Lecture 1: Particles In Electromagnetic Fields.

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.

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:

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

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

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:**

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

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

---

* 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.

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

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

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

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

$$H = \sum_{i=1}^{n} P^i \dot q_i - L \equiv \sum_{i=1}^{n} \frac{\partial L}{\partial \dot q_i}\dot q_i - L;\; L = \sum_{i=1}^{n} P^i \dot q_i - H. \qquad (5)$$

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

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

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):

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

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:

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

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:**

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

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. 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:

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

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:

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

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

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

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)

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

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

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

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:

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

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

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

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

(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):

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

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

1.1 Relativistic Mechanics

From here further: i=0,1,2,3 in Minkowki space with (1,-1,-1,-1) metric.

see Appendix A for 4-D metric, vectors and tensors

Let's use Principle of Least Action for a relativistic particle. To determine *the action integral for a free particle* (which does not interact with the rest of the world), we must ensure that the action integral does not depend on our choice of the inertial system. Otherwise, the laws of the particle motion also will depend on the choice of the reference system, which contradicts the first principle of relativity. Therefore, the action must be invariant of Lorentz transformations and rotation in 3D space; i.e., it must depend on a 4D scalar. So far, *from Appendix A*, we know of one 4D scalar for a free particle: the interval. We can employ it as trial function for the action integral, and, by comparing the result with classical mechanics find a constant α connecting the action with the integral of the interval: $ds^2 = dx^i dx_i \equiv \sum_{i=1}^{4} dx^i dx_i = (cdt)^2 - (d\vec{r})^2$

$$S = -\alpha \int_A^B ds = -\alpha \int_A^B \sqrt{(cdt)^2 - d\vec{r}^{\,2}}\,. \tag{16}$$

The minus sign before the integral reflects a natural phenomenon: the law of inertia requires a resting free particle to stay at rest in inertial system. The interval $ds = cdt$ has a maximum possible value ($cdt \geq \sqrt{(cdt)^2 - d\vec{r}^{\,2}}$) and requires for the action to be minimal, that the sign is set to be "-".

The integral (16) is taken along the world line of the particle. The initial point A (event) determines the particle's start time and position, while the final point B (event) determines its final time and position. The action integral (16) can be represented as integral with respect to the time:

$$S = -\alpha \int_A^B \sqrt{(cdt)^2 - d\vec{r}^{\,2}} = -\alpha c \int_A^B dt \sqrt{1 - \vec{v}^{\,2} / c^2} = \int_A^B \mathbf{L} dt\,; \qquad \mathbf{L} = -\alpha c \sqrt{1 - \frac{\vec{v}^2}{c^2}}\,;\ \vec{v} = \frac{d\vec{r}}{dt}\,;$$

where $\mathbf{L}$ signifies the Lagrangian function of the mechanical system. It is important to note that while the action is an invariant of the Lorentz transformation, the Lagrangian is not. It must depend on the reference system because time depends on it. To find coefficient α, we compare the relativistic form with the known classical form by expanding $\mathbf{L}$ by $\vec{v}^{\,2} / c^2$:

$$\mathbf{L} = -\alpha c \sqrt{1 - \frac{\vec{v}^2}{c^2}} \approx -\alpha c + \alpha \frac{\vec{v}^2}{2c}; \qquad \mathbf{L}_{classical} = m \frac{\vec{v}^2}{2}\,;$$

which confirms that α is positive and $\alpha = mc$, where m is the mass of the particle. Thus, we found the action and the Lagrangian for a relativistic particle:

$$S = -mc \int_A^B ds\,; \tag{17}$$

$$\mathbf{L} = -mc^2 \sqrt{1 - \frac{\vec{v}^2}{c^2}}\,; \tag{18}$$

The energy and momentum of the particles are defined by the standard relations eqs. (4) and (5):

$$\vec{p} = \frac{\partial \mathbf{L}}{\partial \vec{v}} = \frac{m\vec{v}}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} = \gamma m\vec{v}; \tag{19}$$

$$E = \vec{p}\vec{v} - L = \gamma mc^2;\ \gamma = 1/\sqrt{1 - \vec{v}^2/c^2} \tag{20}$$

with ratio between them of

$$E^2 = \vec{p}^2 c^2 + \left(mc^2\right)^2. \tag{21}$$

The energy of the resting particle does not go to zero as in classical mechanics but is equal to the famous Einstein value, $E = mc^2$; with the standard classical additions at low velocities ($v << c; p << mc$): $E \cong mc^2 + m\frac{\vec{v}^2}{2} \cong mc^2 + \frac{\vec{p}^2}{2m}$.

Four-momentum, conservation laws. The least-action principle gives us the equations of motion and an expression for the momentum of a system. Let us consider the total variation of an action for a single particle:

$$\delta S = -mc\delta\int_A^B ds = -mc\delta\int_A^B \sqrt{dx^i dx_i} = -mc\{\int_A^B \sqrt{d\delta x^i dx_i} + \sqrt{dx^i d\delta x_i}\} =$$

$$-mc\{\int_A^B \frac{d\delta x^i dx_i}{2\sqrt{dx^i dx_i}} + \frac{dx^i d\delta x_i}{2\sqrt{dx^i dx_i}}\} = -mc\int_A^B \frac{dx^i d\delta x_i}{ds} = -mc\int_A^B u^i d\delta x_i;$$

where $u^i \equiv dx^i/ds$ is 4-velocity. Integrating by parts,

$$\delta S = -mcu^i \delta x_i\Big|_A^B + mc\int_A^B \delta x_i \frac{du^i}{ds} ds; \tag{22}$$

we obtain the expression that can be used for all purposes. First, using the least-action principle with fixed A and B $\delta x_i(A) = \delta x_i(B) = 0$, to derive the conservation of 4-velocity for a free particle: $\frac{du^i}{ds} = 0;\ u^i = const$ or **<u>the inertia law.</u>**

Along a real trajectory $mc\int_A^B \delta x_i \frac{du^i}{ds} ds = 0$ the action is a function of the limits A and B (see eq. (12): $\delta S_{real\ traj} = (-E\delta t + \vec{P}\delta\vec{r})\Big|_A^B$, i.e., $dS_{real\ traj} = -Edt + \vec{P}d\vec{r}$ is the full differential of t and $\vec{r}$ with energy and momentum as the parameters. We note that this form of the action already is a Lorentz invariant:

$$\delta S_{real\ traj} = (-E\delta t + \vec{P}\delta\vec{r})\Big|_A^B = (-P^i\delta x_i)\Big|_A^B;$$

<u>i.e. classical Hamiltonian mechanics always encompassed a relativistic form and a metric: a scalar δS is a 4-product of P^i and δx_i with the metric (1,-1,-1,-1)</u>. Probably one of most remarkable things in physics is that its classic approach detected the metric of 4-D space and time at least a century before Einstein and Poincaré.

To get 4-momentum, we consider a real trajectory $mc\int_A^B \delta x_i \frac{du^i}{ds} ds = 0$ and set $\delta x_i(B) = \delta x_i$:

$$p^i = -\frac{\partial S}{\partial x_i} = -\partial^i S = mcu^i = (\gamma mc, \gamma m\vec{v}) = (E/c, \vec{p}) \tag{23}$$

with an obvious scalar product ($u^i u_i = 1$, see Appendix A. eq. (A.42))

$$p^i p_i = E^2 / c^2 - \vec{p}^2 = m^2 c^2 u^i u_i = m^2 c^2 . \tag{24}$$

Equivalent forms of presentation are

$$p^i = (E/c, \vec{p}) \equiv m\gamma_v (c, \vec{v}) \equiv \frac{(mc, m\vec{v})}{\sqrt{1 - v^2 / c^2}} \tag{25}$$

and, Lorentz transformation (P^i is a 4-vector, K' moves with $\vec{V} = \hat{e}_x V$):

$$E = \gamma_V (E' + c\beta_V p'_x); p_x = \gamma_V (p'_x + \beta_V E'/c); p_{y,z} = p'_{y,z}; \gamma_V = 1/\sqrt{1-\beta_V{}^2}; \beta_V = V/c; \tag{26}$$

where subscripts are used for γ, β to define the velocity to which they are related. .

Equation (24) expresses energy, velocity, and the like in terms of momenta and allows us to calculate all differentials:

$$E = c\sqrt{\vec{p}^2 + m^2c^2}; dE = cd\sqrt{\vec{p}^2 + m^2c^2} = \frac{d\vec{p}\cdot c\vec{p}}{\sqrt{\vec{p}^2 + m^2c^2}} = \frac{c^2\vec{p}\cdot d\vec{p}}{E} = \vec{v}\cdot d\vec{p}; \tag{27}$$

$$\begin{gathered} \vec{v} = \frac{c\vec{p}}{\sqrt{\vec{p}^2 + m^2c^2}}; \vec{a}dt = d\vec{v} = d\frac{c\vec{p}}{\sqrt{\vec{p}^2 + m^2c^2}} = \\ \frac{c\left(d\vec{p}(\vec{p}^2 + m^2c^2) - \vec{p}(\vec{p}d\vec{p})\right)}{\left(\sqrt{\vec{p}^2 + m^2c^2}\right)^3} = c\frac{d\vec{p}\cdot m^2c^2 + \left[\vec{p}\times[d\vec{p}\times\vec{p}]\right]}{\left(\sqrt{\vec{p}^2 + m^2c^2}\right)^3}; \end{gathered} \tag{28}$$

Coefficients $\gamma = E/mc^2$; $\vec{\beta} = \vec{v}/c$ differ from the above by constants, and satisfy similar relations.

The conservation laws reflect the homogeneity of space and time (see Mechanics): these natural laws do not change even if the origin of the coordinate system is shifted by δx. Then, $\delta x_i(A) = \delta x_i(B) = \delta x_i$. We can consider a closed system of particles (without continuous interaction, i.e., for most of the time they are free). Their action is sum of the individual actions, and

$$\sum_a \delta S_a = -(\sum_a m_a c u^i{}_a)\delta x_i\Big|_A^B = -(\sum_a m_a c u^i{}_a)\delta x_i\Big|_A^B = \{\sum_a p^i{}_a(A) - \sum_a p^i{}_a(B)\}\delta x_i = 0 \tag{29}$$

$$\sum_a p^i{}_a(A) = \sum_a p^i{}_a(B) = \left(\sum_a E_a / c, \sum_a \vec{p}\right) = const . \tag{30}$$

1.2 Particles in the 4-potential of the EM field.

The EM field propagates with the speed of light, i.e., it is a natural product of relativistic 4-D space-time; hence, the 4-potential is not an odd notion!

In contrast with the natural use of the interval for deriving the motion of the free relativistic particle, there is no clear guideline on what type of term should be added into action integral to describe a field. It is possible to consider some type of scalar function $\int A(x^i)ds$ [1] to describe electromagnetic fields, but this would result in wrong equations of motion. Nevertheless, the next guess is to use a product of 4-vectors $A^i dx_i$, and surprisingly it does work, even though we do not know why? **Hence**, **the fact that electromagnetic fields are fully described by the 4-vector of potential $\boxed{A^i = (A^0, \vec{A})}$ must be considered as an experimental fact!**

Nevertheless, it looks natural that the interaction of a charge with electromagnetic field is represented by the scalar product of two 4-vectors with the $-e/c$ coefficient chosen by convention:

$$S_{\text{int}} = -\frac{e}{c}\int_A^B A^i dx_i\,;\ A^i \equiv (A^o, \vec{A}) \equiv (\varphi, \vec{A}) \tag{31}$$

where the integral is taken along the particle's world line. A charge e and speed of the light c are moved outside the integral because they are constant; hence, **we use the conservation of the charge e and constancy of the speed of the light !**

IT IS ESSENTIAL THAT FIELD IS GIVEN, SINCE WE ARE CONSIDERING A PARTICLE INTERACTING WITH A GIVEN FIELD.

Turning our attention back to the Least-Action Principle and Hamiltonian Mechanics

The standard presentation of 4-potential is

$$A^i \equiv (A^0, \vec{A}) \equiv (\varphi, \vec{A})\,; \tag{32}$$

where φ is called the scalar potential and $\vec{A}$ is termed the vector potential of electromagnetic field.

Gauge Invariance. As we discussed earlier the action integral is not uniquely defined; we can add to it an arbitrary function of coordinates and time without changing the motion: $S' = S + f(x_i)$. This corresponds to adding the full differential of f in the integral (31)

$$S' = \int_A^B \left(-mcds - \frac{e}{c} A^i dx_i + dx_i \partial^i f \right).$$

This signifies that the 4-potential is defined with sufficient flexibility to allow the addition of any 4-gradient to it (let us choose $f(x_i) = \frac{e}{c} g(x_i)$)

$$A'^i = A^i - \partial^i g(x_i) = A^i - \frac{\partial g}{\partial x_i}; \tag{33}$$

without affecting the motion of the charge, a fact called **THE GAUGE INVARIANCE .**

We should be aware that the evolution of the system does not change but appearance of the

---

[1] You can check that this function will give the equations of motion $(mc - A)\frac{du^i}{ds} + \partial^i A = 0$.

equation of the motion for the system could change. For example, as follows from (33), the canonical momenta will change:

$$P'^{i} = P^{i} - \partial^{i} f \ .$$

Nevertheless, only the appearance of the system is altered, not its evolution. Measurable values (such as fields, mechanical momentum) do not depend upon it. One might consider Gauge invariance as an inconvenience, but, in practice, it provides a great opportunity to find a gauge in which the problem becomes more comprehensible and solvable.

The action is an additive function: therefore, the action of a charge in electromagnetic field is simply the direct sum of a free particle's action and action of interaction: (remember $ds = ds^2 / ds = dx^i dx_i / ds = u^i dx_i$)

$$S = \int_A^B \left(-mcds - \frac{e}{c} A^i dx_i \right) = \int_A^B \left(-mcu^i - \frac{e}{c} A^i \right) dx_i \tag{34}$$

Then the total variation of the action is

$$\begin{aligned} \delta S = \delta \int_A^B \left(-mcds - \frac{e}{c} A^i dx_i \right) = \int_A^B \left(-mc \frac{dx^i d\delta x_i}{ds} - \frac{e}{c} A^i d\delta x_i - \frac{e}{c} \delta A^i dx_i \right) = \\ -\left[\left(mcu^i + \frac{e}{c} A^i \right) \delta x_i \right]_A^B + \int_A^B \left(mc \frac{du^i}{ds} \delta x_i ds + \frac{e}{c} \delta x_i dA^i - \frac{e}{c} \delta A^i dx_i \right) = 0. \end{aligned} \tag{35}$$

That gives us a 4-momentum

$$P^i = -\frac{\delta S}{\delta x_i} = \left(mcu^i + \frac{e}{c} A^i \right) = \left(H / c, \vec{P} \right) = p^i + \frac{e}{c} A^i; \tag{36}$$

with

$$\begin{aligned} H = E = c(mcu^0 + \frac{e}{c} A^0) = \gamma mc^2 + e\varphi = c\sqrt{m^2c^2 + \vec{p}^2} + e\varphi; \\ \vec{P} = \gamma m\vec{v} + \frac{e}{c} \vec{A} = \vec{p} + \frac{e}{c} \vec{A}; \Rightarrow \vec{p} = \vec{P} - \frac{e}{c} \vec{A}. \end{aligned} \tag{37}$$

The Hamiltonian must be expressed in terms of generalized 3-D momentum, $\vec{P} = \vec{p} + \frac{e}{c} \vec{A}$ and it is

$$H(\vec{r}, \vec{P}, t) = c\sqrt{m^2c^2 + \left(\vec{P} - \frac{e}{c} \vec{A} \right)^2} + e\varphi; \tag{38}$$

with Hamiltonian equation following from it:

$$\vec{\mathrm{v}} = \frac{d\vec{r}}{dt} = \frac{\partial H}{\partial \vec{P}} = \frac{\vec{P}c - e\vec{A}}{\sqrt{m^2c^2 + \left(\vec{P} - \frac{e}{c} \vec{A} \right)^2}}$$

$$\frac{d\vec{P}}{dt}=\frac{d\vec{p}}{dt}+\frac{e}{c}\frac{d\vec{A}}{dt}=-\frac{\partial H}{\partial \vec{r}}=-e\vec{\nabla}\varphi-e\frac{\{(\vec{P}-\frac{e}{c}\vec{A})\cdot\vec{\nabla}\}\vec{A}}{\sqrt{m^2c^2+\left(\vec{P}-\frac{e}{c}\vec{A}\right)^2}}=-e\vec{\nabla}\varphi-\frac{e}{c}(\vec{v}\cdot\vec{\nabla})\vec{A};$$

From this equation we can derive (without any elegance!) the equation for mechanical momentum $\vec{p}=\gamma m\vec{v}$. We will not do it here, but rather we will use easier way to obtain the 4D equation of motion via the least-action principle. We fix A and B to get from equation (35)

$$\delta S=\int_A^B\left(mcu^i\delta x_i ds+\frac{e}{c}\delta x_i dA^i-\frac{e}{c}\delta A^k dx_k\right)=\int_A^B\left(mc\frac{du^i}{ds}\delta x_i ds+\frac{e}{c}\frac{\partial A^i}{\partial x_k}\delta x_i dx_k-\frac{e}{c}\frac{\partial A^k}{\partial x_i}\delta x_i dx_k\right)=$$
$$\int_A^B\left(\frac{dp^i}{ds}+\frac{e}{c}\left\{\frac{\partial A^i}{\partial x_k}-\frac{\partial A^k}{\partial x_i}\right\}u_k\right)\delta x_i ds=0. \tag{39}$$

As usual, the expression inside the round brackets must be set at zero to satisfy (39); i.e., we have the equations of charge motion in an electromagnetic field:

$$mc\frac{du^i}{ds}\equiv\frac{dp^i}{ds}=\frac{e}{c}F^{ik}u_k; \tag{40}$$

wherein we introduce an anti-symmetric **electromagnetic field tensor**

$$F^{ik}=\frac{\partial A^k}{\partial x_i}-\frac{\partial A^i}{\partial x_k}. \tag{41}$$

Electromagnetic field tensor: The Gauge Invariance can be verified very easily:

$$F'^{ik}=\frac{\partial A'^k}{\partial x_i}-\frac{\partial A'^i}{\partial x_k}=F^{ik}-\frac{\partial^2 g}{\partial x_i\partial x_k}+\frac{\partial^2 g}{\partial x_k\partial x_i}=F^{ik};$$

which means that the equation of motion (40) is not affected by the choice of the gauge, and the **electromagnetic field tensor is defined uniquely!** Using the Landau convention, we can represent the asymmetric tensor by two 3-vectors (see Appendix A):

$$F^{ik}=(-\vec{E},\vec{B});F_{ik}=(\vec{E},\vec{B});$$
$$F^{ik}=\begin{bmatrix}0 & -E_x & -E_y & -E_z\\ E_x & 0 & -B_z & B_y\\ E_y & B_z & 0 & -B_x\\ E_z & -B_y & B_x & 0\end{bmatrix}. \tag{42}$$

$\vec{E}$ is the so-called vector of the electric field and $\vec{B}$ is the vector of the magnetic field. Note the occurrence of the Lorentz group generator (see Appendix B) in (42). The 3D expressions of the field vectors can be obtained readily:

$$E^\alpha=F^{\alpha 0}=\frac{\partial A^0}{\partial x_\alpha}-\frac{\partial A^\alpha}{\partial x_0}-=-\frac{\partial\varphi}{\partial r_\alpha}-\frac{1}{c}\frac{\partial A^\alpha}{\partial t};\ \alpha=1,2,3;\vec{E}=-\frac{1}{c}\frac{\partial\vec{A}}{\partial t}-grad\varphi; \tag{43}$$

$$B^{\alpha} = -\frac{1}{2}e^{\alpha\kappa\lambda}F^{\kappa\lambda} = e^{\alpha\kappa\lambda}\left(\frac{\partial A^{\lambda}}{\partial x_{\kappa}} - \frac{\partial A^{\kappa}}{\partial x_{\lambda}}\right); \ \vec{B} = curl\vec{A}; \ F^{\kappa\lambda} = e^{\lambda\kappa\alpha}H_{\alpha}. \tag{44}$$

A 3D asymmetric tensor $e^{\alpha\kappa\lambda}$ and the *curl* definition are used to derive last equation and use Greek symbols for the spatial 3D components. The electric and magnetic fields are also Gauge invariant being components of Gauge invariant tensor.

We have the first pair of Maxwell's equations without further calculation using the fact that differentiation is symmetric operator ($\partial^i \partial^k \equiv \partial^k \partial^i$):

$$e_{iklm}\partial^k F^{lm} = e_{iklm}\partial^k\left(\partial^l A^m - \partial^m A^l\right) = 2e_{iklm}(\partial^k \partial^l)A^m = 0; \tag{45}$$

or explicitly:

$$\partial^k F^{lm} + \partial^l F^{mk} + \partial^m F^{kl} = 0. \tag{46}$$

A simple exercise gives the 3D form of the first pair of Maxwell equations. They also can be attained using (43) and (44) and known 3D equivalencies: $div(curl\vec{A}) \equiv 0$; $curl(grad\varphi) \equiv 0$:

$$\begin{aligned} &\vec{E} = -grad\varphi - \frac{1}{c}\frac{\partial\vec{A}}{\partial t}; && curl\vec{E} = -curl(grad\varphi) - \frac{1}{c}curl\frac{\partial\vec{A}}{\partial t} = -\frac{1}{c}\frac{\partial\vec{B}}{\partial t}; \\ &\vec{B} = curl\vec{A}; && div\vec{B} = div(curl\vec{A}) \equiv 0; \end{aligned} \tag{47}$$

I note that (47) is the exact 3D equivalent of invariant 4D Maxwell equations (45) that you may wish to verify yourself. There are 4 equations in (45): i=0,1,2,3. The *div* is one equation and *curl* gives three (vector components) equations. Even the 3D form looks very familiar; the beauty and relativistic invariance of the 4D form makes it easy to remember and to use.

EM Fields transformation, Invariants of the EM field. The 4-potential was defined as 4-vector and it transforms as 4-vector. The electric and magnetic fields, as components of the asymmetric tensor, follow its transformation rules (See Appendix A).

$$\begin{aligned} &\varphi = \gamma(\varphi' + \beta A_x'); \ A_x = \gamma(A_x' + \beta\varphi'); \\ &E_y = \gamma(E_y' + \beta B_z'); \ E_z = \gamma(E_z' - \beta B_y'); \\ &B_y = \gamma(B_y' - \beta E_z'); \ B_z = \gamma(B_z' + \beta E_y'). \end{aligned} \tag{48}$$

and the rest is unchanged. An important repercussion from these transformations is that the separation of the electromagnetic field in two components is an artificial one. They translate into each other when the system of observation changes and **MUST** be measured in the same units (Gaussian). The rationalized international system of units (SI) system measures them in V/m, Oe, A/m and T. Why not use also a horse power per square mile an hour, the old British thermal units as well? This makes about the same sense as using Tesla or A/m.

While the values and directions of 3D field components are frame-dependent, two 4-scalars can be build from the EM 4-tensor $F^{ik} = (-\vec{E}, \vec{B})$

$$F^{ik}F_{ik} = inv; \quad e^{iklm}F_{ik}F_{lm} = inv; \tag{49}$$

which in the 3D-form appear as

$$\vec{B}^2 - \vec{E}^2 = inv; \ (\vec{E} \cdot \vec{B}) = inv. \tag{50}$$

This conveys a good sense what can and cannot be done with the 3D components of electromagnetic fields. Any reference frame can be chosen and both fields transferred in a minimal number of components limited by (50). For example; 1) if $\left|\vec{E}\right| > \left|\vec{B}\right|$ in one system it is true in all systems and vice versa; and (2) if fields are perpendicular in one frame, $(\vec{E}\cdot\vec{B})=0$, this is true in all frames. When $(\vec{E}\cdot\vec{B})=0$ a frame can always be found where E or B are equal to zero (locally!).

Lorentz form of equation of a charged particle's motion.

The equations of motion (40) can rewritten in the form:

$$\frac{d\mathrm{E}}{dt}=c\frac{dp^0}{dt}=eF^{0k}\mathrm{v}_k=e\vec{E}\cdot\vec{\mathrm{v}};\qquad \mathrm{v}_k=\frac{dx_k}{dt}=(c,-\vec{\mathrm{v}})$$
$$\frac{d\vec{p}}{dt}=e\left(\hat{e}_\alpha F^{\alpha k}\frac{\mathrm{v}_k}{c}\right)=\frac{e}{c}\left(\hat{e}_\alpha\cdot cF^{\alpha 0}-\hat{e}_\alpha\cdot F^{\alpha\kappa}\mathrm{v}_k\right)=e\vec{E}+\hat{e}_\alpha e^{\alpha\kappa\lambda}B_\lambda\frac{\mathrm{v}_k}{c}=e\vec{E}+\frac{e}{c}\left[\vec{\mathrm{v}}\times\vec{B}\right]. \tag{51}$$

So, we have expressions for the generalized momentum and energy of the particle in an electromagnetic field. Generalized momentum is equal to the particle's mechanical momentum plus the vector potential scaled by *e/c.* The total energy of the charged particle is its mechanical energy, γmc^2, plus its potential energy ,$e\varphi$, in an electromagnetic field. The Standard Lorentz (not Hamiltonian!) equations of motion for $\vec{p}=\gamma m\vec{v}$ are

$$\frac{d\vec{p}}{dt}=e\vec{E}+\frac{e}{c}\left[\vec{\mathrm{v}}\times\vec{B}\right]. \tag{52}$$

with the force caused by the electromagnetic field (Lorentz force) comprised of two terms: the electric force, which does not depend on particle's motion, and, the magnetic force that is proportional to the vector product of particle velocity and the magnetic field, i.e., it is perpendicular to the velocity. Accordingly, the magnetic field does not change the particle's energy. We derived it in Eq. (51):

$$mc^2\frac{d\gamma}{dt}=e\vec{E}\cdot\vec{\mathrm{v}}; \tag{53}$$

Eqs. (52) and (53) are generalized equations. Using directly standard Lorentz equations of motion in a 3D form is a poor option. The 4D form is much better (see below) and, from all points of view, the Hamiltonian method is much more powerful!

It is worth noting that the 4D form of the charge motion (40) and its matrix form is the most compact one,

$$u^i=\frac{dx^i}{ds};\ mc\frac{du^i}{ds}=\frac{e}{c}F^i{}_k u^k;\Rightarrow\frac{d}{ds}[x]=[I]\cdot[u];\ \frac{d}{ds}[u]=\frac{e}{mc^2}[F]\cdot[u] \tag{54}$$

and, in many cases, it is very useful. We treat the ***x, u*** as a vectors, and ***[F]*** as the 4x4 matrix. ***[I]*** is just the unit 4x4 matrix It has interesting formal solution in the matrix form:

$$[u]=e^{\int\frac{e}{mc^2}[F]ds}[u_0];\ [x]=[x_o]+\left\lfloor\int dse^{\int\frac{e}{mc^2}[F]ds}\right\rfloor[u_0] \tag{55}$$

Its resolution is well defined when applied to the motion of a charged particle in uniform, constant EM field:

$$[u]=e^{\frac{e}{mc^2}[F](s-s_0)}[u_0];\ [x]=[x_o]+\left\lfloor\int e^{\frac{e}{mc^2}[F](s-s_0)}dse\right\rfloor[u_0] \tag{56}$$

The Lorentz group of theoretical physics (see Appendix B) is fascinating, and the fact that EM field tensor has the same structure as the generator of Lorentz group is no coincidence – rather, it is indication that physicists have probably come very close to the roots of nature in this specific direction. This statement is far from truth for other fundamental forces and interactions.

To conclude this subsection, we will take one step further from (54) and write a totally linear evolution equation for a combination of 4D vectors

$$\frac{d}{ds}\begin{bmatrix} x \\ u \end{bmatrix} = [\Lambda]\cdot\begin{bmatrix} x \\ u \end{bmatrix};\ [\Lambda] = \begin{bmatrix} 0 & I \\ 0 & \frac{e}{mc^2}F \end{bmatrix} \tag{57}$$

where [Λ] is an 8x8 degenerated matrix. Similarly to (55) and (56)

$$\begin{bmatrix} x \\ u \end{bmatrix} = e^{\int[\Lambda]ds}\cdot\begin{bmatrix} x \\ u \end{bmatrix}_o;\ \begin{bmatrix} x \\ u \end{bmatrix} = e^{[\Lambda](s-s_o I}\cdot\begin{bmatrix} x \\ u \end{bmatrix}_o \ for\,[\Lambda] = const; \tag{58}$$

First pair of Maxwell's equations (a little more of juice)

We will derive full set of Maxwell equations using the least action principle. Nevertheless, you can consider the Maxwell equation as given - in any case they were derived originally from numerous experimental laws!

First pair of Maxwell's equations is the consequence of definitions of electric and magnetic field through the 4-potential:

$$\vec{E} = -grad\varphi - \frac{1}{c}\frac{\partial\vec{A}}{\partial t};\quad \vec{H} = curl\vec{A};$$

it is equivalent to

$$curl\vec{E} = -curl(grad\varphi) - \frac{1}{c}curl\frac{\partial\vec{A}}{\partial t} = \frac{1}{c}\frac{\partial\vec{H}}{\partial t};\quad div\vec{H} = div(curl\vec{A}) \equiv 0; \tag{59}$$

Nevertheless, it is very important to remember that they are actually originated from experiment. First Maxwell equation is the Faraday law and the second is nothing else that absence of magnetic charge! You should remember all time that inclusion of the term $S_{\text{int}} = -\frac{e}{c}\int_A^B A^i dx_i$ into action integral is consequence of experiment! Thus, the first pair of Maxwell equations governing the electromagnetic fields is:

$$curl\vec{E} = -\frac{1}{c}\frac{\partial\vec{H}}{\partial t}; \tag{60}$$

$$div\vec{H} = 0; \tag{61}$$

with well known integral ratios following it:

Gauss' theorem:

$$\oint\vec{H}d\vec{a} = \int div\vec{H}dV = 0; \tag{62}$$

Stokes' theorem:

$$\oint\vec{E}d\vec{l} = \int curl\vec{E}d\vec{a} = -\frac{1}{c}\frac{\partial}{\partial t}\int\vec{H}d\vec{a}; \tag{63}$$

where $d\vec{a}$ is vector of the element of the surface and $d\vec{l}$ is a vector of a contour length. Integral equations read: the

1) Flux of the of the magnetic field though the surface covering any volume V is equal zero;
2) The circulation of electric field around the contour (electromotive force) is equal to the derivative of the magnetic flux though the contour scaled down by "-c" - the Faraday law.

7.2 Action of the electromagnetic field.

As we discussed earlier, in the relativistic picture of the world, the field acquires its it's own physical reality. Therefore, the action of whole system including a particle and a field must consist of three parts: the action of free particle, the action of free field and the action their interaction:

$$S = S_p + S_f + S_{pf}. \tag{64}$$

We already got first and last term. For a several free particles, the action is the direct sum of individual actions:

$$S_p = -\sum_p mc\int ds; \tag{65}$$

and interaction with the field is the sum of their individual interactions:

$$S_{pf} = -\sum_p \frac{e}{c}\int A^i dx_i. \tag{66}$$

The sum of (65) and (66) gives us equation of particle's motion in "external", i.e. pre-defined electromagnetic fields. Now we want to know how charged particles influence the EM field and how EM field evolves on its own? We do not know, also, what defines properties of a free field? First pair of Maxwell equations gives us only two connections: the time derivative of the magnetic field and its divergence (zero). We still don't know what is time derivative of electric field and what is its divergence?

Please remember that all following discussion must be considered as a logical excise. Final form of the field action has to have the most important property: it must satisfy the experimental observations! Where to start to get them?

One of the most important properties of the field confirmed by experiments is **the Principle of Superposition:**

the resulting field produced by various sources is a simple composition (the direct sum) of the fields produced by individual sources! It means that resulting electric and magnetic fields are vector sum of individual fields. Thus, we have a clue that we should look for type of equations, which allows superposition of solutions, i.e. linear differential field equations[1]. In order to generate linear differential equations, the action should contain quadratic expression of the field components[2], which described by field 4-tensor F^{ik}.

[1] In field theory the 4-vector of the field A^i is coordinate of the field. Therefore, field's 4-tensor is first order derivative of the coordinates. According to Hamiltonian principle, the action could have under integral only coordinates and their first derivatives. This requirement excludes derivatives of F^{ik} from the action's integral.

[2] 4-vector of the field A^i is not unique (Gauge transformation) and trial function comprising 4-vector of the field will give non-unique equation of the field. The difference with interaction term is that last includes first order of 4-potential and non-uniqueness does not affect equation of motions. Situation is not the same for quadratic term! A variation acts in similar manner as a differentiation - "to get linear ($2x^1$) we need to differentiate (x^2)".

In addition, the action must be 4-invariant (4-scalar, not pseudo-scalar!), which leaves us with $F^{ik}F_{ik} = 2(\vec{H}^2 - \vec{E}^2)$. Finally, the field is "an entity leaving" in space and time coordinates. In order to describe total field we should integrate over all space between two "time" events $d\Omega = dx^0 dx^1 dx^2 dx^3 = cdtdV$ which is 4-invariant: $d\Omega = e_{iklm} dx_a^i dx_b^k dx_c^l dx_d^m$ where a,b,c,d four 4-vectors defining element of 4-volume. Therefore, a probable form of the action of the EM field is:

$$S_f = -a\int F^{ik} F_{ik} d\Omega. \tag{67}$$

The choice of the coefficient before integral is equivalent to the choice of the units to measure the field. In the Gaussian system of units, which we are using, fields are measured in Gs and coefficient is

$$a = \frac{1}{16\pi \cdot c}. \tag{68}$$

The total action is:

$$S = -\sum_p mc\int ds - \sum_p \frac{e}{c}\int A^i dx_i - \frac{1}{16\pi c}\int F^{ik} F_{ik} d\Omega. \tag{69}$$

4-current and equation of continuity. The conservation of the charge should affect our equations. Let's make a glance on this issue and write a charge conservation law in the form useful for future derivation of the field equations. It is very useful to describe charges by a distribution function. The charge density ρ is defined as the charge contained in unit volume:

$$de = \rho dV; \tag{70}$$

and microscopic (exact in classical EM) definition of ρ is sum of Dirac's delta-functions:

$$\rho = \sum_a e_a \delta(\vec{r} - \vec{r}_a); \tag{71}$$

where index *a* is index to count particles. 4-vector of current is defined as:

$$j^i = \rho \frac{dx^i}{dt}. \tag{72}$$

The fact that j^i is a 4-vector comes from equivalence:

$$dedx^i = \rho \frac{dx^i}{dt} \cdot dtdV = \rho \frac{dx^i}{dt} \cdot d\Omega; \tag{73}$$

and the fact that charge is 4-scalar or invariant (experimental fact) and $d\Omega \equiv dVdt$ is the 4-scalar. Thus:

$$j^i = (\rho c, \vec{j}); \vec{j} = \rho \vec{v}. \tag{74}$$

To be exact, for point charges, the 4--current is:

$$j^i = \sum_a e_a \delta(\vec{r} - \vec{r}_a) \frac{dx_a^i}{dt}. \tag{75}$$

It is the microscopic 4-current for ensemble of particles. When it is necessary, it can be averaged over a "small volume" for macroscopic description. We do not need averaging now and can comfortably use Eq. (75). Our goal is to get the equation of continuity:

$$\partial_i j^i = \frac{\partial j^i}{\partial x^i} = \left(\frac{\partial \rho}{\partial t} + div\vec{j}\right) = 0; \tag{76}$$

which is resulting from charge conservation. It is easy to do for microscopic distribution (75):

$$\partial_i j^i = \sum_a e_a \left\{ \frac{\partial}{c\partial t}\left(\delta(\vec{r} - \vec{r}_a(t))c\right) + div\left(\left(\delta(\vec{r} - \vec{r}_a(t))\vec{v}_a\right)\right) \right\} = \sum_a e_a \vec{\nabla}\delta(\vec{r} - \vec{r}_a)\cdot\left\{ -\frac{\partial \vec{r}_a(t)}{\partial t} + \vec{v}_a \right\} \equiv 0; \tag{77}$$

with $\partial^i = (\partial / \partial ct, \partial / \partial \vec{r}); \partial / \partial \vec{r}(r_a(t)) \equiv 0$ and we use derivative of Dirac's delta-function. Now we are ready for next trick, i.e. to present action of the interaction as integral of 4-current:

$$\begin{gathered} e_a = \int e_a \delta(\vec{r} - \vec{r}_a) dV; \quad \frac{1}{c}\int \sum_a e_a A_k dx^k = \\ \frac{1}{c}\int A_k \sum_a e_a dx^k \delta(\vec{r} - \vec{r}_a) dV = \frac{1}{c}\int A_k j^k \, dt \, dV = \frac{1}{c^2}\int A_k j^k \, d\Omega \end{gathered} \tag{78}$$

Side note: Today we are using the method, which is standard for all modern filed theories: QED, QCD, SUSY, etc. In self-consistent theories, particles become fields as well. In QED, an electron is not a point particle but a "wave" described by 4-spinor ψ. We can include this into our action very easy by writing correct QED current in the interaction term (78)

$$j^i = \overline{\psi}\gamma^i\psi .$$

In this case, the current is a continuous function of the space and time. It is a better way that having Dirac's delta-function. The nature of the current, as we would see, does not change equation of the field motion. It means that Maxwell equations do not change when we introduce quantum description of charges! In this case, the equation of the charges motion should be also proper, i.e. those derived by Dirac:

$$\left\lfloor \gamma^i \left(p_i - \frac{eA_i}{c} \right) - m \right\rfloor \gamma^0 \psi = 0 .$$

I would not go into details of Dirac's description of electron and his 4x4 γ-matrices. If you are interested, look through one of many QED books. Thus, equivalent form of (7-12) is:

$$S = -\sum_p mc \int ds - \frac{1}{c^2}\int A_k j^k d\Omega - \frac{1}{16\pi c}\int F^{ik} F_{ik} d\Omega. \tag{79}$$

Second pair of Maxwell's equations: more of the least action…

We already found equation of charges motion in the field. Let's consider all charges following their equation of motion

$$\delta_{for\ particles}\left(\sum_p mc\int ds + \int A_k j^k d\Omega\right) = 0 . \tag{80}$$

Let's changes move along their real trajectories. Now we will vary only the field to find its equations of motion:

$$\delta S = -\frac{1}{16\pi c^2}\int\left(16\pi\delta A_i j^i + c\delta(F^{ik}F_{ik})\right)d\Omega = -\frac{1}{8\pi c^2}\int\left(8\pi\delta A_i j^i + cF^{ik}\delta F_{ik}\right)d\Omega = 0; \tag{81}$$

where we use

$$F^{ik}\delta F_{ik} = \delta F^{ik}F_{ik} . \tag{82}$$

It is important to remember that we can vary both particle's trajectories and field if we wish. It will give us two terms in the variation of the action: one containing variation of the trajectories

$$\delta S_{part} = \sum_a \int_A^B \left(\frac{dp_a{}^i}{ds} + \frac{e}{c}\left\{\frac{\partial A^i}{\partial x_k} - \frac{\partial A^k}{\partial x_i}\right\}u_k\right)\delta x_{a\,i} ds \tag{83}$$

and the other containing variation of the field. Variations for each particle and the field are independent. Therefore, each independent component of action's variation must be equal zero. (83) will give us again equation of particle's motion, while field terms (81) will bring us to the field equations. Let's rewrite second term in (81):

$$F_{ik} = \partial_i A_k - \partial_k A_i$$

$$F^{ik}\delta F_{ik} = F^{ik}\partial_i\delta A_k - F^{ik}\partial_k\delta A_i = -F^{ki}\partial_i\delta A_k - F^{ik}\partial_k\delta A_i = -2F^{ik}\partial_k\delta A_i . \tag{84}$$

Now we can integrate by parts:

$$\begin{aligned}\delta S &= -\frac{1}{4\pi c^2}\int\left(4\pi\delta A_i j^i - cF^{ik}\partial_k\delta A_i\right)d\Omega = \\ &-\frac{1}{4\pi c^2}\int\left(4\pi j^i + c\partial_k F^{ik}\right)\delta A_i\, d\Omega - \frac{1}{4\pi c}\int F^{ik}\delta A_i\, dS_k\bigg|_{suraface\, of\, \Omega} = 0;\end{aligned} \tag{85}$$

with second integral obtained by 4D Gauss theorem:

$$\int div_4 A^i d\Omega = \oint A^i dS_i , \tag{86}$$

where dS_i is element of hyper-surface surrounding 4-volume Ω. It is not so essential, how it looks. One simple case: we integrate over all space and fixed time interval *(t1,t2)*. Surface of the *W* is full 3D space at moments of t_1, t_2. The least action method calls for zero variations on the boundaries $\delta A_i|_{surafaceof\,\Omega} = 0$ and second integral in (85) disappears leaving us with:

$$\delta S = -\frac{1}{4\pi c^2}\int\left(4\pi j^i + c\partial_k F^{ik}\right)\delta A_i d\Omega = 0\,. \tag{87}$$

Please notice that we are left only with variations of 4-potential. It is very natural because variations of 4-potential fully define field's variations. Equation (87) gives us "second pair" of Maxwell equations in 4D form:

$$\frac{\partial F^{ik}}{\partial x^k} = -\frac{4\pi}{c} j^i \tag{88}$$

3D form follows directly from (88) and form of the field tensor:

$$F^{lm} = \begin{vmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -H_z & H_y \\ E_y & H_z & 0 & -H_x \\ E_z & -H_y & H_x & 0 \end{vmatrix}; \tag{89}$$

and yields:

$$div\vec{E} = 4\pi\rho; \tag{90}$$

$$curl\vec{H} = \frac{4\pi}{c}\vec{j} + \frac{1}{c}\frac{\partial\vec{E}}{\partial t}. \tag{91}$$

Integral equations are obvious applications of Stokes and Gauss theorems to Eqs (90-91)

$$\oint \vec{E}d\vec{a} = 4\pi\int\rho dV; \tag{92}$$

$$\oint \vec{H}d\vec{l} = \frac{1}{c}\int(4\pi\vec{j} + \frac{\partial\vec{E}}{\partial t})d\vec{a}\,. \tag{93}$$

Equivalent forms of Maxwell equations:

$$\begin{aligned}\vec{E} &= -grad\varphi - \frac{1}{c}\frac{\partial\vec{A}}{\partial t}; \\ \vec{H} &= curl\vec{A};\end{aligned} \quad \Leftrightarrow \quad F^{ik} = \frac{\partial A^k}{\partial x_i} - \frac{\partial A^i}{\partial x_k}. \tag{94}$$

Compact 4-D:

$$e^{iklm}\frac{\partial F_{lm}}{\partial x^k}=0; \qquad \oplus \qquad \frac{\partial F^{ik}}{\partial x^k}=-\frac{4\pi}{c}j^i; \tag{95}$$

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:

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

Consider a non-degenerated transformation in 4D space

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

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

and allowing the inverse transformation

$$\begin{aligned} &X = X(X') \\ &x^i = x^i(x'^0, x'^1, x'^2, x'^3);\ i = 0,1,2,3 \end{aligned} \tag{A-4}$$

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

$$\frac{\partial x'^i}{\partial x^j};\ \frac{\partial x^j}{\partial x'^i}; \tag{A-5}$$

and are orthogonal to each other

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

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

$$a^i b_i \equiv \sum_{i=0}^{i=3} a^i b_i\ . \tag{A-7}$$

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

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

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 alsoare 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:

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

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

$$A \cdot B = A_i B^i ; \tag{A-16}$$

It is the invariant of transformations:

$$A'_i B'^i = \frac{\partial x^j}{\partial x'^i}\frac{\partial x'^i}{\partial x^k} A_j B^k = \frac{\partial x^j}{\partial x^k} A_j B^k = \delta^j_k A_j B^k = A_k B^k ; \tag{A-17}$$

where

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

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

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

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:

$$ds^2 = dx^{0^2} - dx^{1^2} - dx^{2^2} - dx^{3^2} = dx_0{}^2 - dx_1{}^2 - dx_2{}^2 - dx_3{}^2 ; \tag{A-20}$$

and the diagonal metric tensor g^{ik} :

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

$$g_{ik} = g^{ik}; \; g^{00} = 1; g^{11} = -1; g^{22} = -1; g^{33} = -1; \tag{A-21}$$

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

$$x^i = g^{ik}x_k\,;\; x_i = g_{ik}x^k\,; \tag{A-22}$$

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

$$F^{\dots k \dots\dots}_{\dots\dots i\,\dots} = g^{kj}F^{\dots\dots\dots\dots}_{\dots j\,.i\dots} = g^{kj}g_{il}F^{\dots\dots l\,\dots}_{\dots j\,\dots};etc. \tag{A-23}$$

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 :

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

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

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 $\delta_i^{\ k}$:

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

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 $J=\det\left[\frac{\partial x'}{\partial x}\right]$.

$$e'^{iklm} = \frac{\partial x'^i}{\partial x^j}\frac{\partial x'^k}{\partial x^n}\frac{\partial x'^l}{\partial x^p}\frac{\partial x'^m}{\partial x^q}e^{jnpq} = \det\left[\frac{\partial x'}{\partial x}\right]e^{jnpq}\delta^i_j\delta^k_n\delta^l_p\delta^m_q = e^{iklm}; \tag{A-28}$$

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

$$e'^{ik} = \frac{\partial x'^i}{\partial x^j}\frac{\partial x'^k}{\partial x^n}e^{jn} = \frac{\partial x'^i}{\partial x^1}\frac{\partial x'^k}{\partial x^2}e^{12} + \frac{\partial x'^i}{\partial x^2}\frac{\partial x'^k}{\partial x^1}e^{21} = \frac{\partial x'^i}{\partial x^1}\frac{\partial x'^k}{\partial x^2} - \frac{\partial x'^i}{\partial x^2}\frac{\partial x'^k}{\partial x^1} = \det\begin{Bmatrix} \frac{\partial x'^i}{\partial x^1} & \frac{\partial x'^i}{\partial x^2} \\ \frac{\partial x'^k}{\partial x^1} & \frac{\partial x'^k}{\partial x^2} \end{Bmatrix}; \quad \text{(A-29)}$$

Therefore:

$$e'^{ii} = \det\begin{Bmatrix} \frac{\partial x'^i}{\partial x^1} & \frac{\partial x'^i}{\partial x^2} \\ \frac{\partial x'^i}{\partial x^1} & \frac{\partial x'^i}{\partial x^2} \end{Bmatrix} = 0 = e^{ii}; e'^{12} = \det\begin{Bmatrix} \frac{\partial x'^1}{\partial x^1} & \frac{\partial x'^1}{\partial x^2} \\ \frac{\partial x'^2}{\partial x^1} & \frac{\partial x'^2}{\partial x^2} \end{Bmatrix} = 1 = e^{12}; e'^{21} = \det\begin{Bmatrix} \frac{\partial x'^2}{\partial x^1} & \frac{\partial x'^2}{\partial x^2} \\ \frac{\partial x'^1}{\partial x^1} & \frac{\partial x'^1}{\partial x^2} \end{Bmatrix} = -1 = e^{21}; \quad \text{(A-30)}$$

for rotations when $\det\begin{Bmatrix} \frac{\partial x'}{\partial x} \end{Bmatrix} = 1$. 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 $\vec{A} \to -\vec{A}$; $\vec{B} \to -\vec{B}$; $\vec{C} \Rightarrow \vec{C}$!

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

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

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.

$$A^{*ik} = e^{iklm}A_{lm} \quad \text{(A-32)}$$

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:

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

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

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

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:

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

Appendix B: Lorentz group http://en.wikipedia.org/wiki/Lorentz_group

Lorentz Group - Matrix representation

Jackson's Classical Electrodynamics, Section 11.7 [CED] has an excellent discussion of this topic. Here, we will review it briefly with some attention to the underlying mathematics. Generic Lorentz transformation involves a boost (a transformation from K to K' moving with some velocity $\vec{V}$) and an arbitrary rotation in 3D space. Matrix representation is well suited to describe 4-vectors transformations. The coordinate vector is defined as

$$X = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}; \tag{B-29}$$

and standard scalar product of 4-vectors is defined by $(a,b) = \tilde{a}b$, where $\tilde{a}$ is the transposed vector. The 4-scalar product involves the metric tensor (matrix):

$$a \cdot b \equiv a^i \cdot b_i = (a, gb) = (ga, b) = \tilde{a}gb; \tag{B-30}$$

$$g = \tilde{g} = \{g^{ik}\} = \{g_{ik}\} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{B-31}$$

Lorentz transformations A (or the group of Lorentz transformations[1]) are linear transformations that preserve the interval, or scalar product (B-30):

$$X' = AX;\ \tilde{X}'gX' = \tilde{X}\tilde{A}gAX = \tilde{X}gX; \Rightarrow \tilde{A}gA = g. \tag{B-33}$$

Using standard ratios for matrices

$$\det(\tilde{A}gA) = \det^2 A \det g = \det g \ \Rightarrow \det A = \pm 1; \tag{B-34}$$

we find that the matrices of Lorentz transformation have det=±1. We will consider only *proper Lorentz transformations* with unit determinants $\det A = +1$. Improper Lorentz transformations, like space- and time-inversions, should be considered as special transformations and added to the proper ones.

A 4x4 matrix has 16 elements. Equation (B-33) limited number of independent elements in matrix A of Lorentz transformations. Matrices on both sides are symmetric. Thus, there are 10 independent conditions on matrix A, leaving six independent elements there. This is unsurprising since

---

[1] Group G is defined as a set of elements , with a definition of a product of any two elements of the group; $P = A \bullet B \in G$; $A, B \in G$. The product must satisfy the associative law : $A \bullet (B \bullet C) = (A \bullet B) \bullet C$; there is an unit element in the group $E \in G; E \bullet A = A \bullet E = A$: $\forall A \in G$; and inverse elements: $\forall A \in G; \exists B(called\ A^{-1}) \in G$: $A^{-1}A = AA^{-1} = E$.

Matrices NxN with non-zero determinants are examples of the group. Lorentz transformations are other examples : the product of two Lorentz is defined as two consequent Lorentz transformations. Therefore, the product also is a Lorentz transformation whose velocity is defined by rules discussed in previous lectures. The associative law is straightforward: unit Lorentz transformation is a transformation into the same system. Inverse Lorentz transformation is a transformation with reversed velocity. Add standard rotation s, to constitute the Lorentz Group rotation in 3D space is represented by 3 angles and a boost is represented by 3 components of velocity. Intuitively, then there are six independent rotations: (xy), (yz), (zx), (t, x), (t, y), and (t, z). No other combinations of 4D coordinates are possible: $C_4^2 = \frac{4!}{2!2!} = 6$.

We next consider the properties of A in standard way, representing A through a generator L:

$$A = e^L; \tag{B-35}$$

where we use matrix exponent defined as the Taylor expansion:

$$e^L \downarrow_{def} \equiv \sum_{n=0}^{\infty} \frac{L^n}{n!}; L^0 = I; I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \tag{B-36}$$

where I is the unit matrix. Using (B-35) and $g^2 = I$ we find how to compose the inverse matrix for A:

$$\tilde{A} g A = g \Rightarrow A^{-1} = g \tilde{A} g; \tag{B-37}$$

which, in combination with

$$\tilde{A} = transpose(e^L) = \sum_{n=0}^{\infty} \frac{\tilde{L}^n}{n!} = e^{\tilde{L}}; e^{gUg} = \sum_{n=0}^{\infty} \frac{(gUg)^n}{n!} = \sum_{n=0}^{\infty} g \frac{U^n}{n!} g; \tag{B-38}$$

gives

$$A^{-1} = g\tilde{A}g = e^{g\tilde{L}g}. \tag{B-39}$$

We can show that matrix exponent has similar properties as the regular exponent, i.e. $e^U e^{-U} = I$ by explicitly using Taylor expansion to collect the powers of U:

$$e^U e^{-U} = \left(\sum_{n=0}^{\infty} \frac{U^n}{n!} \right) \left(\sum_{k=0}^{\infty} (-1)^k \frac{U^k}{k!} \right) = \sum_{k=0, n=0}^{\infty} (-1)^k \frac{U^{n+k}}{n!k!} = I + \sum_{m=1}^{\infty} c_m U^m; \tag{B-40}$$

and the well-known expansion of $(1-x)^m$. Our goal is to show that all c_m are zero:

$$(1-x)^m = \sum_{n=0}^{m} \frac{(-1)^n m!}{n!(m-n)!} x^n \Rightarrow m! c_m = \sum_{n=0}^{m} \frac{(-1)^n m!}{n!(m-n)!} = (1-1)^m = 0. \tag{B-41}$$

Now (B-39) can be rewritten

$$A^{-1} = g\tilde{A}g = e^{g\tilde{L}g} = e^{-L} \Rightarrow g\tilde{L}g = -L; \Rightarrow \widetilde{gL} = -gL \tag{B-42}$$

Hence, gL is an asymmetric matrix and has six independent elements as expected:

$$gL = \begin{bmatrix} 0 & L_{01} & L_{02} & L_{03} \\ -L_{01} & 0 & -L_{12} & -L_{13} \\ -L_{02} & L_{12} & 0 & -L_{23} \\ -L_{03} & L_{13} & L_{23} & 0 \end{bmatrix}; L = g(gL) = \begin{bmatrix} 0 & L_{01} & L_{02} & L_{03} \\ L_{01} & 0 & L_{12} & L_{13} \\ L_{02} & -L_{12} & 0 & L_{23} \\ L_{03} & -L_{13} & -L_{23} & 0 \end{bmatrix}. \tag{B-43}$$

Each independent element represents an irreducible (fundamental) element of the Lorentz group or rotations and boosts, as discussed above. The six components of the L can be considered as six components of 3-vectors in the form ("-" is a convention):

$$L = -\vec{\omega}\vec{S} - \vec{\varsigma}\vec{K}; A = e^{-\vec{\omega}\vec{S} - \vec{\varsigma}\vec{K}}; \tag{B-44}$$

with

$$\vec{S} = \hat{e}_x \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} + \hat{e}_y \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} + \hat{e}_z \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \tag{B-45}$$

$$\vec{K} = \hat{e}_x \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \hat{e}_y \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \hat{e}_z \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}; \tag{B-46}$$

where $\vec{\omega}\vec{S}$ represents the orthogonal group of rotations in 3D space (O_3^+), and $\vec{\varsigma}\vec{K}$ represents the boosts caused by transformation into a moving system. It is easy to check that these matrices satisfy commutation rules of

$$[S_i, S_k] = e_{ikl}S_l; [S_i, K_k] = e_{ikl}K_l; [K_i, K_k] = -e_{ikl}S_l;\ [A,B] \equiv AB - BA; \tag{B-47}$$

where e_{ikl} is the totally asymmetric 3D-tensor. You should be familiar with 3D rotation $e^{-\vec{\omega}\vec{S}}$ by $\vec{\omega}$: the direction of $\vec{\omega}$ is the axis of rotation and the value of $\vec{\omega}$ is the angle of rotation.

For the arbitrary unit vector $\hat{e}$

$$(\hat{e}\vec{S})^3 = -\hat{e}\vec{S};\ (\hat{e}\vec{K})^3 = \hat{e}\vec{K}. \tag{B-48}$$

Therefore, $\vec{S}$ "behaves" as an imaginary "i" and we should expect *sin* and *cos* to be generated by exp(..$\vec{S}$.); exp(.. $\vec{K}$.) should generate hyperbolic functions *sinh* and *cosh.* It is left for your homework to show, in particular, that boost transformation is:

$$A(\vec{\beta} = \vec{V} / c) = e^{-\vec{\beta}\vec{K}\tanh^{-1}\beta}. \tag{B-49}$$

Finally, all fully relativistic phenomena naturally have six independent parameters. For example, electromagnetic fields are described by two 3D vectors: the vector of the electric field and that of the magnetic field, or in equivalent form of an asymmetric 4-tensor of an electromagnetic field with six components. Furthermore, electric fields give charged particles energy boosts, while magnetic field rotates them without changing the energy….

Not surprisingly, the EM fields reflect the structure of the 4D space and its transformations.

Books where this and more material can be found:

L.D. Landau, E.M. Lifshitz, "The Classical Mechanics", Pergamon Press

L.D. Landau, E.M. Lifshitz, "The Classical Theory of Fields", Pergamon Press, 4th Revised Edition, 1975

J.D.Jackson, "Classical Electrodynamics", Second edition, p.536