Lecture 7. *Last updated 07.03.10*

VII. RIGOROUS DERIVATION OF EFES

A. The principle of the least action

The derivation of EFEs is very important material for understanding GR. In this lecture we will derive rigorously the Einstein Field equations (EFEs) from the principle of the least action.
This principle says that

$$\delta(S_g + S_m) = 0, \tag{VII.1}$$

where S_g and S_m are the actions of gravitational field and matter respectively. Taking into account that we are going to derive EFEs, the subject of variations is all components of the metric tensor.
To derive EFEs we should understand what are S_g and S_m .

B. The action function for the gravitational field

First of all S_g should depend on configuration of gravitational field, or geometry, in the whole space-time, hence it should be expressed in terms of a scalar integral over the all space and over the time coordinate between two given moments of time

$$S_g = \int G d\tilde{\Omega}, \tag{VII.2}$$

where $d\tilde{\Omega}$ is invariant element of 4-volume (see Lecture 3) and G is some scalar function called the action density.
We know that the final equations should contain derivatives of g_{ik} no higher than the second. Otherwise we could not obtain Newtonian Poisson's equation (see the previous lecture). In other words, G must contain only g_{ik} and Γ^l_{mn}, i.e

$$G = G(g_{ik}, \Gamma^i_{kl}). \tag{VII.3}$$

Immediately we confront with the following problem : this is impossible to construct the scalar from g_{ik} and Γ^l_{mn} . The only scalar in gravitational field, the scalar curvature R, contains the second derivatives of g_{ik}. Fortunately, there is rather simple resolution of this paradox: R is linear with respect to the second derivatives and for this reason, as we will see later, all terms containing second derivatives don't contribute to the variations of the action. Let us write the action function in the following form

$$S_g = \alpha \int R\sqrt{-g} d\Omega, \tag{VII.4}$$

where α is a constant which will be determined later.

Because of the linearity of R with respect to the second derivatives, the invariant action function can be transformed in the following way

$$S_g = \alpha \int R\sqrt{-g}d\Omega = \alpha \int G\sqrt{-g}d\Omega + \alpha \int w^l_{,l}d\Omega, \tag{VII.5}$$

where G contains only g_{ik} and $g_{ik,n}$, w is a function which we can be obtained by straightforward calculations:

$$\sqrt{-g}R = \sqrt{-g}g^{ik}R_{ik} = \sqrt{-g}\left\{g^{ik}\Gamma^l_{ik,l} - g^{ik}\Gamma^l_{il,k} + g^{ik}\Gamma^l_{ik}\Gamma^m_{lm} - g^{ik}\Gamma^m_{il}\Gamma^l_{km}\right\}, \tag{VII.6}$$

obviously

$$\sqrt{-g}g^{ik}\Gamma^l_{ik,l} = (\sqrt{-g}g^{ik}\Gamma^l_{ik})_{,l} - \Gamma^l_{ik}(\sqrt{-g}g^{ik})_{,l} \tag{VII.7}$$

and

$$\sqrt{-g}g^{ik}\Gamma^l_{il,k} = (\sqrt{-g}g^{ik}\Gamma^l_{il})_{,k} - \Gamma^l_{il}(\sqrt{-g}g^{ik})_{,k} = (\sqrt{-g}g^{il}\Gamma^k_{ik})_{,l} - \Gamma^k_{ik}(\sqrt{-g}g^{il})_{,l}. \tag{VII.8}$$

Then we obtain

$$\sqrt{-g}R = (\sqrt{-g}g^{ik}\Gamma^l_{ik} - \sqrt{-g}g^{il}\Gamma^k_{ik})_{,l} + \sqrt{-g}G = w^l, l + \sqrt{-g}G, \tag{VII.9}$$

where

$$w^l = \sqrt{-g}(g^{ik}\Gamma^l_{ik} - g^{il}\Gamma^k_{ik}) \tag{VII.10}$$

and

$$\sqrt{-g}G = \Gamma^m_{im}(\sqrt{-g}g^{ik}), k - \Gamma^l_{ik}(\sqrt{-g}g^{ik})_{,l} - (\Gamma^m_{il}\Gamma^l_{km} - \Gamma^l_{ik}\Gamma^m_{lm})\sqrt{-g}g^{ik} \tag{VII.11}$$

$$\Gamma^i_{ki} = \frac{1}{2}g^{im}\frac{\partial g_{im}}{\partial x^k}. \tag{VII.12}$$

According to the Gauss' theorem the volume integral of a full derivative is reduced to the integral over boundary. Taking into account that our objective is to obtain proper equations by applying the principle of the least action, we should keep all boundary conditions fixed. Hence, w disappears after variation. As a result

$$\delta \int R\sqrt{-g}d\Omega = \delta \int G\sqrt{-g}d\Omega. \tag{VII.13}$$

Thus we don't need G any more, because we proved that the variation of the integral with R is the same as the variation of the integral with G, hence we can work with R only.

$$\delta \int R\sqrt{-g}d\Omega = \delta \int g^{ik}R_{ik}\sqrt{-g}d\Omega = \int \{R_{ik}\sqrt{-g}\delta g^{ik} + g^{ik}R_{ik}\delta(\sqrt{-g}) + g^{ik}\sqrt{-g}\delta R_{ik}\}d\Omega. \tag{VII.14}$$

There are three terms in the variation of the action function. Let us first calculate the second term.

$$\delta(\sqrt{-g}) = -\frac{1}{2\sqrt{-g}}\delta g = -\frac{1}{2\sqrt{-g}}\frac{\partial g}{\partial g_{ik}}\delta g_{ik} = -\frac{1}{2\sqrt{-g}}M^{ik}\delta g_{ik}, \tag{VII.15}$$

where M^{ik} is the minor of the determinant g corresponding to the component g_{ik}. Indeed, the determinant g depends on all components g_{ik}. Calculating g with the help, say the first raw, one can write $g = M^{1i}g_{1i}$, where M^{1i} are minors of the components in the first row. Obviously M^{1i} do not contain g_{1i}. Hence

$$\frac{\partial g}{\partial g_{1i}} = M^{1i}. \tag{VII.16}$$

This is true for any row in determinant:

$$\frac{\partial g}{\partial g_{ni}} = M^{ni}. \tag{VII.17}$$

Taking into account that g^{ik} is reciprocal to g_{ik}, i.e. $g_{ik}g^{kn} = \delta_i^n$, ($g^{ik}$ is inverse matrix of g_{ik}), one can write $g^{ik} = M^{ik}/g$, i.e. $M^{ik} = gg^{ik}$. Thus

$$dg = \frac{\partial g}{\partial g_{ik}} dg_{ik} = M^{ik} dg_{ik} = gg^{ik} dg_{ik}, \tag{VII.18}$$

hence

$$g^{ik} dg_{ik} = \frac{dg}{g} = d\ln|g| = d\ln(-g) = 2\ln\sqrt{-g}. \tag{VII.19}$$

Then $g^{ik}dg_{ik} = d(g^{ik}g_{ik}) - g_{ik}dg^{ik} = d\delta_i^i - g_{ik}dg^{ik} = -g_{ik}dg^{ik}$.
Thus

$$\delta(\sqrt{-g}) = -\frac{1}{2\sqrt{-g}} gg^{ik}\delta g_{ik} = \frac{1}{2\sqrt{-g}} gg_{ik}\delta g^{ik} = -\frac{1}{2}\sqrt{-g} g_{ik}\delta g^{ik}. \tag{VII.20}$$

Now we can rewrite the variation of action as

$$\delta \int R\sqrt{-g} d\Omega = \int [(R_{ik} - \frac{1}{2} g_{ik} R)\sqrt{-g}\delta g^{ik} + g^{ik}\sqrt{-g}\delta R_{ik}] d\Omega. \tag{VII.21}$$

Let us consider now the last term in the variation. For the calculation of δR_{ik} we can use the fact that although Γ^i_{kn} is not a tensor, its variation, $\delta\Gamma^i_{kn}$, is a tensor.
Proof: Let A^i is an arbitrary vector at the point x^i. After the parallel transport From the point x^i to the point $x^i + dx^i$, as we know, its components are

$$A^i(x^n + dx^n) = A^i(x^n) + (A^i_{,m}(x^n) + \Gamma^i_{mp}(x^n) A^p(x^n)) dx^m. \tag{VII.22}$$

Then

$$\delta A^i(x^n + dx^n) = \delta\Gamma^i_{mp}(x^n) A^p(x^n)) dx^m. \tag{VII.23}$$

The left side is a vector because it is the difference between two vectors in the same point, hence the right side is also a vector. Thus $\delta\Gamma^i_{mp}(x^n)$ is a tensor.
In a locally galilean frame of reference

$$g^{ik}\delta R_{ik} = g^{ik}\left\{\delta\Gamma^l_{ik,l} - \delta\Gamma^l_{il,k}\right\} = g^{ik}\delta\Gamma^l_{ik,l} - g^{il}\delta\Gamma^k_{ik,l} = W^l_{,l}, \tag{VII.24}$$

where

$$W^l = g^{ik}\delta\Gamma^l_{ik} - g^{il}\delta\Gamma^k_{ik}, \tag{VII.25}$$

obviously W^l is a vector.
Now let us prove that the covariant divergence of an arbitrary vector can be written as follows

$$A^n_{;\,n} = \frac{1}{\sqrt{-g}} (\sqrt{-g} A^n)_{,n}. \tag{VII.26}$$

Proof:

$$A^n_{;n} = A^n_{,n} + \Gamma^n_{ni} A^i = A^n_{,n} + \frac{1}{2} g^{nm}(g_{nm,i} + g_{mi,n} - g_{in,m}) A^i = A^n_{,n} + \frac{1}{2}(g^{nm}g_{nm,i} + g^{nm}g_{mi,n} - g^{nm}g_{ni,m}) A^i =$$

$$= A^n_{,n} + \frac{1}{2} g^{nm} g_{nm,i} A^i. \tag{VII.27}$$

Taking into account (VII.18), one obtains

$$A^i_{;i} = A^n_{,n} + \frac{g_{,n}}{2g} A^n = \frac{1}{\sqrt{-g}} [\sqrt{-g} A^n_{,n} + (\sqrt{-g})_{,n} A^n] = \frac{1}{\sqrt{-g}} (\sqrt{-g} A^i)_{,i}. \tag{VII.28}$$

As follows from the proof above, in local galilean frame of reference, where $g = -1$

$$A^i_{;i} = A^i_{,i}, \tag{VII.29}$$

hence, returning back to δR_{ik}, in local galilean frame of reference we have

$$g^{ik}\delta R_{ik} = W^l_{,l} = W^l_{;l}. \tag{VII.30}$$

Since this is a relation between two tensors (of 0-rank), once this is valid in one frame of reference it is valid in an arbitrary frame of reference. Hence

$$\sqrt{-g}g^{ik}\delta R_{ik} = \sqrt{-g}W^l_{;l} = (\sqrt{-g}W^l)_{,l}, \tag{VII.31}$$

this means that according to the Gauss theorem the contribution of the third term in the variation of the action function is equal to zero.
Finally we obtain

$$\delta S_g = \alpha \int (R_{ik} - \frac{1}{2}g_{ik}R)\delta g^{ik}\sqrt{-g}d\Omega. \tag{VII.32}$$

C. The action function for matter

Similar to the action function for gravitational field, the action function for matter can be written as

$$S_m = \int \Psi\sqrt{-g}d\Omega, \tag{VII.33}$$

where Ψ is a scalar action density (by matter we mean any substance including all physical fields, for example, electromagnetic field).
Let us calculate the variation of S_m. Immediately the following problem arises. Obviously Ψ can depend on many physical parameters describing the physical system to which we are trying to apply the least action method. let us denote all of them as q_a, $a = 1, 2, 3, 4, ...$ Should we take into account the variations of all these q_a? The answer is no, all these variations should cancel each other by virtue of the "equations of motion" of the physical system under consideration, since these equations are obtained, according to the principle of the least action, from the condition that the variations of S_m, related with the variations of q_a, are equal to zero. Thus it is enough to take into account the variations of the metric tensor only. Then we have

$$\delta S_m = \int \left\{ \frac{\partial\sqrt{-g}\Psi}{\partial g^{ik}}\delta g^{ik} + \frac{\partial\sqrt{-g}\Psi}{\partial(g^{ik}_{,l})}\delta(g^{ik}_{,l}) \right\} d\Omega. \tag{VII.34}$$

Then taking into account that

$$\delta(g^{ik}_{,l}) = (\delta g^{ik})_{,l}, \tag{VII.35}$$

which means that the partial differentiation, obviously, commutates with the procedure of taking variations, we can integrate the second term in the previous formula by parts, as a result we obtain

$$\delta S_m = \int \left\{ \frac{\partial\sqrt{-g}\Psi}{\partial g^{ik}} - \frac{\partial}{\partial x^l}\frac{\partial\sqrt{-g}\Psi}{\partial(g^{ik}_{,l})} \right\} \delta g^{ik} d\Omega. \tag{VII.36}$$

Let us introduce the following notation

$$\sqrt{-g}\ A_{ik} = \frac{\partial\sqrt{-g}\Psi}{\partial g^{ik}} - \frac{\partial}{\partial x^l}\frac{\partial\sqrt{-g}\Psi}{\partial(g^{ik}_{,l})}. \tag{VII.37}$$

Then δS_m takes the following form

$$\delta S_m = \int A_{ik}\delta g^{ik}\sqrt{-g}d\Omega. \tag{VII.38}$$

D. The stress-energy tensor and the action density

One can prove that the tensor A_{ik} introduced in the previous section, is proportional to the stress-energy tensor T_{ik} introduced in the previous lecture.
Proof: Let us carry out infinitesimally small translation from the coordinates x^i to the coordinates $x'^i = x^i + \xi^i$, where ξ^i are infinitesimally small quantities. Considering this translation as a transformation of coordinates, we can see that the contravariant metric tensor is transformed under these translations as

$$g'^{ik}(x'^l) = g^{lm}(x^l)\frac{\partial x'^i}{\partial x^l}\frac{\partial x'^k}{\partial x^m} = g^{lm}(\delta^i_l + \frac{\partial \xi^i}{\partial x^l})(\delta^l_m + \frac{\partial \xi^k}{\partial x^m}) = g^{ik}(x^l) + g^{im}\xi^k_{,m} + g^{kl}\xi^i_{,l}. \tag{VII.39}$$

On other hand, using the usual Tailor expansion we have

$$g'^{ik}(x'^l) = g^{ik}(x^l + \xi^l) = g'^{ik}(x^l) + \xi^l\frac{\partial g^{ik}}{\partial x^l} = g'^{ik}(x^l) + \xi^l g^{ik}_{,l}, \tag{VII.40}$$

hence

$$g^{ik}(x^l) + g^{im}\xi^k_{,m} + g^{kl}\xi^i_{,l} = g'^{ik}(x^l) + \xi^l g^{ik}_{,l}. \tag{VII.41}$$

We obtain that

$$g'^{ik}(x^l) = g^{ik}(x^l) - \xi^l g^{ik}_{,l} + g^{il}\xi^k_{,l} + g^{kl}\xi^i_l \ \ \textbf{or} \ \ g'^{ik} = g^{ik} + \delta g^{ik}, \tag{VII.42}$$

where

$$\delta g^{ik} = -\xi^l g^{ik}_{,l} + g^{il}\xi^k_{,l} + g^{kl}\xi^i_l. \tag{VII.43}$$

It easy to show that

$$\delta g^{ik} = g^{il}\xi^k_{;l} + g^{kl}\xi^i_{;l} \equiv \xi^{i;k} + \xi^{k;i}. \tag{VII.44}$$

Indeed,

$$\delta g^{ik} = -\xi^l(g^{ik}_{;l} - \Gamma^i_{nl}g^{nk} - \Gamma^k_{nl}g^{in}) + g^{il}(\xi^k_{;l} - \Gamma^k_{ln}\xi^n) + g^{kl}(\xi^i_{;l} - \Gamma^i_{ln}\xi^n) =$$

$$= \xi^l(\Gamma^i_{nl}g^{nk} + \Gamma^k_{nl}g^{in}) + g^{il}\xi^k_{;l} + g^{kl}\xi^i_{;l} - \xi^n(\Gamma^k_{ln}g^{il} + \Gamma^i_{ln}g^{kl}) =$$

$$= \xi^l(\Gamma^i_{nl}g^{nk} + \Gamma^k_{nl}g^{in} - \Gamma^k_{nl}g^{in} - \Gamma^i_{nl}g^{kn}) + g^{il}\xi^k_{;l} + g^{kl}\xi^i_{;l} =$$

$$= g^{il}\xi^k_{;l} + g^{kl}\xi^i_{;l} \equiv \xi^{i;k} + \xi^{k;i}. \tag{VII.45}$$

Now we know what is the variation of the contravariant metric tensor under infinitesimally small translation. If we substitute this variation into Eq.(VII.38), we obtain

$$\delta S_m = \int A_{ik}(\xi^{i;\,k} + \xi^{k;\,i})\sqrt{-g}d\Omega. \tag{VII.46}$$

From the definition of A_{ik} follows that it is a symmetric tensor. From the fact that S_m is scalar follows that the variation of S_m under translation (which is the sort of transformation of coordinates) is equal to zero, hence, we obtain

$$0 = \int A_{ik}\xi^{i;\,k}\sqrt{-g}d\Omega = \int (A^k_i\xi^i)_{;\,k}\sqrt{-g}d\Omega - \int A^k_{i;\,k}\xi^i\sqrt{-g}d\Omega. \tag{VII.47}$$

The first term in the last expression can be written as

$$(A^k_i\xi^i)_{;\,k}\sqrt{-g} = \sqrt{-g}A^k_{;k}, \ \text{ where } \ A^k = A^k_i\xi^i. \tag{VII.48}$$

As follows from Eq. (VII.26)

$$\sqrt{-g}A^{k}_{;\,k} = (\sqrt{-g}A^{k})_{,k}, \tag{VII.49}$$

and gives zero contribution to the variation. As a result we obtain that

$$\int A^{k}_{i;\,k}\xi^{i}\sqrt{-g}d\Omega = 0 \tag{VII.50}$$

nd because of arbitrariness of ξ^i we conclude that

$$A^{k}_{i;\,k} = 0. \tag{VII.51}$$

Taking into account that the covariant divergence of the stress-energy tensor T^i_k (see the previous lecture) is also equal to zero, one can identify A^k_i with the physical stress energy tensor within a constant factors, β and Λ:

$$A^i_k = \beta(T^i_k + \Lambda\delta^i_k). \tag{VII.52}$$

E. The final EFEs

Finally, from the principle of least action we have

$$\delta(S_g + S_m) = 0, \tag{VII.53}$$

or

$$\int \left[\alpha\left(R_{ik} - \frac{1}{2}g_{ik}R\right) + \beta(T_{(phys)ik} + \Lambda g_{ik})\right]\delta g^{ik}\sqrt{-g}d\Omega = 0. \tag{VII.54}$$

Taking into account the arbitrariness of δ and dropping label "(phys)" and putting $\Lambda = 0$ [because discussion of this famous Λ-terms is out of the scope of this course] we obtain

$$R_{ik} - \frac{1}{2}g_{ik}R = \kappa T_{ik}, \tag{VII.55}$$

where

$$\kappa = -\frac{\beta}{\alpha}. \tag{VII.56}$$

The value of κ called the Eistein constant, can be easily obtained from the weak field and slow motion limit. As we will see later

$$\kappa = \frac{8\pi G}{c^4}. \tag{VII.57}$$

This is the end of the rigorous derivation of the EFEs.
One can see that the EFEs can be rewritten in mixed components as

$$R^i_k - \frac{1}{2}\delta^i_k R = \kappa T^i_k. \tag{VII.58}$$

Contracting indices one can obtain

$$R - \frac{1}{2}4R = \kappa T, \;\; R = -\kappa T, \;\; T = T^i_k. \tag{VII.59}$$

Hence

$$R_{ik} = \kappa(T_{ik} - \frac{1}{2}g_{ik}T). \tag{VII.60}$$

In empty space-time

$$T^i_k = 0, \quad \text{hence} \quad R^i_k = 0. \tag{VII.61}$$

However, it could happen that

$$R_{iklm} \neq 0. \tag{VII.62}$$

The tidal accelerations related with non zero components of the Riemann tensor in empty space are produced by gravitational waves. From

$$T^i_{;i} = 0 \tag{VII.63}$$

follows that

$$(R^i_k - \frac{1}{2}\delta^i_k R)_{;i} = R^i_{k;i} - \frac{1}{2}R_{,k} = 0. \tag{VII.64}$$

This is actually the case as it follows from the Bianchi identity. And vice versa, from pure geometrical Bianchi identity one can obtain the full description of motion of all forms of matter and fields. This means that the EFEs is complete and self-consistent description of the interaction between matter and geometry, i.e. gravitational field.

Back to Content **Previous Lecture** **Next Lecture**