau = \beta\left(t - \frac{v}{\sqrt{V^2}}x\right),$$
$$\xi = \beta(x - vt),$$
$$\eta = y,$$
$$\zeta = z,$$

where

$$\beta = \sqrt{1 - \left(\frac{v}{V}\right)^2}.$$

§4. The physical meaning of the equations obtained concerning moving rigid bodies and moving clocks

We consider a rigid sphere[3] of radius R that is at rest relative to the moving system k and whose center lies at the origin of k. The equation of the surface of this sphere, which moves with velocity v relative to the system k, is

$$\xi^2 + \eta^2 + \zeta^2 = R^2.$$

Expressed in x, y, z, the equation of this surface at time $t = 0$ is

$$\frac{x^2}{\left(1 - \left(\frac{v}{V}\right)^2\right)^2} + y^2 + z^2 = R^2.$$

---

[3]I.e., a body possessing the shape of a sphere when investigated at rest.

A rigid body that has a spherical shape when measured in the state of rest thus in the state of motion—observed from a system at rest—has the shape of an ellipsoid of revolution with axes

$$\frac{x^2}{\left(\sqrt{1-\left(\frac{v}{V}\right)^2}\right)^2} + y^2 + z^2 = R^2.$$

$$R\sqrt{1-\left(\frac{v}{V}\right)^2}, R, R.$$

Thus, while the Y and Z dimensions of the sphere (and hence also of every rigid body, whatever its shape) do not appear to be altered by motion, the X dimension appears to be contracted in the ratio $1 : \sqrt{1-(v/V)^2}$, i.e., the greater the value of v, the greater the contraction. At $v = V$, all moving objects—observed from the system "at rest-—shrink into plane structures. For superluminal velocities our considerations become meaning-less; we shall see in the considerations that follow that in our theory the velocity of light physically plays the part of infinitely great velocities.

It is clear that the same results apply for bodies at rest in a system "at rest"that are observed from a uniformly moving system.

We further imagine that one of the clocks that is able to indicate time t when at rest relative to the system at rest and time τ when at rest relative to the system in motion, is placed in the origin of k and set such that it indicates the time τ. What is the rate of this clock when observed from the system at rest?

The quantities x, t, and τ, which refer to the position of this clock, are obviously related by the equations

$$\tau = \sqrt{1-\left(\frac{v}{v}\right)^2}(t - \frac{v}{V^2}x)$$

and

$$x = vt$$

We thus have

$$\tau = t\sqrt{1-\left(\frac{v}{V}\right)^2} = t - \left(1 - \sqrt{1-\left(\frac{v}{V}\right)^2}\right)t$$

which shows that the clock (observed in the system at rest) is retarded each second by $(1 - \sqrt{1-(v/V^2)})$ sec or, with quantities of the fourth and higher orders neglected, by $\frac{1}{2}(v/V)^2$ sec.

This yields the following peculiar consequence: If at the points A and B of K there are located clocks at rest which, observed in a system at rest, are synchronized, and if the clock in A is transported to B along the connecting

line with velocity v, then upon arrival of this clock at B the two clocks will no longer be synchronized; instead, the clock that has been transported from A to B will lag $\frac{1}{2}tv^2/V^2$ sec (up to quantities of the fourth and higher orders) behind the clock that has been in B from the outset, if t is the time needed by the clock to travel from A to B.

We see at once that this result holds even when the clock moves from A to B along any arbitrary polygonal line, and even when the points A and B coincide.

If we assume that the result proved for a polygonal line holds also for a continuously curved line, then we arrive at the following proposition: If there are two synchronous clocks in A, and one of them is moved along a closed curve with constant velocity until it has returned to A, which takes, say, t sec, then this clock will lag on its arrival at A $\frac{1}{2}(v/V)^2$ sec behind the clock that has not been moved. From this we conclude that a balance-wheel clock that is located at the Earth's equator must be very slightly slower than an absolutely identical clock, subjected to otherwise identical conditions, that is located at one of the Earth's poles.

§5. The addition theorem of velocities

In the system k moving with velocity v along the X-axis of the system K let there be a point moving according to the equation

$$\xi = w_\xi \tau$$
$$\eta = w_\eta \tau$$
$$\zeta = 0$$

where w_ξ and w_η denote constants.

We seek the motion of the point relative to the system K. Introducing the quantities x, y, z, t into the equations of motion of the point by means of the transformation equations derived in §3, we obtain