LECTURE 3: SPACETIME AND GEOMETRY:
AN INTRODUCTION TO SPECIAL RELATIVITY

AS204 February 11 2008

I. The Geometry of Space

You are used to dealing with vectors in ordinary 3-dimensional space, and to specify locations in space by means of *coordinates*, like ordinary x, y, z *Cartesian* coordinates, though there are, of course, many other more complex coordinate systems, such as *Cylindrical* — r, θ, z and *Sphericopolar* — r, θ, ϕ, and many, many others. The prototype of a vector is the entity which connects neighboring points in space:

$$\mathbf{dr} = (dx, dy, dz).$$

Associated with this vector (or any vector) is its *length*, the square of which is

$$ds^2 = dx^2 + dy^2 + dz^2,$$

and an associated notion of *dot product* of two vectors $\mathbf{dr} = (dx, dy, dz)$ and $\mathbf{u} = (u, v, w)$,

$$\mathbf{dr} \cdot \mathbf{u} = u\ dx + v\ dy + w\ dz.$$

It is important to understand that vectors have existence quite independent of the coordinates used to describe them; their dot products and their lengths are geometric properties of the vectors themselves which must be calculable in any coordinate system and must be independent of the coordinate system used to describe them. In particular, different cartesian coordinate systems are related to each other by a *translation* — *i. e.* a choice of origin —and a *rotation*— a choice of the directions of the axes. The latter is not quite arbitrary, of course, since ordinary cartesian coordinates are *orthogonal* — *i. e.* the coordinate axes are mutually perpendicular. The transformation from one cartesian coordinate system to another related by an arbitrary rotation is called an orthogonal transformation:

$$(u') = O(u)$$

where (u) is the array of the components of the vector **u** in a particular coordinate system, and (u') the array of the components of the **same vector u** in the new rotated coordinate system: O is a 3×3 matrix, which, if and only if it is a proper orthogonal transformation, has the property that the matrix product of O and its transpose is the identity:

$$O^T O = I$$

where I is the identity matrix, $I_{ab} = 1$ for $a = b; 0$, for $a \neq b$.

It is easy to see why $O^T O = I$ is the defining property of orthogonal transformations:

Let $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{x} = (x_1, x_2, x_3)$ be two vectors as described in some coordinate system. Their dot product in the original coordinates is calculated by

$$\mathbf{u} \cdot \mathbf{x} = u_1 x_1 + u_2 x_2 + u_3 x_3 = u_a x_a,$$

where we introduce the shorthand convention that repeated *latin* indices are summed over from 1 to 3, *i. e.* $u_a x_a = \sum_{i=1,3} u_a x_a$. Since $u'_a = O_{ab} u_b$, and $x'_c = O_{cd} x_d$ we must have

$$\mathbf{u} \cdot \mathbf{x} = u_a x_a = u'_b x'_b = O_{bc} u_c O_{bd} x_d$$

Since u and x are completely arbitrary, it must be that, since the first and last expressions must be equal, $O_{bc} O_{bd}$ must be 1 for $c = d$ and zero if $c \neq d$, *i. e.*, the identity. but $O_{bc} O_{bd}$ is just the matrix product $O^T O$.

Enough. Let's talk about a famous experiment which was not *supposed* to be an experiment about geometry but turned out to be.

II. Michelson-Morley and the Death of the Ether

The classical experiment which sounded the death knell of Newtonian ideas of space and time was the Michelson-Morley experiment, which was designed to measure the velocity of the earth with respect to the ether, a hypothetical substance which was imagined to supported the propagation of electromagnetic waves and was supposed to fill all of space. Michelson and Morley designed an L-shaped interferometer which could measure the difference in the propagation time of light along the two equal-length arms of the device. It is easy to show that if the device is moving with velocity v in the direction of one of the arms, the time which light requires to move from the intersection of the arms to a mirror on the end of that arm and back is

$$t_{\parallel} = \frac{2L}{c(1 - v^2/c^2)} \tag{1}$$

and this is the same whether the velocity is in the direction the arm points or opposite it. The time required for the light to travel from the intersection to a mirror on the end of the perpendicular arm and back is different,

$$t_{\perp} = \frac{2L}{c(1 - v^2/c^2)^{1/2}} \tag{2}$$

Now the orbital velocity of the earth around the sun is 30 km/sec. If we expand these expressions noting that v^2/c^2 is very small (about 10^{-8} if the solar system is not moving too fast with respect to the ether), we find

$$t_{\parallel} \simeq \frac{2L}{c}\left(1 + \frac{v^2}{c^2}\right), \tag{3}$$

while

$$t_\perp \simeq \frac{2L}{c}\left(1+\frac{v^2}{2c^2}\right). \quad (4)$$

The difference is $\Delta t \simeq Lv^2/c^3$, which if v is the orbital velocity of the earth, is about 3.3×10^{-17} seconds for an apparatus with arms of length $L = 1$m, a very tiny interval. Light in this time travels about 10^{-6} cm, which is also small, but is about a fiftieth of the wavelength of ordinary green light. Michelson's and Morley's apparatus in fact was an *interferometer*, which had a light source and a beamsplitter, which allowed the light from the two arms to interfere upon its return to the intersection, and shifts of this magnitude can be measured with some care. (For the practically minded among you, the experimental technique actually involved nulling the interferometer–so that the difference in travel times along the two arms was an integral number of periods of the monochromatic light used—and then rotating the whole apparatus, which sat in a vat of mercury, 90 degrees, thus switching the roles of the two arms of the device. Remembering that the velocity of the sun about the center of the Galaxy is about 250 km/sec, the signal would have been expected to be even a hundred times larger, perhaps.)

The result was very surprising. The velocity was *always* zero, no matter where the earth was in its orbit–at times separated by 6 months, the velocity should have been 60 km/s different, no matter what it was the first time–but the result was always that the experiment was at rest with respect to the ether. It became more and more evident that *absolute motion* was something that could not be measured. Various *ad hoc* ideas were put forward which could explain the phenomenon, but none were very natural, and the simplicity of the results demanded a simple explanation. Einstein provided one.

II. The Geometry of Spacetime

Einstein's incredible realization that space and *time* together could be regarded as an arena for *geometry* allowed him to think about the problems with conventional Newtonian physical ideas which were emerging in the early 1900s in an entirely new way. Many of the things he predicted and described had already been put forward by Lorenz and others, but only as *ad hoc* ideas, devoid of any real framework in which to think about them and extend them. Einstein provided that framework, but, interestingly, evidently never learned to think about it as having geometrical reality beyond its mathematical description. We owe this development, which not only allows the development of powerful intuition but also leads often to very simple solutions of otherwise very difficult problems, largely to John Wheeler and his outstanding students, including most notably Feynman, Misner, and Thorne.

To think about this new idea, let us consider the four dimensional world– three dimensions of ordinary space and one of time. The atomic notion in this space that corresponds to a *point* in 3-space is an *event*— a point in space and an instant of time, which describes both *where* and *when*. In Newtonian physics, there is 3-space, and there is time, but they are separate and not part of any unified geometrical system. Einstein realized that space and time were inseparable, and indeed different manifestations of the same thing, in a sense which I hope will become clear as we go on.

We can describe an event by a four numbers: $E = (t, x, y, z)$. Just as there are vectors in 3-space which have a direction and a length, and are typified by the separations of neighboring points, there are *4-vectors* in spacetime which are typfied by the separation of neighboring *events*:

$$\vec{dx} = (dt, dx, dy, dz).$$

is the 4-vector which connects (t, x, y, z) with $(t+dt, x+dx, y+dy, z+dz)$. So far this is not very profound. The profundity begins when we assign a *length* to a 4-vector, and specify what rules it obeys. Einstein saw that several violations of Newtonian physics which had either been observed or suspected could be explained if the geometry of spacetime were described by a geometrical length-like quantity (the **interval**), the square of which is

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2, \tag{5}$$

which Einstein hypothesized was the same in every coordinate system—that is to say, to all observers. We will see what that means presently.

But wait—everybody knows we measure time in seconds or some such and distance in centimeters or meters, so the equation (5) makes no sense, because quantities added or subtracted from each other must have the same units, no?

Yes, which means that there is some special relation between intervals of time and distances in space. There is in the universe a special *velocity*, namely the velocity of light, which is the propagation velocity of all electromagnetic waves and is fundamental to all physics. So we can either write

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2, \tag{5a}$$

if we want to measure the interval in length units, or

$$ds^2 = dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2), \tag{5b}$$

if we want to measure the interval in time units, or we can be clever and measure time in seconds (say) and length in light-seconds (or years and light-years, or any other such) and stick with the simplicity of equation (5). This is more than convenience; it stresses that the geometry is unified–who ever heard of measuring different axes with different units?

It is perhaps worth an aside here—why the velocity of *light*? Why not some other? Sound? The average walking speed of a sacred bull? Well, first of all, it works—if Einstein had hypothesized that there was *some* special velocity then to explain the results of the Michelson-Morley experiment that velocity would have had to be the velocity of light. But the reasons are far deeper. Physics at that time consisted of Newtonian-Lagrangian-Hamiltonian mechanics, a magnificent mathematical and practical edifice which had *no* special velocities built into it, the first whispers of quantum mechanics, and Maxwell's electromagnetic theory, into every fiber of which was woven the velocity of light. Einstein and other workers quickly realized that though Newtonian physics had to be modified to understand the Michelson-Morley experiment, Maxwellian electromagnetism did *not*—it was already a relativistic theory, even though the framework into which it fit had not been known at the time Maxwell wrote down his equations.

It is also clear that the reason that hundreds of years went by before anyone noticed any departures from Newtonian physics is that the velocity of light is so large that all ordinary velocities are very small with respect to it. What finally tipped people off, of course, was not the ability to accelerate particles to very high velocities and thereby observe the completely bizarre effects which occur, (which we do routinely today) but the very accurate measurements in the Michelson-Morley experiment involving the effect of the earth's orbital velocity around the sun, which is a ten-thousandth of the velocity of light.

Let us return to Equation (5). Consider two events–one, the emission of a photon from the source at the center of the interferometer, the second, the reflection of the photon from the mirror on the leading arm. Suppose the arm is along the x axis, and the apparatus is at rest. If we put the origin at the center and the photon is emitted at time t_0, it arrives at the mirror at time $t_0 + L/c$, $x = L$. The interval between these two events is

$$s^2 = c^2t^2 - L^2 = 0.$$

The tenet of special relativity is that the interval is *invariant*– that is, it is measured to be the same by ALL observers. That means that if there is another observer moving with some velocity v, so that in *her* coordinate system $x' = x - vt$, she will also get zero for the interval between these events. This clearly has the staggering implication that *the velocity of light is the same to all observers, no matter what the state of their relative motion.* Two events which occur along a light ray always have zero interval, and that means that any observer measures a time difference and a separation which corresponds to the same velocity, unity in our clever coordinates and c in ordinary ones. This explains the Michelson-Morley experiment, because the velocity of light is not affected by the motion; it is not additive to some velocity with respect to some fixed ether.

Other sacred or at least completely taken–for–granted ideas fall by the wayside. Consider a Michelson-Morley arrangement with two arms at 180 degrees from each other. If the arms are the same length, a flash of light from the center arrives at the ends of the two arms simultaneously. To an observer which is moving with respect to the apparatus along its axis at velocity v, the travel time along the leading arm is $L/(c - v)$ and along the trailing one is $L/(c + v)$, and thus the two events which are *simultaneous* to the observer at rest with respect to the apparatus are *not* to the moving observer. *Simultaneity has no meaning for events separated in space. There is no absolute time.*

Now consider the original experiment and an observer moving at velocity v along one of the arms. A flash of light from the center goes to the ends of the arms and bounces back and arrives at the center simultaneously– same place, same time. All observers must clearly see this event as simultaneous; the experimenter could light a green "success" light if they arrive simultaneously at his detector and a red "failure" light if not. Either the green light or the red light comes on. The green light comes on. So what is wrong with our calculation? The only way the moving observer can get the parallel photon where it belongs at the right time is for the moving parallel arm to be *shorter*, $L_m = L_0\sqrt{1 - v^2/c^2}$. This is called *Lorenz contraction.* We see that *there is no absolute space.*

In the next lecture we will learn how to calculate some of these effects simply, and learn a bit more about the far-reaching changes to physics that the idea of relativity entails.