# Variational theory in the context of general relativity

Hiroki Takeda[1,*]

[1] *Department of Physics, University of Tokyo, Bunkyo, Tokyo 113-0033, Japan*

(Dated: October 25, 2018)

## Abstract

This is a brief note to summarize basic concepts of variational theory in general relativity for myself. This note is mainly based on Hawking(2014)"Singularities and the geometry of spacetime" and Wald(1984)"General Relativity".

* takeda@granite.phys.s.u-tokyo.ac.jp

## I. BRIEF REVIEW OF GENERAL RELATIVITY AND VARIATIONAL THEORY

### A. Variation

**defnition I.1 (Action)** *Let $\mathscr{M}$ be a differentiable manifold and $\Psi$ be some tensor fields on $\mathscr{M}$. $\Psi$ are simply called the fields. Consider the function of $\Psi$, $S[\Psi]$. This $S$ is a map from the set of the fields on $\mathscr{M}$ to the set of real numbers and called the action.*

$\Psi$ actually represents some fields ${\Psi_{(i)}}^{a\cdots b}{}_{c\cdots d}(i=1,2,\cdots,n)$. I left out the indices $i$ that indicate the number of fields and the abstract indices for the tensors. The tensor of the tensor fields ${\Psi_{(i)}}^{a\cdots b}{}_{c\cdots d}(i=1,2,\cdots,n)$ at $x\in\mathscr{M}$ is written by ${\Psi_{(i)}|_x}^{a\cdots b}{}_{c\cdots d}(i=1,2,\cdots,n)$. Similarly I denote the tensor of the tensor fields $\Psi$ at $x\in\mathscr{M}$ by $\Psi|_x$.

**defnition I.2 (Variation)** *Let $\mathscr{D}\subset\mathscr{M}$ be a submanifold of $\mathscr{M}$ and $\Psi|_x(u)$, $u\in(-\epsilon,\epsilon)$, $x\in\mathscr{M}$ be the tensors of one-parameter family of the tensor fields $\Psi(u)$, $u\in(-\epsilon,\epsilon)$ at $x\in\mathscr{M}$ such that*

$$(1)\ \Psi|_x(0)=\Psi \text{ in } \mathscr{M}, \tag{1}$$

$$(2)\ \Psi|_x(u)=\Psi \text{ in } \mathscr{M}-\mathfrak{D}. \tag{2}$$

*Then I define the variation of the fields as follows,*

$$\delta\Psi := \left.\frac{d\Psi(u)}{du}\right|_{u=0}. \tag{3}$$

In this note, I suppose the existence of the derivatives $dS/du|_{u=0}$ for all one-parameter family $\Psi(u)$.

**defnition I.3 (Functional derivative)** *I assume the existence of the smooth tensor fields $\chi$ which is dual to $\psi$ such that*

$$\left.\frac{dS}{du}\right|_{u=0} = \int_{\mathfrak{D}} \chi\delta\Psi, \tag{4}$$

*for any one-parameter family.* [1] *Then we say that $S$ is functionally differentiable at $\Psi(0)$. We call $\chi$ the functional derivative of $S$ and denote it as*

$$\left.\frac{\delta S}{\delta\Psi}\right|_{\Psi_0} := \chi.\text{[2]} \tag{5}$$

---

[1] If the type of tensor field of $\Psi$ is $(k,l)$, that of $\chi$ is $(l,k)$. Here contraction of all indices in the integral is understood.

[2] More generally, if there exists a tensor distribution $\chi$ such that $dS/du|_{u=0}=\chi[\delta\Psi]$, we say that $S$ is

**defnition I.4 (Lagrangian)** *Consider a functional form of S*

$$S[\Psi] = \int_{\mathfrak{D}} L[\Psi] \tag{6}$$

*where L is a local function of $\psi$ and a finite number of its derivatives,*

$$L|_x = L(\Psi(x), \nabla\Psi(x), \cdots, \nabla^k\Psi(x)). \tag{7}$$

*L is called by Lagrangian density.*

**defnition I.5 (Variational principle)** *Let S be functionally differentiable. We require that the fields $\Psi$ maximize S,*

$$\left.\frac{dS}{du}\right|_{u=0} = 0. \tag{8}$$

*This principle is called by variational principle. Hence, the following equation is hold by the definition of the functional derivative and the following fundamental lemma of variational calculus.*

$$\left.\frac{\delta S}{\delta \Psi}\right|_{\Psi_0} = 0. \tag{9}$$

*This gives us the fundamental equations for the fields.*

**lemma I.1 (Fundamental lemma of variational calculus)** *Let $\mathscr{D} \subset \mathscr{M}$ be a submanifold of $\mathscr{M}$ and $\Psi|_x(u)$, $u \in (-\epsilon, \epsilon)$, $x \in \mathscr{M}$ be any one-parameter family satisfying a proper boundary conditions. The following is held.*

$$\left.\frac{dS}{du}\right|_{u=0} = \int_{\mathfrak{D}} \chi\delta\Psi = 0 \Rightarrow \chi = 0. \tag{10}$$

*This statement is called by fundamental lemma of variational calculus.*

### B. General Relativity

**defnition I.6 (Space-time)** *Space-time is the set of events. In general relativity, the mathematical model of space-time is ($\mathscr{M}$, g) where $\mathscr{M}$ is a connected four-dimensional Hausdorff $C^\infty$ manifold and g is a Lorentz metric [3] on $\mathscr{M}$.*

---

functionally differentiable and call $\chi$ the functional derivative of $S$ at $\Psi(0)$. Here the tensor distribution is a distribution whose type is tensor, so called generalized function in Japanese. The bracket on the right hand side denotes the scalar product (distribution) which is defined by $\chi[\delta\Psi] =< \chi, \delta\Psi >= \int_{\mathfrak{D}} \chi\delta\Psi$. (Schwarz, 2008)

[3] Its signature is $+2$

First, I derive the fundamental equations for fields by variational method. Hereafter I refer to the fields $\Psi$ as $\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d}$, $(i=1,2,\cdots,n)$ without omitting the indices. I assume the form of the action is Lagrangian form and the Lagrangian depends on the fields and their first covariant derivatives.

Consider the variation of the fields $\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d}$, $(i=1,2,\cdots,n)$. Then

$$\left.\frac{dS}{du}\right|_{u=0} = \sum_i \int_{\mathfrak{D}} \left( \frac{\partial L}{\partial \Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d}} \delta\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d} - \frac{\partial L}{\partial \Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d;e}} \delta(\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d;e}) \right). \tag{11}$$

Since$\delta(\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d;e}) = (\delta\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d})_{;e}$, the second term can be written as

$$\sum_i \int_{\mathfrak{D}} \left[ \left( \frac{\partial L}{\partial \Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d;e}} \delta\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d} \right)_{;e} - \left( \frac{\partial L}{\partial \Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d;e}} \right)_{;e} \delta\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d} \right]. \tag{12}$$

According to Stokes' theorem, the first term can be written as

$$\int_{\mathfrak{D}} Q^a{}_{;a} = \int_{\partial\mathfrak{D}} Q^a d\sigma_a, \tag{13}$$

where

$$Q^e := \sum_i \frac{\partial L}{\partial \Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d;e}} \delta\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d}. \tag{14}$$

This integral have to vanish because $\delta\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d} = 0$ is held at the boundary $\partial\mathfrak{D}$ from the second condition of variation Eq. (2). Finally we get the following expression,

$$\left.\frac{dS}{du}\right|_{u=0} = \sum_i \int_{\mathfrak{D}} \left( \frac{\partial L}{\partial \Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d}} - \left( \frac{\partial L}{\partial \Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d;e}} \right)_{;e} \right) \delta\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d} = 0. \tag{15}$$

Thus, we can get the following Euler-Lagrange equations from fundamental lemma of variational calculus Eq. (10),

$$\chi = \frac{\partial L}{\partial \Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d}} - \left( \frac{\partial L}{\partial \Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d;e}} \right)_{;e} = 0. \tag{16}$$

Then we get the energy-momentum tensor from the Lagrangian by considering a variation with respect to the metric. Suppose a variation of the metric $g_{ab}(u,x)$ does not change the fields $\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d}$. Then,

$$\left.\frac{dS}{du}\right|_{u=0} = \int_{\mathfrak{D}} \left( \sum_i \frac{\partial L}{\partial \Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d;e}} \delta(\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d;e}) + \frac{\partial L}{\partial g_{ab}} \delta g_{ab} \right) + \int_{\mathfrak{D}} L \frac{\partial \boldsymbol{\epsilon}}{\partial g_{ab}} \delta g_{ab}. \tag{17}$$

The last term comes from the fact that the integral is defined by $\int f = \int f\boldsymbol{\epsilon}$. Since $\boldsymbol{\epsilon} = (4!)^{-1}\boldsymbol{\eta}$, $\eta_{abcd} = (-g)^{1/2}4!\delta_{[a}{}^1\delta_b{}^2\delta_c{}^3\delta_{d]}{}^4$ and $g := \det(g_{ab})$,

$$\frac{\partial \eta_{abcd}}{\partial g_{ef}} = -\frac{1}{2}(-g)^{-1/2}\frac{\partial g}{\partial g_{ef}}4!\delta_{[a}{}^1\delta_b{}^2\delta_c{}^3\delta_{d]}{}^4. \tag{18}$$

Thus, we obtain

$$\frac{\partial \boldsymbol{\epsilon}}{\partial g_{ab}} = \frac{1}{2}g^{ab}\boldsymbol{\epsilon}. \tag{19}$$

Here I use the the fact that the cofactor of the matrix can be expressed as a product of the component of the inverse matrix and the determinant od the matrix, that is the relation $\partial g/\partial g_{ef} = g^{ef}g$ to evaluate $\partial g/\partial g_{ef}$. The first term arises because $\delta(\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d;e})$ does not necessarily become zero unlike $\delta\Psi_{(i)}{}^{a\cdots b}{}_{c\cdots d}$ due to a variation of the components of the connection $\Gamma^a{}_{bc}$ induced by a variation of the metric. $\partial\Gamma^a{}_{bc}/\partial u = \delta\Gamma^a{}_{bc}$ should be a tensor because the difference between two connections should be a tensor. This can be calculated as follows,

$$\delta\Gamma^a{}_{bc} = \frac{1}{2}g^{ad}\{(\delta g_{db})_{;c} + (\delta g_{dc})_{;b} - (\delta g_{bc})_{;d}\}. \tag{20}$$

One can verify this equation by checking that this equation is held in normal coordinates at $p$ where the components $\Gamma^a{}_{bc}$ and the coordinate derivatives of the components $g_{ab}$ vanish because this equation is a tensor equation. For Riemann normal coordinate,

$$(l.h.s) = \frac{1}{2}g^{ad}\left(\frac{\partial\delta g_{cd}}{\partial x^b} + \frac{\partial\delta g_{bd}}{\partial x^c} - \frac{\partial\delta g_{bc}}{\partial x^d}\right), \tag{21}$$

and, since the components $\Gamma^a{}_{bc}$ at $p$ induced by the covariant derivatives vanish,

$$(r.h.s) = \frac{1}{2}g^{ad}\{(\delta g_{db})_{;c} + (\delta g_{dc})_{;b} - (\delta g_{bc})_{;d}\} = (l.h.s.). \tag{22}$$

If the tensor equation is held in a specified coordinate, it is held in any coordinates. This method is often used to derive equations in relativity. However, One should be careful to calculations including the variation and the covariant derivative. Note that $\delta(A+B) = \delta A + \delta B$, $\delta(AB) = (\delta A)B + A(\delta B)$ and $\delta(\partial A) = \partial(\delta A)$. Especially, the variation operator does not commute the covariant derivative for any tensors $g$ (here its type is $(0,2)$ as an example),

$$\nabla_a(\delta g_{bc}) = \delta(\nabla_a g_{bc}) + g_{dc}\,\delta\Gamma^d_{ba} + g_{bd}\,\delta\Gamma^d_{ac}, \tag{23}$$

when one consider the variation of the metric because a change of the connection is induced. This can be easily derived by the above three equations. On the other hand, the variation operator does commute the covariant derivative when you consider the variation of the fields except for the metric as I mentioned above. Eq. (19) can be also derived by this equation straightforward. One also should be careful to that the since the metricity condition holds only for the unperturbed metric $g_{ab}$, $\nabla(g(0)) = 0$ does not mean that $\nabla(g(\Delta u)) = 0$. Using to leave only terms multiplied by $\delta g_{ab}$, The first term in Eq. (17) can be expressed by the terms involving only $\delta\Gamma$ by using the similar relation as Eq. (23). Then, using Eq. (19) and the integration by parts, it finally can be expressed by the only terms involving only $\delta g$. Thus, the energy-momentum tensor $T^{ab}$ can be defined as follows,

$$\left.\frac{dS}{du}\right|_{u=0} = \frac{1}{2}\int_{\mathfrak{D}} (T^{ab}\delta g_{ab}). \tag{24}$$

The energy-momentum tensor satisfies the conservation equations. Let $\phi$ be a diffeomorphism which is is the identity except in the interior of $\mathscr{D}$. Then,

$$S = \int_{\mathfrak{D}} L = \frac{1}{4!}\int_{\mathfrak{D}} L\boldsymbol{\eta} = \frac{1}{4!}\int_{\phi(\mathfrak{D})} L\boldsymbol{\eta} = \frac{1}{4!}\int_{\mathfrak{D}} \phi^*(L\boldsymbol{\eta}). \tag{25}$$

Here I used the condition that $\phi$ is the identity except in the interior of $\mathscr{D}$ for the third equality, and the invariance of integrals under a differential map for the forth equality. Thus,

$$\frac{1}{4!}\int_{\mathfrak{D}} (L\boldsymbol{\eta} - \phi^*(L\boldsymbol{\eta})) = 0. \tag{26}$$

If $\phi$ is generated by a vector field $\boldsymbol{X}$ which is non-zero only in the interior of $\mathfrak{D}$,

$$\frac{1}{4!}\int_{\mathfrak{D}} L_{\boldsymbol{X}}(L\boldsymbol{\eta}) = 0, \tag{27}$$

by the definition of the Lie derivatives. On the other hand,

$$\begin{aligned}(l.h.s.) &= \sum_i \int_{\mathfrak{D}} \left( \frac{\partial L}{\partial {\Psi_{(i)}}^{a\cdots b}{}_{c\cdots d}} - \left( \frac{\partial L}{\partial {\Psi_{(i)}}^{a\cdots b}{}_{c\cdots d;e}} \right)_{;e} \right) L_{\boldsymbol{X}} {\Psi_{(i)}}^{a\cdots b}{}_{c\cdots d} \\ &\quad + \frac{1}{2}\int_{\mathfrak{D}} T^{ab} L_{\boldsymbol{X}} g_{ab} \\ &= \frac{1}{2}\int_{\mathfrak{D}} T^{ab} L_{\boldsymbol{X}} g_{ab} \\ &= \int_{\mathfrak{D}} ((T^{ab}X_a)_{;b} - {T^{ab}}_{;b} X_a).\end{aligned} \tag{28}$$

I used the field equation for the second equality and the relation $L_{\boldsymbol{X}} g_{ab} = 2X_{(a;b)}$ for the third equality. The first term can be an integral over $\partial\mathfrak{D}$ by Stokes' theorem. It can vanish because $\boldsymbol{X}$ vanishes there. Thus,

$$\int_{\mathfrak{D}} T^{ab}{}_{;b} X_a = 0, \tag{29}$$

is satisfied for all vector fields $\boldsymbol{X}$. Hence, the conservation law of the energy-momentum tensor is derived,

$$T^{ab}{}_{;b} = 0. \tag{30}$$

**ACKNOWLEDGEMENTS**

Thank you, everyone.