PHY 564
Advanced Accelerator Physics
Lecture 2
Particles In Electromagnetic Fields

Vladimir N. Litvinenko
Yichao Jing
Gang Wang

CENTER for ACCELERATOR SCIENCE AND EDUCATION
Department of Physics & Astronomy, Stony Brook University
Collider-Accelerator Department, Brookhaven National Laboratory

In likely event of us running out of time in class – the home reading is paragraphs 1-7 of Classical theory of fields of L.D. Landau and E.M. Lifshitz

Hamiltonian formalism

Fundamentals of Hamiltonian Mechanics

http://en.wikipedia.org/wiki/Hamilton_principle

1.0. Least-Action Principle and Hamiltonian Mechanics

Let us refresh our knowledge of some aspects of **the Least-Action Principle** (LAP is humorously termed the *coach potato principle*) **and Hamiltonian Mechanics**. **The Principle of Least Action** is the most general formulation of laws governing the motion (evolution) of systems of particles and fields in physics. In mechanics, it is known as **the Hamilton's Principle,** and states the following**:**

1) A mechanical system with n degrees of freedom is fully characterized by a monotonic generalized coordinate, ***t***, the full set of n coordinates $q = \{q_1, q_2, q_3 \ldots q_n\}$ and their derivatives $\dot{q} = \{\dot{q}_1, \dot{q}_2, \dot{q}_3 .. \dot{q}_n\}$ that are denoted by dots above a letter. We study the dynamics of the system with respect to ***t***. All the coordinates, $q = \{q_1, q_2, q_3 \ldots q_n\}$; $\dot{q} = \{\dot{q}_1, \dot{q}_2, \dot{q}_3 .. \dot{q}_n\}$ should be treated as a functions of ***t*** that itself should be treated as an independent variable.

2) Each mechanical system can be fully characterized by the **Action Integral**:

$$S(A,B) = \int_A^B L(q,\dot{q},t)dt \tag{1}$$

that is taken between two events A and B described by full set of coordinates * (q,t). The function under integral $L(q,\dot{q},t)$ is called the system's **Lagrangian function.** Any system is fully described by its action integral.

* For one particle, the full set of event coordinates is the time and location of the particle. The integral is taken along a particle's world line (its unique path through 4-dimentional space-time) and is a function of both the end points and the intervening trajectory.

After that, applying **Lagrangian mechanics involves just** n second -order ordinary differential equations**:** $\ddot{q} = f(q,\dot{q})$.

We can find these equations, setting variation of δS_{AB} to zero:

$$\delta S_{AB} = \delta\left(\int_A^B L(q,\dot{q},t)dt\right) = \int_A^B \left\{\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q}\right\}dt = \int_A^B \left\{\frac{\partial L}{\partial q}\delta q dt + \frac{\partial L}{\partial \dot{q}}\delta dq\right\} =$$

$$\left[\frac{\partial L}{\partial \dot{q}}\delta q\right]_A^B + \int_A^B \left\{\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right\}\delta q dt = 0 \qquad ; \qquad (2)$$

and taking into account $\delta q(A) = \delta q(B) = 0$. Thus, we have integral of the function in the brackets, multiplied by an arbitrary function $\delta q(t)$ equals zero.

Therefore, we must conclude that the function in the brackets also equals zero and thus obtain **Lagrange's equations:**

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = 0. \qquad (3)$$

Explicitly, this represents a set of n second-order equations

$$\frac{d}{dt}\frac{\partial L(q,\dot{q},t)}{\partial \dot{q}_i} = \frac{\partial L(q,\dot{q},t)}{\partial q_i} \qquad \Leftrightarrow \sum_{j=1}^{n}\left(\ddot{q}_j \frac{\partial^2 L(q,\dot{q},t)}{\partial \dot{q}_i \partial \dot{q}_j} + \dot{q}_j \frac{\partial^2 L(q,\dot{q},t)}{\partial \dot{q}_i \partial q_j} + \frac{\partial^2 L(q,\dot{q},t)}{\partial \dot{q}_i \partial t}\right) = \frac{\partial L(q,\dot{q},t)}{\partial q_i}.$$

The partial derivative of the Lagrangian over $\dot{q}$ is called generalized (**canonical**) momentum:

$$P^i = \frac{\partial L(q,\dot{q},t)}{\partial \dot{q}_i} \; or \; P = \frac{\partial L(q,\dot{q},t)}{\partial \dot{q}} \; ; \tag{4}$$

and the partial derivative of the Lagrangian over q is called the generalized force: $f^i = \dfrac{\partial L(q,\dot{q},t)}{\partial q_i}$: (4) can be rewritten in more familiar form: $\dfrac{dP^i}{dt} = f^i$. Then, by a definition, the energy (Hamiltonian) of the system is:

$$H = \sum_{i=1}^{n} P^i \dot{q}_i - L \equiv \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L; L = \sum_{i=1}^{n} P^i \dot{q}_i - H. \tag{5}$$

Even though the Lagrangian approach fully describes a mechanical system it has some significant limitations. It treats the coordinates and their derivatives differently, and allows only coordinate transformations $q' = q'(q,t)$. There is more powerful method, the **Hamiltonian or Canonical Method**. The Hamiltonian is considered as a function of coordinates and momenta, which are treated equally. Specifically, pairs of coordinates with their conjugate momenta (4) (q_i,P_i) or (q^i,P_i) are called canonical pairs. The Hamiltonian method creates many links between classical and quantum theory wherein it becomes an operator. Before using the Hamiltonian, let us prove that it is really function of (q,P,t) : i.e., that the full differential of the Hamiltonian is

$$dH(q,P,t) = \sum_{i=1}^{n}\left(\frac{\partial H}{\partial q_i}dq_i + \frac{\partial H}{\partial P^i}dP^i\right) + \frac{\partial H}{\partial t}dt. \tag{6}$$

Using equation (5) explicitly, we can easily prove it:

$$dH = d\sum_{i=1}^{n}\frac{\partial L}{\partial \dot q_i}\dot q_i - dL \equiv \sum_{i=1}^{n}\left\{\frac{\partial L}{\partial \dot q_i}d\dot q_i + \dot q_i d\frac{\partial L}{\partial \dot q_i} - \frac{\partial L}{\partial q_i}dq_i - \frac{\partial L}{\partial \dot q_i}d\dot q_i - \frac{\partial L}{\partial t}dt\right\} =$$

$$\sum_{i=1}^{n}\left\{\dot q_i d\frac{\partial L}{\partial \dot q_i} - \frac{\partial L}{\partial q_i}dq_i - \frac{\partial L}{\partial t}dt\right\} = \sum_{i=1}^{n}\left\{\dot q_i dP^i - \frac{\partial L}{\partial q_i}dq_i - \frac{\partial L}{\partial t}dt\right\} = \sum_{i=1}^{n}\left(\frac{\partial H}{\partial q_i}dq_i + \frac{\partial H}{\partial P^i}dP^i\right) + \frac{\partial H}{\partial t}dt.$$

wherein we substitute $d(\partial L/\partial \dot q_i) = dP^i$ with the expression for generalized momentum. In addition to this proof, we find some ratios between the Hamiltonian and the Lagrangian:

$$\left.\frac{\partial H}{\partial q_i}\right|_{P=const} = -\left.\frac{\partial L}{\partial q_i}\right|_{\dot q=const};\ \left.\frac{\partial H}{\partial t}\right|_{P,q=const} = -\left.\frac{\partial L}{\partial t}\right|_{q,\dot q=const};\ \dot q_i = \frac{dq_i}{dt} = \frac{\partial H}{\partial P^i};$$

wherein we should very carefully and explicitly specify what type of partial derivative we use. For example, the Hamiltonian is function of (q,P,t): thus, partial derivative on q must be taken with constant momentum and time. For the Lagrangian, we should keep $\dot q, t = const$ to partially differentiate on q .

The last ratio gives us the first Hamilton's equation, while the second one comes from Lagrange's equation (5-11):

$$\begin{aligned}&\dot q_i = \frac{dq_i}{dt} = \frac{\partial H}{\partial P^i};\\ &\frac{dP^i}{dt} = \frac{d}{dt}\frac{\partial L(q,\dot q,t)}{\partial \dot q_i} = \frac{dP^i}{dt} = \left.\frac{\partial L}{\partial q_i}\right|_{\dot q=const} = -\left.\frac{\partial H}{\partial q_i}\right|_{P=const};\end{aligned} \tag{7}$$

both of which are given in compact form below in (11).

Now, to state this in a formal way. The Hamiltonian or Canonical Method uses a Hamiltonian function to describe a mechanical system as a function of coordinates and momenta:

$$H = H(q,P,t) \tag{8}$$

Then using eq. (5), we can write the action integral as

$$S = \int_A^B \left(\sum_{i=1}^n P^i \frac{dq_i}{dt} - H(q,P,t) \right) dt = \int_A^B \left(\sum_{i=1}^n P^i dq_i - H(q,P,t) dt \right); \tag{9}$$

The total variation of the integral can be separated into the variation of the end points, and the variation of the integral argument:

$$\delta \int_A^B f(x)dt = \int_{A+\delta A}^{B+\delta B} f(x+\delta x)dt - \int_A^B f(x)dt = \int_B^{B+\delta B} f(x+\delta x)dt + \int_{A+\delta A}^{A} f(x+\delta x)dt + \int_A^B f(x+\delta x)dt - \int_A^B f(x)dt =$$

$$= f(B)\Delta t_B - f(A)\Delta t_A + \int_A^B (f(x+\delta x) - f(x))dt; \;\; \Delta t_C = t(C+\delta C) - t(C); \;\; for C = A,B.$$

The first term represents the variation caused by a change of integral limits (events), while the second represents the variation of the integral between the original limits (events). The total variation of the action integral (9) can separated similarly:

$$\begin{aligned} \delta S &= \left[\sum_{i=1}^n P^i \Delta q_i - H\Delta t \right]_A^B + \int_A^B \left(\delta \sum_{i=1}^n P^i dq_i - \sum_{i=1}^n \left(\frac{\partial H}{\partial q_i} \delta q_i dt + \frac{\partial H}{\partial P^i} \delta P^i dt \right) \right) = \\ &= \left[\sum_{i=1}^n P^i \Delta q_i - H\Delta t \right]_A^B + \sum_{i=1}^n \int_A^B \left(\delta P^i dq_i + P^i d\delta q_i - \left(\frac{\partial H}{\partial q_i} \delta q_i dt + \frac{\partial H}{\partial P^i} \delta P^i dt \right) \right); \end{aligned} \tag{10}$$

This equation encompasses everything: The expressions for the Hamiltonian and the momenta through the action and **Hamiltonian equations of motion.** Now we consider variation in both the coordinates and momenta that are treated equally: $\delta q; \delta P$.

To find the equation of motion we set constant events and $\delta q(A) = \delta q(B) = 0;$ the first term disappears, and the minimal-action principle gives us

$$\delta S = \sum_{i=1}^{n} \int_A^B \left(\frac{\partial H}{\partial q_i} \delta q_i dt + \frac{\partial H}{\partial P^i} \delta P^i dt - \delta P^i dq_i - P^i d\delta q_i \right) = 0 ,$$

which, after integration by parts of the last term translates into

$$\delta S = \left[-\sum_{i=1}^{n} P^i \delta q_i \right]_A^B + \sum_{i=1}^{n} \int_A^B \left(\frac{\partial H}{\partial q_i} \delta q_i dt + \frac{\partial H}{\partial P^i} \delta P^i dt - \delta P^i dq_i + dP^i \delta q_i \right) =$$

$$= \sum_{i=1}^{n} \int_A^B \left(\left\{ \frac{\partial H}{\partial q_i} + \frac{dP^i}{dt} \right\} \delta q_i dt + \left\{ \frac{\partial H}{\partial P^i} - \frac{dq_i}{dt} \right\} \delta P^i dt \right) = 0;$$

where the variation of coordinates and momenta are considered to be independent. Therefore, both expressions in brackets must be zero at a real trajectory. This gives us the **Hamilton's equations of motion:**

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial P^i}; \quad \frac{dP^i}{dt} = -\frac{\partial H}{\partial q_i}. \tag{11}$$

It is easy to demonstrate that these equations are exactly equivalent to the Lagrange's equation of motion. This is not surprising because they are obtained from the same principle of least action and describe the motion of the same system.

Conservation laws, invariants

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial P^i};\ \frac{dP^i}{dt} = -\frac{\partial H}{\partial q_i}.$$

Let us also look at the full derivative of the Hamiltonian:

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_{i=1}^{n}\left(\frac{\partial H}{\partial P^i}\frac{dP^i}{dt} + \frac{\partial H}{\partial q_i}\frac{dq_i}{dt}\right) = \frac{\partial H}{\partial t} + \sum_{i=1}^{n}\left(-\frac{\partial H}{\partial q_i}\frac{\partial H}{\partial q_i} + \frac{\partial H}{\partial q_i}\frac{\partial H}{\partial P^i}\right) = \frac{\partial H}{\partial t}.$$

This equation means that the Hamiltonian is constant if it does not depend explicitly on ***t***. It is an independent derivation of energy conservation for closed system. The conservation of momentum is apparent from equation (11), viz., if the Hamiltonian does not depend explicitly on the coordinates, then momentum is constant. All these conservation laws result from the general theorem by **Emmy Noether** : *Any one-parameter group of dimorphisms operating in a phase space* $((q,\dot{q},t)$ *for Lagrangian* $((q,P,t)$ *for Hamiltonian) and preserving the Lagrangian/Hamiltonian function equivalent to existence of the (first order) integral of motion.* (Informally, it can be stated as, for every differentiable symmetry created by local actions there is a corresponding conserved current).

Returning to the Eq. (10), we now can consider motion **along real trajectories**. Here, the variation of the integral is zero and the connection between the action and the Hamiltonian variables is obtained by differentiation of the first term:

$$H = -\left.\frac{\Delta S}{\Delta t}\right|_{\Delta q=0} = -\frac{\partial S}{\partial t};\ P^i = \left.\frac{\Delta S}{\Delta q_i}\right|_{\Delta q_{k\neq i}=0} = \frac{\partial S}{\partial q_i};\quad S = \int\limits_{\substack{Along\ \ real\\ Trajectrory}} (P_i dq_i - Hdt); \tag{12}$$

Thus, knowing the action integral we can find the Hamiltonian and canonical (generalized) momenta from solving (12) without using the Lagrangian. All conservation laws emerge naturally from (10): if nothing depends on ***t*,** then ***H*** is conserved (i.e., the energy). If nothing depends on position, then the momenta are conserved: $P^i(A) = P^i(B)$. Finally, we write the Hamiltonian equations for one particle using the Cartesian frame:

$$S = \int \left(\vec{P}d\vec{r} - H(\vec{r},\vec{P},t)dt\right) \tag{13}$$

$$H(\vec{r},\vec{P},t) = -\frac{\partial S}{\partial t};\ \vec{P} = \frac{\partial S}{\partial \vec{r}};$$

$$\frac{d\vec{r}}{dt} = \frac{\partial H}{\partial \vec{P}};\qquad \frac{d\vec{P}}{dt} = -\frac{\partial H}{\partial \vec{r}};\qquad \frac{dH}{dt} = \frac{\partial H}{\partial t}$$

Canonical transformations

Hamiltonian method gives us very important tool – the general change of variables: $\{P_i, q_i\} \to \{\tilde{P}_i, \tilde{q}_i\}$, called **Canonical transformations**. From the least-action principle, two systems are equivalent if they differ by a full differential: (we assume the summation on repeating indices i=1,2,3, $a_i b_i \equiv \sum_i a_i b_i;\ a^\alpha b_\alpha \equiv \sum_\alpha a^\alpha b_\alpha$ and the use of co- and contra-variant vector components for the non-unity metrics tensor)

$$\delta \int P_i\, dq_i - Hdt = 0 \propto \delta \int \tilde{P}_i\, d\tilde{q}_i - \tilde{H}dt = 0 \to P_i dq_i - Hdt = \tilde{P}_i d\tilde{q}_i - \tilde{H}dt + dF \qquad (14)$$

where F is the so-called generating function of the transformation. Rewriting (14), reveals that $F = F(q_i, \tilde{q}_i, t)$:

$$dF = P_i dq_i - \tilde{P}_i d\tilde{q}_i + (H' - H)dt;\ \tilde{P}_i = -\frac{\partial F}{\partial q_i}; P_i = \frac{\partial F}{\partial \tilde{q}_i}; H' = H + \frac{\partial F}{\partial t}. \qquad (15)$$

In fact, generating functions on any combination of old coordinates or old momenta with new coordinates of new momenta are possible, totaling 4= 2 x 2 combinations:

Canonical transformations

$$F(q,\tilde{q},t) \quad \Rightarrow dF = P_i dq_i - \tilde{P}_i d\tilde{q}_i + (H' - H)dt;\; P_i = \frac{\partial F}{\partial q_i};\; \tilde{P}_i = -\frac{\partial F}{\partial \tilde{q}_i};\; H' = H + \frac{\partial F}{\partial t}.$$

$$\Phi(q,\tilde{P},t) = F + \tilde{q}_i\tilde{P}_i \quad \Rightarrow d\Phi = P_i dq_i + \tilde{q}_i d\tilde{P}_i + (H' - H)dt;\; P_i = \frac{\partial \Phi}{\partial q_i};\; \tilde{q}_i = \frac{\partial \Phi}{\partial \tilde{P}_i};\; H' = H + \frac{\partial \Phi}{\partial t};$$

$$\Omega(P,\tilde{q},t) = F - P_i q_i \quad \Rightarrow d\Omega = -q_i dP_i - \tilde{P}_i d\tilde{q}_i + (H' - H)dt;\; q_i = -\frac{\partial \Omega}{\partial P_i};\; \tilde{P}_i = -\frac{\partial \Omega}{\partial \tilde{q}_i};\; H' = H + \frac{\partial \Omega}{\partial t};$$

$$\Lambda(P,\tilde{P},t) = \Phi - P_i q_i \quad \Rightarrow d\Lambda = \tilde{q}_i d\tilde{P}_i - q_i dP_i + (H' - H)dt;\; q_i = -\frac{\partial \Lambda}{\partial P_i};\; \tilde{q}_i = \frac{\partial \Lambda}{\partial \tilde{P}_{ii}};\; H' = H + \frac{\partial \Lambda}{\partial t};$$

(15’)

The most trivial canonical transformation is $\tilde{q}_i = P_i;\ \tilde{P}_i = -q_i$ with trivial generation function of

$$F(q,\tilde{q}) = q_i\tilde{q}_i \quad P_i = \frac{\partial F}{\partial q_i} = \tilde{q}_i;\ \tilde{P}_i = -\frac{\partial F}{\partial \tilde{q}_i} = -\tilde{q}_i;\ H' = H$$

Hence, this is direct proof that in the Hamiltonian method the coordinates and momenta are treated equally, and that the meaning of canonical pair (and its connection to Poisson brackets) has fundamental nature.

The most non-trivial finding from the Hamiltonian method is that the motion of a system, i.e., the evolution of coordinates and momenta also entails a Canonical transformation:

$$q_i(t+\tau) = \tilde{q}_i(q_i(t),P_i(t),t);\ P_i(t+\tau) = \tilde{P}_i(q_i(t),P_i(t),t);$$

with generation function being the action integral along a real trajectory (12):

$$S = S = \int_A^{t+\tau}\left(P_i dq_i - Hdt\right) - \int_A^{t}\left(P_i dq_i - Hdt\right);$$

$$dS = P_i(t+\tau)dq_i - P_i(t)dq_i + (H_{t+\tau} - H_t)dt$$

Special relativity

1.1 Einstein principle of relativity.

There is nothing more un-natural than "non-relativistic" electrodynamics. And there are very few thing in our world as natural as relativistic electrodynamics. We can consider non-relativistic classical or quantum mechanics for objects which can rest or move slowly. But how we can describe electromagnetic wave without using speed of the light? which is the universal, as far as we know, physical constant:

$$\mathbf{c = 2.99792458(1.2)\cdot 10^{10}\, cm/sec}; \qquad (1\text{-}1)$$

The "**c**" does not depend on the system of reference . The standard non-relativistic Galileo's relativity principle claims

1. Free particle propagates with constant velocity (the law of inertia) $\vec{v} = conts$;

2. Time does not depend on the choice of inertial frame moving with velocity $\vec{V}$ with respect to initial frame of reference:

$$t = t'; \vec{r} = \vec{r}\,' + \vec{V}t \qquad (1\text{-}2)$$

and velocity transformation is

$$\vec{v} = \vec{v}\,' + \vec{V}. \qquad (1\text{-}3)$$

Many modern experimental facts disagree with Galileo's principle and confirm that:

The speed of the light does not depend of the reference frame.

Galileo assumed that we are leaving in Euclidean world. What is wrong in Galileo's principle is the assumption that time and distance between two points in 3-D space are absolute, i.e. independent from the reference frame.

In 1905 Einstein modified principle of relativity to satisfy new experimental data. The **Einstein principle of relativity** comprises of two postulates:

1. **POSTULATE OF RELATIVITY** (the same as Galileo):

The laws of nature and results of all experiments are independent of translational motion of the system (reference frame) as whole. Precisely: there are a triply infinite set of equivalent Euclidean (3D) reference frames moving with constant velocities in rectilinear paths relative to one other in which all physical phenomena occur in an identical manner.

2. **POSTULATE OF THE CONSTANCY OF THE SPEED OF THE LIGHT** (Einstein):

The speed of the light (maximum velocity of propagation of interaction) is independent on the motion of its source. In other words: there is maximum velocity of propagation of any physical object (a particle, a wave, etc.), which interact with our world.

Galileo principle and formulae for velocity transformation (1-3) do not satisfy second Einstein postulate. Therefore, Newton (or classical) mechanics based on the Galileo principles must be modified to satisfy experimental results. The most of famous experimental result contradicting to Galileo principle was Michelson-Morley experiment (1887). They tried to measure "ether drift" (the ether is imaginary substance in which electromagnetic waves are propagating; similar to the air for acoustic waves). They tried to measure difference between speed of the light in the direction of the Earth rotation and the opposite direction. According to the Galeleo law (1-3), there must be difference of $\pm v_{rot}$. The result showed no difference.

1.2 Events, 4-vectors, 4D-Intervals.

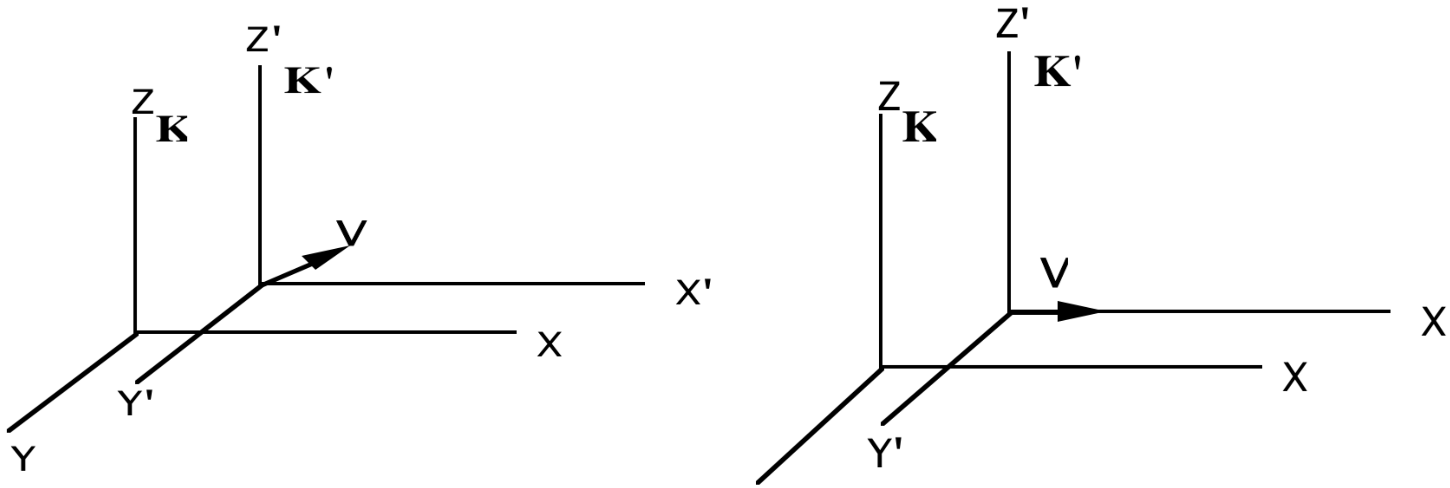

Fig. 1. Two Inertial Reference Frames: system K' moves with velocity $\vec{V}$ with respect to system K. By choice of coordinate system (rotation in 3D space) we can make $\vec{V}$ parallel to the X axis.

Let's introduce an important object in relativistic theory - an *EVENT*. An event is described by the location (in 3D coordinate system) **where** it occurred and by time **when** it occurred. As far as we know, it is full description of any event. We do not have any firm prove about the existence of other coordinates, so far...

Therefore, an event is defined by four coordinates (4-vector) in 4-dimensional time-space:

$$x^i = (x^0, x^1, x^2, x^3) \equiv (x^0, \vec{r}); \qquad \begin{array}{l} x_o = ct;\, x^1 = x;\; x^2 = y;\; x^3 = z \\ (x^4 = ix_o \;\; - Minkowski\;\; metric) \end{array}. \tag{1-4}$$

Let's look at two event A and B: A is the event when we sent a signal propagating with maximum possible speed c, B is the event when signal arrived in different point of space. Both events can be described in any reference system:

K-system: Event A: the signal was sent from location $\vec{r}_A = \hat{e}_x x_A + \hat{e}_y y_A + \hat{e}_z z_A$ at time t'_A :
$X_A^i = (x_A^0, \vec{r}_A)$;

Event B: the signal was observed in location $\vec{r}_B = \hat{e}_x x_B + \hat{e}_y y_B + \hat{e}_z z_B$ at time t_B:
$X_B^i = (x_B^0, \vec{r}_B)$.

K'-system: Event A: the signal was sent from location $\vec{r}'_A = \hat{e}_x x'_A + \hat{e}_y y'_A + \hat{e}_z z'_A$ at time t'_B :
$X_A'^i = (x_A'^0, \vec{r}'_A)$;

Event B: the signal was observed in location $\vec{r}'_B = \hat{e}_x x'_B + \hat{e}_y y'_B + \hat{e}_z z'_B$ at time t'_B:
$X_B'^i = (x_B'^0, \vec{r}'_B)$.

Signal propagates with the speed of the light in both systems. Therefore:

$$c^2(t_B - t_A)^2 - (\vec{r}_B - \vec{r}_A)^2 = c^2(t_B - t_A)^2 - (x_B - x_A)^2 - (y_B - y_A)^2 - (z_B - z_A)^2 = 0; \qquad (1\text{-}5)$$

$$c^2(t'_B - t'_A)^2 - (\vec{r}'_B - \vec{r}'_A)^2 = c^2(t'_B - t'_A)^2 - (x'_B - x'_A)^2 - (y'_B - y'_A)^2 - (z'_B - z'_A)^2 = 0. \qquad (1\text{-}5')$$

Interval

The quantity for any arbitrary events A and B, defined as:

$$s_{AB} = \sqrt{c^2(t_B - t_A)^2 - (x_B - x_A)^2 - (y_B - y_A)^2 - (z_B - z_A)^2}; \quad \text{(1-6)}$$

is of special importance in special relativity. It is called ***the interval between two events***. We have found that if interval is equal zero in one system it is equal to zero in all inertial system of references (eqs. (1-5) and (1-5')). Let's look at to events, which are infinitely close to each another: $\vec{r}_B = \vec{r}_A + d\vec{r}$; $t_B = t_A + dt$; and interval ds between them:

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2. \quad \text{(1-7)}$$

If ds^2=0, then it is equal zero in any other system ds'^2=0. In addition, ds and ds' are infinitesimals of the same order. Therefore, ds^2, ds'^2 must be proportional to each other:

$$ds^2 = a ds'^2. \quad \text{(1-8)}$$

The coefficient a can not depend on time or position not to violate homogeneity of the space and time. Similarly, it can not depend on direction of relative velocity not to contradict the isotropy of the space. Therefore, it can depend only on absolute value of relative velocity of the systems $a = a\left(\left|\vec{V}\right|\right)$.

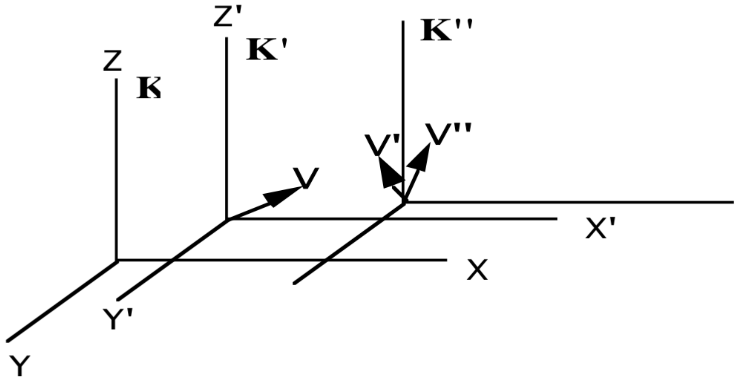

Fig. 2

Three inertial reference systems K.K'.K". K' moves with velocity $\vec{V}$ with respect to K, K" moves with velocity $\vec{V}'$ with respect to K' and with velocity $\vec{V}''$ with respect to K. $\vec{V}''$ depends on both values and direction of $\vec{V}, \vec{V}'$.

Using relation (1-8) we have for K-system:

$$ds^2 = a\left(\left|\vec{V}\right|\right)ds'^2;\;\; ds^2 = a\left(\left|\vec{V}''\right|\right)ds''^2;$$

and for K'-system:

$$ds'^2 = a\left(\left|\vec{V}'\right|\right)ds''^2;$$

yields the ratio:

$$a\left(\left|\vec{V}''\right|\right) = a\left(\left|\vec{V}\right|\right)a\left(\left|\vec{V}'\right|\right).$$

Left side depends on value of $\vec{V}''$ which depends on both values and direction of $\vec{V}, \vec{V}'$, while right side depends only on absolute values of $\vec{V}, \vec{V}'$. Therefore, we should conclude that a does not depend on velocity at all: $a = const$. The above relation reduces to $a = a^2$, i.e. $a = 1$ (we drop trivial $a = 0$). This great ratio gives us equality of infinitesimal intervals:

$$ds^2 = ds'^2; \tag{1-9}$$

and as result invariance of any finite intervals:

$$s_{AB} = \int_A^B ds = \int_A^B ds' = s'_{AB}. \tag{1-10}$$

Definitions

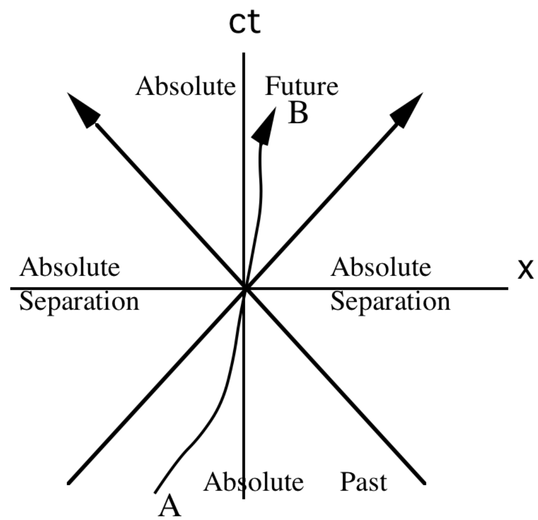

Fig. 3 World line (A-B) of the system and the light cone.

There are three distinctive values of $s^2{}_{AB}$: positive, negative and zero. The sign and the value of $s^2{}_{AB}$ does not depend on system of reference:

$$\left\{ \begin{array}{l} s^2{}_{AB} < 0,\ \textit{spacelike separation} \\ s^2{}_{AB} > 0,\ \textit{timelike separation} \\ s^2{}_{AB} = 0,\ \textit{lightlike separation} \end{array} \right\}$$

Spacelike interval: there is a system K' where two events occur at the same time, but in different points of space $s^2{}_{AB} = c^2(t_B - t_A)^2 - (\vec{r}_B - \vec{r}_A)^2 < 0; \Rightarrow s^2{}_{AB} = -(\vec{r}\,'_B - \vec{r}\,'_A)^2 < 0$;

Timelike interval: there is a system K' where two events occur at the same place, but in different points of time $s^2{}_{AB} = c^2(t_B - t_A)^2 - (\vec{r}_B - \vec{r}_A)^2 > 0; \Rightarrow s^2{}_{AB} = c^2(t'_B - t'_A)^2 < 0$;

Lightlike interval: two events can be connected by light signal $s^2{}_{AB} = 0$.

If we put event O in the origin, then $\vec{r}^{\,2} = c^2t^2$ will define the light cone. All events inside the light cone (closer to t axis) can or could be connected with event O in future or in the past. Events outside this cone are absolutely remote with respect to this event: any exchange of information between these events and the event O is impossible. Fig. 3 illustrates this puncture for 1D space with light cone equation of $x = \pm ct$.

1.3 Lorentz transformations.

Transformation related to the change of reference system **must** preserve the value of interval $s^2{}_{AB}$ between two arbitrary events: $s^2{}_{AB} = c^2(t_B - t_A)^2 - (\vec{r}_B - \vec{r}_A)^2$. An example of such transformation is rotation in 3D space which does not change time and preserves $(\vec{r}_B - \vec{r}_A)^2$. We should look for some type of rotation in 4D space which preserves the interval. There are six independent rotation in 4D space: for example in planes xy, yz, zx, xt, yt, zt. Three of them are 3 independent rotation in 3D space. The rest are special - they rotate **THE TIME**. Let's consider xt "rotation", which does not change values of y and z. To preserve interval we should use hyperbolic functions instead of trigonometric:

$$\begin{aligned} x &= x'\cosh\psi + ct'\sinh\psi;\ \ y = y'; \\ ct &= ct'\cosh\psi + x'\sinh\psi;\ z = z'; \end{aligned} \tag{1-11}$$

$$\begin{aligned} s^2 &= (ct'\cosh\psi + x'\sinh\psi)^2 - (x'\cosh\psi + ct'\sinh\psi)^2 - y'^2 - z'^2 = \\ &(ct')^2\left(\cosh^2\psi - \sinh^2\psi\right) - x'^2\left(\cosh^2\psi - \sinh^2\psi\right) - y'^2 - z'^2 = s'^2 \end{aligned}.$$

Let's relate the angle of "rotation" and the movement of K' origin $x' = 0$ (i.e. its velocity):

$$x = ct'\sinh\psi; ct = ct'\cosh\psi;\ \Rightarrow \frac{V}{c} = \frac{x}{ct} = \tanh\psi;$$

and yields final expression for Lorentz transformation:

$$\sinh\psi = \frac{V}{c}\Big/\sqrt{1 - \frac{V^2}{c^2}} = \beta\gamma;\ \cosh\psi = 1\Big/\sqrt{1 - \frac{V^2}{c^2}} = \gamma$$

with conventional dimensionless parameters $0 \le \beta < 1;\ \ 1 \le \gamma < \infty$:

$$\beta = \frac{V}{c}; \vec{\beta} = \frac{\vec{V}}{c};\ \ \gamma = 1\Big/\sqrt{1 - \frac{V^2}{c^2}} = 1\Big/\sqrt{1-\beta^2}. \tag{1-12}$$

Therefore, the Lorentz transformation in compact form is:

$$x = \gamma(x' + \beta ct'); \;\; ct = \gamma(ct' + \beta x'); \; y = y'; z = z'; \tag{1-13}$$

gives us all necessary relation to proceed further. The inverse Lorentz transformation is following from (1-13):

$$x' = \gamma(x - \beta ct); \;\; ct' = \gamma(ct - \beta x); \; y' = y; z' = z; \tag{1-14}$$

which gives us identity relations if combined with (1-13):

$$\begin{aligned} x &= \gamma(x' + \beta ct') = \gamma(\gamma(x - \beta ct) + \beta\gamma(ct - \beta x)) = \gamma^2(1-\beta^2)x = x; \\ ct &= \gamma(ct' + \beta x') = \gamma(c\gamma(ct - \beta x) + \beta\gamma(x - \beta ct)) = \gamma^2(1-\beta^2)ct = ct; \end{aligned} \tag{1-15}$$

using identity ratio:

$$\gamma^2(1-\beta^2) = \frac{1-\beta^2}{1-\beta^2} = 1. \tag{1-16}$$

More general approach to the same derivation (we leave aside y and z which do not transform). In matrix form interval is:

$$s^2 = X^T SX; \qquad S = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; \tag{1-17}$$

and arbitrary Lorentz transformation in (x,t) is:

$$X = L \cdot X'; \quad L = \begin{bmatrix} a & b \\ c & d \end{bmatrix}; \tag{1-18}$$

with condition to preserve 4-interval (we chose +):

$$L^T SL = S \Rightarrow \det L = \pm 1; \quad "+" \quad ad - bc = 1; \tag{1-19}$$

$$X' = L^{-1} \cdot X; L^{-1} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Applying standard conditions : coordinates move with ±V:

$$x' = 0; \quad x = \beta ct; \; c = \beta a; \; \beta = V/c; \; x = 0; \quad x' = -\beta ct'; \; c = -\beta d; \Rightarrow a = d;$$

we got

$$L = \begin{bmatrix} a & b \\ \beta a & a \end{bmatrix}.$$

Constant speed of light gives the symmetry of (x,ct):

$$x = ct; x' = ct'; \begin{bmatrix} ct' \\ ct' \end{bmatrix} = L \begin{bmatrix} ct \\ ct \end{bmatrix}; \Rightarrow a + b = a + \beta a; \Rightarrow b = \beta a$$

Finally, detL=1 resolves the rest of puzzle:

$$L = a \begin{bmatrix} 1 & \beta \\ \beta & 1 \end{bmatrix}; \det L = 1 \Rightarrow a = \gamma = \frac{1}{\sqrt{1 - \beta^2}} \tag{1-20}$$

2.1 Proper Time, Proper Length and Proper Volume.

Proper time is defined in moving system K', i.e. in the rest frame of an object. (a clock). Let's consider a clock located in the origin of K'. Therefore, $d\vec{r}\,'=0$ and we can write proper time for moving object:

$$ds^2 = c^2dt^2 - d\vec{r}^{\,2} = c^2dt'^2$$

$$dt' = dt\sqrt{1-\frac{d\vec{r}^{\,2}}{c^2dt^2}} = dt\sqrt{1-\frac{v^2}{c^2}} = dt\sqrt{1-\beta^2}\,; \qquad (2\text{-}1)$$

$$t'_B - t'_A = \int_A^B dt\sqrt{1-\frac{v^2}{c^2}} = \int_A^B \frac{dt}{\gamma}.$$

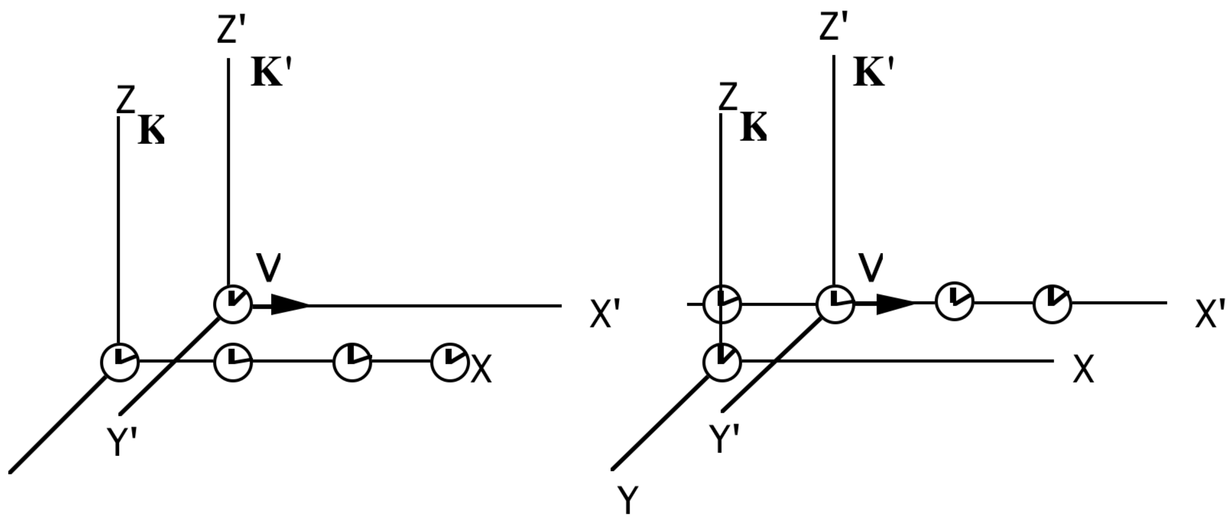

Fig. 4 To find the proper time at origin of K', we compare one clock in K' with set of clocks in K (left); to find proper time at origin of K', we compare one clock in K with set of clocks in K (right). This process is asymmetric and a clock compared with a set of clocks always lags behind.

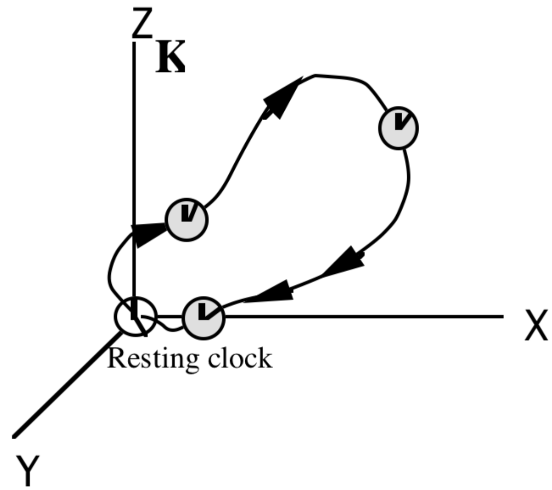

Fig. 5 The only correct way to compare clocks: use two clocks, start them at the same point of space, leave one at the rest and bring second at the same point to compare elapsed time. The clock at rest will show more time then moving clock. Why?

It is impossible to return clock using rectilinear motion; i.e. moving clock must be accelerated. Therefore, the system related to traveling clock is not inertial and is not identical to inertial system where first clock rests. Thus, a moving clock will show less time elapsed then a resting one. On other hand, we can look for the motion of K system from point of view of K'. Now we should locate a clock at the origin of K, and $d\vec{r} = 0$. Similar to eq. (2-1) we have:

$$dt = dt'\sqrt{1 - \frac{d\vec{r}\,'^2}{c^2 dt'^2}} = dt\sqrt{1 - \frac{v^2}{c^2}}; t_B - t_A = \int_A^B dt'\sqrt{1 - \frac{v^2}{c^2}} = \int_A^B \frac{dt'}{\gamma}. \tag{2-2}$$

It looks as a contradiction: time in K' system is both faster and slower then in K system. What is not correct is to compare different clocks in the resting system with fixed one in the moving system. The solution of “paradox” is illustrated by Figures 4 and 5.

"Time paradox" is directly related to the *Lorentz contraction*. Suppose that there is a rod at rest in K system measured $l = x_B - x_A$ where x_B, x_A are coordinates of two end of the rod. We should determine length of the same rod in K' system: $x_A = \gamma(x'_A + \beta c t'); x_B = \gamma(x'_B + \beta c t')$; at the same moment of time t':

$$l' = x'_B - x'_A = \left(x_B - x_A\right) / \gamma = l / \gamma . \tag{2-3}$$

Therefore, observed from a moving system the resting rod contracts by factor γ. The same will be correct if we look from K system on the rot resting in K' system at the same moment of time t using $x'_A = \gamma(x_A - \beta ct); x'_B = \gamma(x_B - \beta ct)$;

$$l = l' / \gamma . \tag{2-4}$$

Again, there is no contradiction. We are looking for the length of the rod by observing its ends at the same moment of time, but in different systems. The source of "asymmetry": time and space coordinates depend of the system of observation.

As we derived, coordinates transverse to the relative velocity of the system do not change $y' = y; z' = z$. Therefore, the volume of the body will decrease proportionally to the contraction of coordinate parallel to the relative velocity of the system (x). This volume is called proper volume:

$$V = V_0 / \gamma \tag{2-5}$$

To finish discussion, let's consider a synchronization procedure of the clocks. The natural way to set clocks located at different positions x in K system is to send periodical light signal from the origin and set them at time $t = x / c$ when light reach them. The traveling clock, fixed at origin of K', sees the distances in K system contracting by factor γ, and therefore the clock "thinks" that elapse time is $t' = x / \gamma c$.

What is most important that 4-dimentional volume

$$d\Omega = cdtdV \equiv dx^0 dx^1 dx^2 dx^3$$

is invariant of Lorentz transformations (we will discuss it at next lecture). It is direct consequence of the unit determinant of Lorentz transformation matrix:

$$d\Omega = \det[L] d\Omega';$$

$$L = \begin{bmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \ \det[L] = \gamma^2(1-\beta^2) = 1$$

2.2 Transformation of velocities.

Lorentz transformation of coordinates and time give us all necessary information to calculate velocity of the particles is arbitrary inertial system:

$$\vec{v}=\frac{d\vec{r}}{dt}=\hat{e}_x\frac{dx}{dt}+\hat{e}_y\frac{dy}{dt}+\hat{e}_z\frac{dz}{dt};\vec{v}'=\frac{d\vec{r}'}{dt'}=\hat{e}_x\frac{dx'}{dt'}+\hat{e}_y\frac{dy'}{dt'}+\hat{e}_z\frac{dz'}{dt'};.$$

Let's rewrite (1-13) in form of differentials:

$$cdt=\gamma(cdt'+\beta dx');dx=\gamma(dx'+\beta cdt');dy=dy';dz=dz'; \tag{2-6}$$

and divide coordinate differential by time differential:

$$v_x=\frac{dx}{dt}=\frac{c\gamma(dx'+\beta cdt')}{\gamma(cdt'+\beta dx')}=\frac{\frac{dx'}{dt'}+\beta c}{1+\beta/c\cdot\frac{dx'}{dt'}}=\frac{v'_x+V}{1+v'_xV/c^2};$$

$$v_y=\frac{dy}{dt}=\frac{cdy'}{\gamma(cdt'+\beta dx')}=\frac{\frac{dy'}{dt'}}{\gamma(1+\frac{\beta dx'}{cdt'})}=\frac{v'_y\sqrt{1-\frac{V^2}{c^2}}}{1+\frac{v'_xV}{c^2}}; \tag{2-7}$$

$$v_z=\frac{dz}{dt}=\frac{cdz'}{\gamma(cdt'+\beta dx')}=\frac{dz'/dt'}{\gamma(1+\beta/c\cdot dx'/dt')}=\frac{v'_z\sqrt{1-\frac{V^2}{c^2}}}{1+v'_xV/c^2}.$$

The transformation of velocities is more complex then transformation of space-time coordinates. It should not be of any surprise; e.g. the 3-D velocity is not a 4D object and it combines time and coordinates in "unnatural way for 4D world".

Geometry in Special relativity

Appendix A: 4-D metric of special relativity

"Tensors are mathematical objects - you'll appreciate their beauty by using them"

4-scalars, 4 vectors, 4- tensors. (closely follows [CTF])

An event is fully described by coordinates in 4D-space (time and 3D-space), i.e., by a 4 vector:

$$X^i = (x^0, x^1, x^2, x^3) \equiv (x^0, \vec{r}) \; ; \; x^0 = ct; \; x^1 = x; x^2 = y; x^3 = z \; . \qquad \text{(A-1)}$$

Consider a non-degenerated transformation in 4D space

$$X' = X'(X) \; ; \qquad \text{(A-2)}$$

$$x'^i = x'^i(x^0, x^1, x^2, x^3); \; i = 0,1,2,3 \; ; \qquad \text{(A-3)}$$

and allowing the inverse transformation

$$X = X(X') \qquad \text{(A-4)}$$
$$x^i = x^i(x'^0, x'^1, x'^2, x'^3); \; i = 0,1,2,3$$

Jacobian matrices describe the local deformations of the 4D space:

$$\frac{\partial x'^i}{\partial x^j}; \; \frac{\partial x^j}{\partial x'^i}; \qquad \text{(A-5)}$$

and are orthogonal to each other

$$\sum_{j=0}^{j=3} \frac{\partial x'^i}{\partial x^j} \cdot \frac{\partial x^j}{\partial x'^k} = \frac{\partial x'^i}{\partial x^j} \cdot \frac{\partial x^j}{\partial x'^k} = \frac{\partial x'^i}{\partial x'^k} = \delta^i_k \; ; \qquad \text{(A-6)}$$

Here, we start with the convention to "silently" summate the repeated indexes:

$$a^i b_i \equiv \sum_{i=0}^{i=3} a^i b_i \; . \qquad \text{(A-7)}$$

A 4-scalar is defined as any scalar function that preserves its value while undergoing Lorentz transformation (including rotations in 3D space):

$$f(X') = f(X); \forall X' = L \otimes X \qquad \text{(A-8)}$$

Contravariant 4-vector $A^i = (A^0, A^1, A^2, A^3)$ is defined as an object for which the transformation rule is the same as for the 4D-space vector:

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j \tag{A-9}$$

i.e.,

$$A'^i = \frac{\partial x'^i}{\partial x^j} A^j ; \tag{A-10}$$

or explicitly

$$A'^i = \frac{\partial x'^i}{\partial x^0} A^0 + \frac{\partial x'^i}{\partial x^1} A^1 + \frac{\partial x'^i}{\partial x^2} A^2 + \frac{\partial x'^i}{\partial x^3} A^3 ; \tag{A-11}$$

Covariant 4-vector $A_i = (A_0, A_1, A_2, A_3)$ is defined as an object for which the transformation rule is

$$A'_i = \frac{\partial x^j}{\partial x'^i} A_j ; \tag{A-12}$$

i.e., the inverse transformation is used for covariant components.

Contravariant F^{jl} **and Covariant** G_{jl} **4-tensors of rank 2** are similarly defined :

$$F'^{ik} = \frac{\partial x'^i}{\partial x^j} \frac{\partial x'^k}{\partial x^l} F^{jl} ; G'_{ik} = \frac{\partial x^j}{\partial x'^i} \frac{\partial x^l}{\partial x'^k} G_{jl} ; \tag{A-13}$$

Mixed tensors with co- and contra- variant indexes are transformed by mixed rules:

$$F'^i{}_k = \frac{\partial x'^i}{\partial x^j} \frac{\partial x^l}{\partial x'^k} F^j{}_l ; G'^k_i = \frac{\partial x^j}{\partial x'^i} \frac{\partial x'^k}{\partial x^l} G_j{}^l . \tag{A-14}$$

Tensors of higher rank also are defined in this way. Thus, a tensor of rank n has 4^n components: 4-scalar - n=0, 4^0=1 component; 4-vector - n=1, 4^1=4 components; a tensor of rank 2 - n=2, 4^2=16 components; and so on. Some components may be dependent ones. For example, symmetric- and asymmetric-tensors of rank 2 are defined as $S^{ik} = S^{ki}$; $A^{ik} = -A^{ki}$. A symmetric tensor has 10 independent components: four diagonal terms S^{ii} , and six $S^{i,k\neq i} = S^{k\neq i,i}$ non-diagonal terms. An asymmetric tensor has six independent components: $A^{i,k\neq i} = -A^{k\neq i,i}$, while all diagonal terms are zero $A^{ii} = -A^{ii} \equiv 0$. Any tensor of second rank can be expanded in symmetric- and asymmetric-parts:

$$F^{ik} = \frac{1}{2}\left(F^{ik} + F^{ki}\right) + \frac{1}{2}\left(F^{ik} - F^{ki}\right). \qquad \text{(A-15)}$$

The scalar product of two vectors is defined as the product of the co- and contra-variant vectors:

$$A \cdot B = A_i B^i \,; \qquad \text{(A-16)}$$

It is the invariant of transformations:

$$A'_i B'^i = \frac{\partial x^j}{\partial x'^i}\frac{\partial x'^i}{\partial x^k} A_j B^k = \frac{\partial x^j}{\partial x^k} A_j B^k = \delta^j_k A_j B^k = A_k B^k \,; \qquad \text{(A-17)}$$

where

$$\delta^j_k = \begin{Bmatrix} 1; j = k \\ 0; j \neq k \end{Bmatrix} \qquad \text{(A-18)}$$

is the unit tensor. Note that the trace of any tensor is a trivial 4-scalar .

$$Trace(F) = F^i{}_i \equiv F^0{}_0 + F^1{}_1 + F^2{}_2 + F^3{}_3 = F'^i{}_i \,; \qquad \text{(A-19)}$$

Metrics

The metrics (or norm that must be a 4-scalar) defines the geometry of the 4-space. The traditional (geometric) way is to define it as $ds^2 = dx^i dx_i$. The 4-scalar is defining interval between events, details on which can be found in any text on relativity (see additional material to the course or in you favorite book, for example, *L.D. Landau, E.M. Lifshitz, "The Classical Theory of Fields"*)

An infinitesimal interval defines the norm of our "flat" space-time in special relativity:

$$ds^2 = dx^{0\,2} - dx^{1\,2} - dx^{2\,2} - dx^{3\,2} = dx_0{}^2 - dx_1{}^2 - dx_2{}^2 - dx_3{}^2 ; \qquad \text{(A-20)}$$

and the diagonal metric tensor g^{ik} :

$$ds^2 = g_{ik} dx^i dx^k = g^{ik} dx_i dx_k ;$$

$$g_{ik} = g^{ik};\ g^{00} = 1; g^{11} = -1; g^{22} = -1; g^{33} = -1; \qquad \text{(A-21)}$$

in which all non-diagonal term are zero $;g^{i\neq k} = 0$. The metric (A-21) is a consequence of the Euclidean space- frame. In general, it suffices that g^{ik} must be symmetric $g^{ik} = g^{ki}$. Note that the contraction of the metric tensor yield the unit tensor $g_{ij} g^{jk} = \delta_i^{\ k}$. Comparing (A-21) and (A-20) we conclude that

$$x^i = g^{ik} x_k ;\ x_i = g_{ik} x^k ; \qquad \text{(A-22)}$$

i.e., the metric tensor g^{ik} raises indexes and g_{ik} lowers them, transforming the co- and contra-variant components

$$F^{..k......}_{......i..} = g^{kj} F^{.........}_{....j.i..} = g^{kj} g_{il} F^{.......l..}_{...j....} ; etc. \qquad \text{(A-23)}$$

For 4-vectors, the lowering or rising indexes change the sign of spatial components. There is no distinction between co- and contra- variants; they can be switched without any consequences. Convention defines them as follows :

$$A^i = (A^0, \vec{A}) = (A^0, A^1, A^2, A^3)$$
$$A_i = (A_0, -\vec{A}) = (A_0, -A_1, -A_2, -A_3) \quad ; \tag{A-24}$$

$$A \cdot B = A^i \cdot B_i = A^0 B^0 - \vec{A} \cdot \vec{B}$$

The $g^{kj}, g_{il}, g_i^{\,k} \equiv \delta_i^k$ tensors are special as they are identical in all inertial frames (coordinate systems). This is apparent for δ_i^k :

$$\delta'^i_j = \frac{\partial x^k}{\partial x'^j} \frac{\partial x'^i}{\partial x^l} \delta_l^{\,k} = \frac{\partial x^k}{\partial x'^j} \frac{\partial x'^i}{\partial x^k} = \frac{\partial x'^i}{\partial x'^j} = \delta_j^{\,i}; \tag{A-25}$$

while g^{ik} invariance is obvious from the invariance of the interval (A-20). Hence, it is better to say that the preservation of g^{ik} determines an allowable group of transformations in the 4D-space - the Lorentz group (see Appendix B). There is one more special tensor: the totally asymmetric 4-tensor of rank 4: e^{iklm} . Its components change sign when any if indexes are interchanged:

$$e^{iklm} = -e^{kilm} = -e^{ilkm} = -e^{ikml} . \tag{A-26}$$

meaning that the components with repeated indexes are zero: $e^{..i..k..} = 0;\ i = k;$ and only non-zero components are permutations of $\{0,1,2,3\}$.

By convention

$$e^{0123} = 1; \tag{A-27}$$

So that $e^{1023} = -1$. The tensor e^{iklm} also is invariant of Lorentz transformation that is directly related to the determinant of the Jacobian matrix of Lorentz transformations $J = \det\left[\frac{\partial x'}{\partial x}\right]$.

$$e'^{iklm} = \frac{\partial x'^i}{\partial x^j} \frac{\partial x'^k}{\partial x^n} \frac{\partial x'^l}{\partial x^p} \frac{\partial x'^m}{\partial x^q} e^{jnpq} = \det\left[\frac{\partial x'}{\partial x}\right] e^{jnpq} \delta_j^i \delta_n^k \delta_p^l \delta_q^m = e^{iklm}; \tag{A-28}$$

For Lorentz transformations $J=1$. In the best courses on linear algebra, the above equation is used as the definition of the matrix determinant. For details, see Section 3.4 (pp. 132-134) and section 4.1 in G. Arfken's "Mathematical Methods for Physicists" (where Eq. 4.2 is equivalent to $a^i_j a^k_n a^l_p a^m_q e^{jnpq} = \det[a]e^{jnpq}\delta^i_j\delta^k_n\delta^l_p\delta^m_q$). As mentioned in Landau CSF (footnote in §6), the invariance of a totally asymmetric tensor of rank equal to the dimension of the space with respect to rotations is the general rule. This is very easy to prove for 2D space. The 2D totally asymmetric tensor of rank 2 is $e^{ik} = \begin{Bmatrix} 0 & 1 \\ -1 & 0 \end{Bmatrix}$ has transformations of

$$e'^{ik} = \frac{\partial x'^i}{\partial x^j}\frac{\partial x'^k}{\partial x^n}e^{jn} = \frac{\partial x'^i}{\partial x^1}\frac{\partial x'^k}{\partial x^2}e^{12} + \frac{\partial x'^i}{\partial x^2}\frac{\partial x'^k}{\partial x^1}e^{21} = \frac{\partial x'^i}{\partial x^1}\frac{\partial x'^k}{\partial x^2} - \frac{\partial x'^i}{\partial x^2}\frac{\partial x'^k}{\partial x^1} = \det\begin{Bmatrix} \dfrac{\partial x'^i}{\partial x^1} & \dfrac{\partial x'^i}{\partial x^2} \\ \dfrac{\partial x'^k}{\partial x^1} & \dfrac{\partial x'^k}{\partial x^2} \end{Bmatrix};$$

(A-29)

Therefore:

$$e'^{ii} = \det\begin{Bmatrix} \dfrac{\partial x'^i}{\partial x^1} & \dfrac{\partial x'^i}{\partial x^2} \\ \dfrac{\partial x'^i}{\partial x^1} & \dfrac{\partial x'^i}{\partial x^2} \end{Bmatrix} = 0 = e^{ii}; e'^{12} = \det\begin{Bmatrix} \dfrac{\partial x'^1}{\partial x^1} & \dfrac{\partial x'^1}{\partial x^2} \\ \dfrac{\partial x'^2}{\partial x^1} & \dfrac{\partial x'^2}{\partial x^2} \end{Bmatrix} = 1 = e^{12}; e'^{21} = \det\begin{Bmatrix} \dfrac{\partial x'^2}{\partial x^1} & \dfrac{\partial x'^2}{\partial x^2} \\ \dfrac{\partial x'^1}{\partial x^1} & \dfrac{\partial x'^1}{\partial x^2} \end{Bmatrix} = -1 = e^{21};$$

(A-30)

for rotations when $\det\begin{Bmatrix} \dfrac{\partial x'}{\partial x} \end{Bmatrix} = 1$. Finally, convolution of absolutely asymmetric tensor of rank n is equal n! - a number of permutations. In particular, $e^{iklm}e_{iklm} = 4! = 24$.

Tensors of any rank can be real tensors or pseudo-tensors, i.e., scalars and pseudo-scalars, vectors and pseudo-vectors, and so forth. They follow the same rules for rotations, but have different properties with respect to the sign inversions of coordinates: special transformations that cannot be reduced to rotations. An example of these transformations is the inversion of 3D coordinates signs.

The totally asymmetric tensor e^{iklm} is pseudo-tensor - it does not change sign when the space or time coordinates are inverted: $e^{0123}=1;$ (it is the same as for 3D version of it, $e^{\alpha\beta\gamma}$; $\vec{C}=\vec{A}\times\vec{B}\Rightarrow C^{\alpha}=e^{\alpha\beta\gamma}A^{\beta}B^{\gamma}$, $e^{123}=1;$). Recall that the vector product in 3D space is a pseudo-vector. Under reflection $\vec{A}\rightarrow-\vec{A};\ \vec{B}\rightarrow-\vec{B};\quad \vec{C}\Rightarrow\vec{C}!$

We can represent six components of an asymmetric tensor by two 3D-vectors;

$$(A^{ik})=(\vec{p},\vec{a})=\begin{bmatrix} 0 & p_x & p_y & p_z \\ -p_x & 0 & -a_z & a_y \\ -p_y & a_z & 0 & -a_x \\ -p_z & -a_y & a_x & 0 \end{bmatrix};(A_{ik})=(-\vec{p},\vec{a}). \tag{A-31}$$

The time-space components of this tensor change sign under the reflection of coordinates, while purely spatial components do not. Hence, $\vec{p}$ is a real (polar) 3 -D vector , and $\vec{a}$ is 3D pseudo-vector (axial) vector.

$$A^{*ik}=e^{iklm}A_{lm} \tag{A-32}$$

is called the dual tensor to asymmetric tensor A^{ik}, and vice versa. The convolution of dual tensors is pseudo-scalar $ps=A^{*ik}A_{ik}$. Similarly, $e^{iklm}A_m$ is a tensor of rank 3 dual to 4 vector A^i.

Differential operators

Next consider differential operators

$$\frac{\partial}{\partial x'^i} = \frac{\partial x^k}{\partial x'^i}\frac{\partial}{\partial x^k}; \tag{A-32}$$

that follow the transformation rule for covariant vectors. Therefore, the differentiation with respect to a contravariant component is a covariant vector operator and vice versa! Accordingly, we can now express standard differential operators:

4-gradient: $$\partial^i \equiv \frac{\partial}{\partial x_i} = \left(\frac{\partial}{\partial x_0}, -\vec{\nabla}\right); \partial_i = \frac{\partial}{\partial x^i} = \left(\frac{\partial}{\partial x_0}, \vec{\nabla}\right);$$

(A-33)

4-divergence $$\partial^i A_i = \partial_i A^i = \frac{\partial A^0}{\partial x^0} + \vec{\nabla}\vec{A};$$

(A-34)

4-Laplacian (De-Lambert-dian): $$\Box = \partial^i \partial_i = \frac{\partial^2}{\partial x^{0^2}} - \vec{\nabla}^2. \tag{A-35}$$

Using differential operators allows us to construct 4-vectors and 4-tensors from 4-scalars. For example:

$$x^i = \partial^i (s^2). \tag{A-36}$$

Other example is the phase of an oscillator: $\exp\left[i\left(\omega t-\vec{k}\vec{r}\right)\right]; \varphi=\omega t-\vec{k}\vec{r};\ \omega=kc$. The phase is 4-scalar; it does not depend on the system of observation. It is very important, but not an obvious fact! Imagine a sine wave propagating in space and a detector that registers when the wave intensity is zero. Zero value of wave amplitude is the event and does not depend on the system of observation. Similarly, we can detect any chosen phase. Therefore, the phase is 4-scalar and

$$k^{i}=\partial^{i}\varphi=(\omega/c,\vec{k}) \tag{A-37}$$

is a 4-wave-vector undergoing standard transformation. Thus, we readily assessed the transformation of frequency and wave-vector from one system to the other, called the Doppler shift:

$$\omega=\gamma(\omega'+c\vec{\beta}\vec{k}');\ \vec{k}_{//}=\gamma(\vec{k}'_{//}+\vec{\beta}\omega'/c);\vec{k}_{\perp}=\vec{k}'_{\perp}. \tag{A-38}$$

then simply applying Lorentz transformations we found as last time:

$$\frac{\partial x'^{i}}{\partial x^{j}}=\begin{bmatrix}\gamma & -\beta\gamma & 0 & 0\\ -\beta\gamma & \gamma & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix};\frac{\partial x^{i}}{\partial x'^{j}}=\begin{bmatrix}\gamma & \beta\gamma & 0 & 0\\ \beta\gamma & \gamma & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{A-39}$$

4-velocity, 4-acceleration

Another way to create new 4-vectors is to differentiate a vector as a function of the scalar function, for example, the interval. Unsurprisingly, 3D velocity transformation rules do not satisfy simple 4-D vector transformation rules; to differentiate over time that is not 4-scalar will be meaningless. 4-velocity is defined as derivative of the coordinate 4-vector x^i over the interval s :

$$u^i = \frac{dx^i}{ds} ; \qquad \text{(A-40)}$$

and ,with simple way to connect it to 3D velocity $dx^i = (c,\vec{v})dt; ds = cdt\sqrt{1-\frac{v^2}{c^2}} = cdt/\gamma$ we obtain :

$$u^i = \gamma(1,\vec{v}/c); \qquad \text{(A-41)}$$

that follows all rules of transformation. The first interesting result is that 4-velocity is dimension-less and has unit 4-length:

$$u^i u_i = 1 \qquad \text{(A-42)}$$

which is evident by taking into account that $ds^2 = dx^i dx_i \equiv u^i u_i ds^2$. Thus, it follows directly that 4-velocity and 4-acceleration

$$w^i = \frac{du^i}{ds} \qquad \text{(A-43)}$$

are orthogonal to each other:

$$u^i w_i = \frac{d(u^i u_i)}{2ds} = 0 . \qquad \text{(A-44)}$$

For next class:

What is more amazing is that simply multiplying 4-velocity by the constant mc yields the 4-momentum:

$$mcu^i = (\gamma mc, \gamma m\vec{v}) = (E/c, \vec{p}) \qquad \text{(A-45)}$$

, furthermore, gives the simple rules to calculate energy and momentum of particles in arbitrary frame (beware of definition of **g** here!):

$$E = \gamma(E' + c\vec{\beta}\vec{p}); \vec{p}_{//} = \gamma(\vec{p}'_{//} + \vec{\beta}E'/c); \vec{p}_\perp = \vec{p}'_\perp . \qquad \text{(A-46)}$$

Integrals and their relations

Transformation rules are needed for elements of hyper-surfaces and for the generalization of Gauss and Stokes theorems. Those who studied have external differential forms in advances math courses will find it trivial, but for those who have not they may not be easy to follow. We will use all necessary relations during the course when we need them. Here is a simple list:

1. The integral along the 4-D trajectory has an element of integration dx^i i.e,. similar to $d\vec{r}$ for the 3D case.

2. An element of the 2D surface in 4D space is defined by two 4-vectors dx_k, dx'_k and an element of the surface is the 2-tensor $df_{ik} = dx_i dx'_k - dx_k dx'_i$. A dual tensor $df^{*ik} = \frac{1}{2} e^{iklm} df_{lm}$; is normal to the surface tensor: $df_{ik} df^{*ik} = 0$. It is similar to 3D case when the surface vector $df_\alpha = \frac{1}{2} e_{\alpha\beta\gamma} f_{\alpha\beta};\ \alpha,\beta = 1,2,3$ is perpendicular to the surface.

3. An element of the 3D surface (hyper-surface or 3D manifold) in 4D space is defined by three 4-vectors dx_k, dx'_k, dx''_k and the three tensor element and dual vector of the 3D surface are

$$dS^{ikl} = \det \begin{bmatrix} dx^i & dx'^i & dx''^i \\ dx^k & dx'^k & dx''^k \\ dx^l & dx'^l & dx''^l \end{bmatrix} = e^{nklm} dS_n ; dS^i = \frac{-1}{6} e^{iklm} dS_{klm} . \qquad \text{(A-47)}$$

Its time component is equal to the elementary 3D-volume $dS^o = dxdydz$ **.**

4. The easiest case is that of a 4D-space volume created by four 4-vectors: $dx_i^{(1)}; dx_j^{(2)}; dx_k^{(3)}; dx_l^{(4)}$ which is a scalar

$$d\Omega = e^{iklm} dx_i^1 dx_j^2 dx_k^3 dx_l^4 \Rightarrow d\Omega = dx_0 dx_1 dx_2 dx_3 = cdtdV ;$$

5. The rules for generalization of the Gauss and Stokes theorems (actually one general Stokes theorem, expressed in differential forms) are similar to those for 3D theorems, but there more of them:

$$\oint A^i dS_i = \int \frac{\partial A^i}{\partial x^i} d\Omega;\ \oint A^i dx_i = \int \frac{\partial A^i}{\partial x^k} df_{ik} ; \int A^{ik} df^*{}_{ik} = \int \frac{\partial A^{ik}}{\partial x^k} dS_i . \qquad \text{(A-48)}$$

What we rehashed

- Least action principle
- Lagrangian and Hamiltonian formalisms
- Special relativity, Lorentz transforamtion
- Geometry of 4D space-time, co- and contra- variant vectors and tensors
- Differential operators and integrals in 4D space-time
- Interval, 4-coordinate, 4-velocity and 4-acceleration

- We do not presume that you can remember everything, but we will use most of the notions we introduced today
- It is a lot of material – please look through your favorite book to remind yourself about this fascinating relativistic world