Reading Notes

General Relativity by Robert Wald

Jinyan Miao

Winter 2023

Contents

1 | The Spacetime Model

Chapter 1 - 3 on Wald's *General Relativity*

These chapters on the book introduce the basic mathematical model used in general relativity. We start with the intuition of the spacetime metric, then we explore Lorentzian manifolds, curvatures, and geodesics in a manner similar to Riemannian geometry.

Modeling Spacetime using Manifolds and Metric Tensors

The word *spacetime* in relativity theory refers to a continuum composed of *events*. Each event is thought of as a point in *space* at an instant of *time*. Consider living in a 3-dimensional space and a 1-dimensional time domain, an event in spacetime can be uniquely represented using four inputs, three for the spatial position and one for the time-tag of the event.

Prior to relativity theory, it was believed that spacetime had the following structure: Given an event p in spacetime, there is a natural, observer independent notion of events occurring at the same time as p. More precisely, given two distinct events p and q, one of the following mutually exclusive possibilities must hold:

1. It is in principle possible for an observer or material body to go from event q to event p, in which case q is said to be the past of p.
2. It is in principle possible for an observer or material body to go from event p to event q, in which case q is said to be the future of p.

It is impossible for an observer or material body to be present at both events p and q, hence the notion of *simultaneity* associated to event p is a 3-dimensional set.

In general relativity, the spacetime causal structure described above turns out to be wrong, the crucial difference is that the notion of simultaneity of a given event forms more than a 3-dimensional set. The causal relation between event p and other events has the lightcone structure sketched in Figure 1.1.

Remark: One might guess that the $\mathbb{R}^4$ Euclidean space is a suitable model for spacetime, but right now we do not know any *global* property of spacetime, thus we shall not assume that spacetime has all the Euclidean structures.

In special relativity, neither the time interval nor the space interval between relatively simultaneous events, that are events determined to be simultaneous by a particular
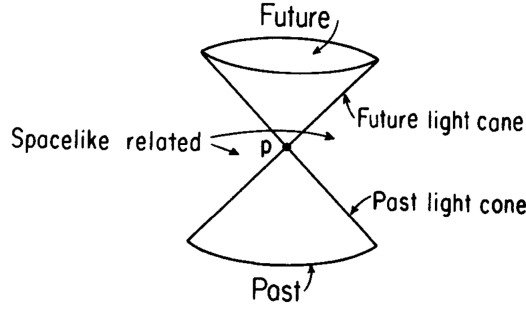
Figure 1.1: Lightcone that shows that causal structure of spacetime in special relativity, the horizontal axes characterize spatial axes, and the vertical axis characterizes the time axis. Events that lie on the cone boundary cannot be reached by a material particle starting at p but can be reached by a light signal emitted from p. The upper cone forms the future lightcone of p, and the lower cone forms the past lightcone of p.

observer, has absolute significance. The quantity which is observer independent is the *spacetime interval*, I, defined by

$$I = -(\Delta t)^2 + \frac{1}{c^2}\left((\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2\right) \tag{1.1}$$

where Δt is the time interval, and Δx, Δy, Δz describe the space interval. The space-time interval I and functions of I are the only observer independent quantities characterizing the spacetime relationships between events. We shall later refer to I as the metric of spacetime in in analogy to an ordinary Euclidean metric.

Spacetime should be a 4-dimensional object as mentioned above, and intuitively, space-time should locally look like Euclidean space, thus mathematically we would like to model spacetime by a 4-dimensional manifold. In this text, an n-dimensional manifold M is equipped with a set of charts $\psi : O_\alpha \to U_\alpha$, where O_α is an open subset of M and U_α is an open subset of $\mathbb{R}^n$, precisely given by the following definition:

Definition 1.0.1
An n-dimensional real manifold M of C^∞ type is a set equipped with a collection of subsets $\mathscr{O} = \{O_\alpha \subseteq M \,|\, \alpha \in \mathscr{J}\}$ satisfying the following properties:[1]
1. *For all $p \in M$, there exists at least one $O_\alpha \in \mathscr{O}$ such that $p \in O_\alpha$.*
2. *For $\alpha \in \mathscr{J}$, there exists a bijection $\psi_\alpha : O_\alpha \to U_\alpha$ where $U_\alpha \subseteq \mathbb{R}_n$ is open.*
3. *If $O_\alpha, O_\beta \in \mathscr{O}$ overlap, then the map $\psi_\beta \circ \psi_\alpha^{-1} : \psi_\alpha(O_\alpha \cap O_\beta) \to \psi_\beta(O_\alpha \cap O_\beta)$ is of C^∞ type, and the sets $\psi_\alpha(O_\alpha \cap O_\beta)$, $\psi_\beta(O_\alpha \cap O_\beta)$ are open in $\mathbb{R}^n$.*

Here the set $\mathscr{O}$ is called the cover of M. The cover $\mathscr{O}$ and the chart family $\{\psi_\alpha \,|\, \alpha \in \mathscr{J}\}$ are assumed to be maximal.

The manifolds are further considered to be Hausdorff and paracompact. In a sense similar to what we have done in Math 635, the set V_p of tangent vectors at a point $p \in M$ can be defined using the one-to-one correspondence between vectors and directional derivatives of the inverse of the charts ϕ.

---

[1] If not specified otherwise, $\mathscr{J}$ in this text denotes an index set.

1.1.1   Vector Spaces on Manifolds

The tangent vectors on the manifold are viewed as the *infinitesimal displacements* about a point on the manifold. In many situations, one is given a manifold without an embedding of it in $\mathbb{R}^n$, thus it is important to define a tangent vector in a way that refers only to the intrinsic structure of the manifold, not to its possible embeddings in $\mathbb{R}^n$. A vector $v = (v^1, v^2, \cdots, v^n) \in \mathbb{R}^n$ defines the directional derivative operator $\sum_\mu v^\mu (\partial/\partial x^\mu)$ and vice versa. Directional derivatives are characterized by the linearity and Leibnitz rule behavior when acting on functions.

Definition 1.0.2
Let M be an n-dimensional real manifold, and let $\mathscr{F}$ denote the collection of C^∞ functions from M to $\mathbb{R}$. We define a tangent vector v at $p \in M$ to be a function $v : \mathscr{F} \to \mathbb{R}$ which is linear and obeys the Leibnitz rule, as characterized by

1. $v(af + bg) = av(f) + bv(g)$ for $f, g \in \mathscr{F}$ and $a, b \in \mathbb{R}$,
2. $v(fg) = f(p)v(g) + g(p)v(f)$ for $f, g \in \mathscr{F}$.

Note: From Definition 1.0.2, we see immediately that if $h \in \mathscr{F}$ is a constant function denoted by $h(q) = c$ for any $q \in M$, then $v(h) = 0$ for any tangent vector v. As we have $v(h \cdot h) = 2cv(h)$ whereas $v(h \cdot h) = v(ch) = cv(h)$.

Theorem 1.1 (Theorem 2.2.1 on Wald's *General Relativity*)
Let M be an n-dimensional manifold, and let $p \in M$, $V_p := T_pM$ denote the tangent space at p, then T_pM has dimension n.

Proof sketch. Let $\mathscr{F}$ denote the collection of C^∞ functions from M to $\mathbb{R}$. The key of the proof is to show that the function $e_\mu : \mathscr{F} \to \mathbb{R}$ defined by the following defines a basis vector for the vector space V_p:

$$e_\mu(f) = \left. \frac{\partial}{\partial x^\mu}(f \circ \psi^{-1}) \right|_{\psi(p)}$$

where ψ is a chart that contains p in its domain of definition, and $f \in \mathscr{F}$. If one uses a different chart ψ', we would obtain a different coordinate basis $\{e'_\mu \mid \mu \in \mathbb{N}_n\}$ with similar definition, using the chain rule, we can write the following:

$$e_\mu = \sum_{\nu=1}^n \left. \frac{\partial (x')^\nu}{\partial x^\mu} \right|_{\psi(p)} e'_\mu$$

where $(x')^\nu$ denotes the ν-th component of the map $\psi' \circ \psi^{-1}$. Hence the components $(v')^\nu$ of a vector v in the coordinate basis $\{e'_\mu \mid \mu \in \mathbb{N}_n\}$ are related to the components v^μ in the coordinate basis $\{e_\mu \mid \mu \in \mathbb{N}_n\}$ by the vector transformation law:

$$(v')^\nu = \sum_\mu^n v^\mu \frac{\partial (x')^\nu}{\partial x^\mu} \tag{1.2}$$

$\square$

Remark: The basis $\{e_\mu \mid \mu \in \mathbb{N}_n\}$ of V_p introduced in Theorem 1.1 is called a *coordinate basis* and is the natural basis equipped on V_p. [2]

---

[2] For notations on Wald's *General Relativity*, e_μ is denoted as X_μ, or more generally, $\partial/\partial x^\mu$.

Definition 1.1.1
A smooth curve on a manifold M is a C^∞ function $c : U \to M$, where U is a connected subset of $\mathbb{R}$. For $f \in C^\infty(M, \mathbb{R})$, let $T(f)$ be the derivative of $f \circ c : \mathbb{R} \to \mathbb{R}$, that is $T(f) = (d/dt)(f \circ c)$, here $T(f)$ is called the tangent vector field of c, and at each $p \in M$, $T(f)|_p$ is called the tangent vector of c at p.

Note: From Definition 1.1.1, for $f \in C^\infty(M, \mathbb{R})$, we can write:

$$T(f) = \frac{d}{dt}(f \circ c) = \sum_\mu \frac{\partial}{\partial x^\mu}(f \circ \psi^{-1})\frac{dx^\mu}{dt} = \sum_\mu \frac{dx^\mu}{dt}e_\mu(f)$$

where e_μ is defined similarly as in Theorem 1.1. Thus in any coordinate basis, the components of T is given by dx^μ/dt.

Definition 1.1.2
A vector field V on a manifold M is a function which associates each point $p \in M$ a vector $v \in T_pM$. A vector field v is smooth provided that $v(f)$ is smooth for all $f \in C^\infty(M, \mathbb{R})$. The commutator $[v, w]$ of two smooth vector fields V, W on M is a vector field defined by $[V, W](f) = V(W(f)) - W(V(f))$.

1.1.2   Tensors on Manifolds

Given the notion of tangent vectors, the use of tensors arises when one wants to compute other quantities. It turns out that many quantities have a linear dependence on displacements.

Definition 1.1.3
Given a tangent space $V = V_p$ of a manifold M at point $p \in M$, a tensor T of type (k, l) over V is a multilinear map of the form given by

$$T : \underbrace{V^* \times V^* \times \cdots V^*}_{k \; times} \times \underbrace{V \times V \times \cdots \times V}_{l \; times} \to \mathbb{R}$$

An assignment of tensor over V_p for each $p \in M$ is called a tensor field on M. The collection of (k, l) tensor is a $n^{(k+l)}$-dimensional vector space denoted as $\mathscr{T}(k, l)$. [3]

Remark: As such, a $(0, 1)$ tensor over a vector field V is a dual vector in V^*, and a $(1, 0)$ tensor over V is a dual of dual vector in V^{**}.

Definition 1.1.4
The contraction of a (k, l) tensor with respect to its i-th dual-vector and j-th vector components is a map $C : \mathscr{T}(k, l) \to \mathscr{T}(k - 1, l - 1)$ defined by:

$$CT = \sum_{\sigma=1}^n T(\cdots, v^{\sigma^*}, \cdots, v_\sigma, \cdots)$$

where $\{v_\sigma \mid \sigma \in \mathbb{N}_n\}$ and $\{v^{\sigma^*} \mid \sigma \in \mathbb{N}_n\}$ are the basis of V and V^*, respectively, and the v^{σ^*} and v_σ are inserted into the i-th dual-vector and j-th vector component of T.

---

[3]Here and thereafter V^* denotes the dual space of the vector space V. V^{**} denotes the dual space of V^*, and is equipped with the canonical isomorphism $M : V \to V^{**}$ which sends $v \in V$ to a map in V^{**} whose value on $w^* \in V^*$ reads $w^*(v)$.

Definition 1.1.5
The outer product of a (k, l) tensor T and a (k', l') tensor T' over a vector space V is a $(k + k', l + l')$ tensor denoted as $T \otimes T'$ defined by:

$$(T \otimes T')(v^{1^*}, v^{2^*} \cdots, v^{k^*}, v^{(k+1)^*}, \cdots, v^{(k+k')^*}, w_1, w_2 \cdots, w_l, w_{l+1}, \cdots, w_{l+l'})$$
$$= T(v^{1^*}, v^{2^*} \cdots, v^{k^*}, w_1, w_2 \cdots, w_l) \cdot T'(v^{(k+1)^*}, \cdots, v^{(k+k')^*}, w_{l+1}, \cdots, w_{l+l'})$$

where v^{n^*} are dual vectors in V^* and w_n are vectors in V.

Lemma 1.1.1
Every tensor T of type (k, l) over an n-dimensional vector space V can be expressed as a sum of simple tensors:

$$T = \sum_{\substack{\mu_1, \mu_2, \cdots \mu_k, \in \mathbb{N}_n \\ \nu_1, \nu_2, \cdots, \nu_l \in \mathbb{N}_n}} (T^{\mu_1 \mu_2 \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \nu_l}) v_{\mu_1} \otimes v_{\mu_2} \otimes \cdots \otimes v_{\mu_k} \otimes v^{\nu_1^*} \otimes v^{\nu_2^*} \otimes \cdots \otimes v^{\nu_l^*}$$

where $\{v_{\mu_i} \mid \mu_i \in \mathbb{N}_n\}$ is a basis of V and $\{v^{\nu_i^*} \mid \nu_i \in \mathbb{N}_n\}$ is the corresponding dual basis of V^*, $T^{\mu_1 \mu_2 \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \nu_l} \in \mathbb{R}$ is called the components of the tensor T with respect to the given basis of V. Here we have taken the convention that, v_{μ_i} and $v^{\nu_i^*}$ satisfy the followings:

$$v^{\mu_j^*} v_{\mu_i} = \delta^{\mu_j}{}_{\mu_i} \qquad and \qquad v^{\nu_i^*} v_{\nu_j} = \delta^{\nu_i}{}_{\nu_j}$$

The contraction of T with respect to its i-th dual-vector and j-th vector components as in Definition 1.1.4 then has components given by:

$$(CT)^{\mu_1 \mu_2 \cdots \mu_{k-1}}{}_{\nu_1 \nu_2 \cdots \nu_{l-1}} = \sum_{\sigma=1}^{n} T^{\mu_1 \mu_2 \cdots \sigma \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \sigma \cdots \nu_l}$$

with σ being inserted into the i-th dual-vector and j-th vector components. The tensor product of a (k, l) tensor T and a (k', l') tensor T' has components given by:

$$(T \otimes T')^{\mu_1 \mu_2 \cdots \mu_{k+k'}}{}_{\nu_1 \nu_2 \cdots \nu_{l+l'}} = T^{\mu_1 \mu_2 \cdots \mu_k}{}_{\nu_1 \nu_2 \cdots \nu_l} \cdot (T')^{\mu_{k+1} \cdots \mu_{k-k'}}{}_{\nu_{l+1} \cdots \nu_{l-l'}}$$

For notations, given an n-dimensional vector space V with it basis $\{\partial/\partial x^i \mid i \in \mathbb{N}_n\}$, the associated dual basis of V^* is denoted as $\{dx^i \mid i \in \mathbb{N}\}$.

Note: Let V be an n-dimensional vector space with its basis $\{v_\nu \mid \nu \in \mathbb{N}_n\}$, using Eq. (1.2), via the Inverse Function Theorem, and by requiring the basis vectors v^{μ^*} for V^* to satisfy $v^{\mu^*} v_\nu = \delta^\mu{}_\nu$, we obtain the *dual vector transformation law* for the components w_μ of a vector w:

$$w'_{\mu'} = \sum_{\mu=1}^{n} w_\mu \frac{\partial x^\mu}{\partial (x')^{\mu'}} \tag{1.3}$$

where x^ν denotes the ν-th component of the chart transition function $\psi \circ (\psi')^{-1}$. Combining one obtains the *tensor transformation law* for components of a (k, l) tensor T:

$$(T')^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} = \sum_{\mu_i, \nu_i \in \mathbb{N}_n} (T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l}) \frac{\partial (x')^{\mu'_1}}{\partial x^{\mu_1}} \cdots \frac{\partial (x')^{\mu'_k}}{\partial x^{\mu_k}} \frac{\partial x^{\nu_1}}{\partial x^{\nu'_1}} \cdots \frac{\partial x^{\nu_l}}{\partial x^{\nu'_l}} \tag{1.4}$$

Note: In the rest of the text, a tensor T of type (k, l) is denoted as $T^{a_1 a_2 \cdots a_k}{}_{b_1 b_2 \cdots b_l}$ or simply $T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l}$, with lower-case-letters subscripts and superscripts. The components of T are denoted using greek-letters subscripts and superscripts $T^{\alpha_1 \alpha_2 \cdots \alpha_k}{}_{\beta_1 \beta_2 \cdots \beta_l}$. Vectors v in the domain of definition of T are denoted using superscripts v^{b_i} such that $T^{a_1 \cdots a_k}{}_{b_1 \cdots b_i \cdots b_l} v^{b_i}$ is a $(k, l-1)$ tensor, that is the i-th vector component of T is evaluated at v^{b_i}. A dual vector w, on the other hand, has vector space as a natural domain of definition, hence is viewed as a $(0, 1)$ tensor, denoted using subscripts w_{a_i}.

1.1.3   The Metric Tensor

Definition 1.1.6
In the theory of general relativity, a metric g on a manifold M is a symmetric non-degenerate tensor field of type $(0, 2)$ defined on M. Components of g are usually denoted as $g_{\mu\nu}$, and g has the form:

$$g = \sum_{\mu,\nu} g_{\mu\nu} dx^\mu \otimes dx^\nu$$

Conventionally, one might employ the compact notation for g:

$$ds^2 = \sum_{\mu,\nu} g_{\mu\nu} dx^\mu dx^\nu$$

where the tensor outer product $\otimes$ notation is omitted.

Remark: The metric can be viewed as an inner product on the tangent space at each point in M. However, from Definition 1.1.6, a metric tensor in the theory of general relativity is not necessarily positive definite.

Theorem 1.2
Given a metric g on a manifold M, one can find an orthonormal basis $\{v_1, v_2, \cdots, v_n\}$ of the tangent space at each p in the spacetime manifold M, that is a basis that makes $g(v_i, v_j) = 0$ if $i \neq j$ and $g(v_i, v_j) = \pm 1$ for $i = j$. It is not hard to check that the number of basis vectors v_i corresponding to $g(v_i, v_i) = +1$ and the number with $g(v_i, v_i) = -1$ are independent of the choice of such orthonormal basis. The number of the $+$ and $-$ signs occuring is called the signiture of the metric. The metric equipped on a spacetime manifold M usually has a signiture $(-, +, +, +)$, as in the next few examples.

Remark: A positive definite metric, that is a metric with all-plus signature is called a Riemannian metric. A metric with a signature like those on spacetime, one minus and the remainder plus, is called a Lorentzian metric.

Definition 1.2.1
A manifold equipped with a Lorentzian metric is called a Lorentzian manifold.

Example: Probably the simplest example of a Lorentzian manifold is the flat spacetime equipped with the Minkowski metric

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 = \eta_{\mu\nu} dx^\mu dx^\nu \,. \tag{1.5}$$

where (t, x, y, z) is a coordinate in the spacetime manifold M.

In matrix form, the Minkowski metric can be represented by

$$\eta = \begin{bmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Here c is a constant, which is interpreted as the speed of light. The flat spacetime is in fact used more often in the theory of special relativity. In special relativity, an event is defined as a single moment in space and time, characterized uniquely by a coordinate (t, x, y, z). One can then introduce the spacetime interval between two events,

$$I = -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 = \eta(z_1 - z_2) = \eta_{\mu\nu}\Delta x^\mu \Delta x^\nu \,,$$

which is a quantity that is invariant for all observers in any inertial frames. [4]

Example: Another example of a Lorentzian manifold is a spacetime equipped with the Schwarzschild metric given by

$$ds^2 = -\left(1 - \frac{2Gm}{rc^2}\right) c^2 \, dt^2 + \left(1 - \frac{2Gm}{rc^2}\right)^{-1} dr^2 + r^2 d\Omega^2 \,, \tag{1.6}$$

where $d\Omega^2 = d\theta^2 + \sin^2(\theta)\, d\phi^2$ is the standard metric on the 2-sphere, G is the gravitation constant, m is a constant with the dimensions of mass. Note that, except at the origin, the Schwarzschild metric approaches the spherical form of the Minkowski metric as m approaches 0, or as r approaches infinity. Note here (1.6) is an exact solution to Einstein's field equations that describes the gravitational field outside a spherical mass, which we shall discuss in next Chapter. Eq. (1.6) is helpful for describing slowly rotating astronomical objects, especially the curved spacetime around a black hole singularity.

---

[4]More details about special relativity in a context without differentiable geometry can be found in Jinyan's notes for the class Physics 360 - Honors Physics, taught by Henriette Elvang at the University of Michigan in Fall 2021.

Derivative Operators on a Manifold

Definition 2.0.1
A derivative operator, or the covariant derivative operator, on a manifold M is an operator ∇_a that associates any (k,l) tensor with a $(k,l+1)$ tensor, and ∇_a itself satisfies the following properties:

1. *Linearity. For $A, B \in \mathcal{T}(k,l)$, and $\alpha, \beta \in \mathbb{R}$, we have:*

$$\nabla_a(\alpha A^{c_1\cdots c_k}{}_{b_1\cdots b_l} + \beta B^{c_1\cdots c_k}{}_{b_1\cdots b_l}) = \alpha \nabla_a A^{c_1\cdots c_k}{}_{b_1\cdots b_l} + \beta \nabla_a B^{c_1\cdots c_k}{}_{b_1\cdots b_l}$$

2. *Sanctification of the Leibnitz Rule. For $A \in \mathcal{T}(k,l)$ and $B \in \mathcal{T}(k',l')$, we have:*

$$\nabla_a \left(A^{c_1\cdots c_k}{}_{b_1\cdots b_l} B^{d_1\cdots d'_k}{}_{e_1\cdots e'_l} \right)$$
$$= \left(\nabla_a A^{c_1\cdots c_k}{}_{b_1\cdots b_l} \right) B^{d_1\cdots d'_k}{}_{e_1\cdots e'_l} + A^{c_1\cdots c_k}{}_{b_1\cdots b_l} \left(\nabla_a B^{d_1\cdots d'_k}{}_{e_1\cdots e'_l} \right)$$

3. *Commutativity with contraction. For all $A \in \mathcal{T}(k,l)$, we have:*

$$\nabla_a(A^{c_1 c_2\cdots d\cdots c_k}{}_{b_1 b_2\cdots d\cdots b_l}) = \nabla_a A^{c_1 c_2\cdots d\cdots c_k}{}_{b_1 b_2\cdots d\cdots b_l}$$

4. *Consistency with the notion of tangent vectors as directional derivatives. For all $f \in C^\infty(M, \mathbb{R})$ and $t = t^a \in V_p$ for some $p \in M$, we have:*

$$t(f) = t^a \nabla_a f$$

5. *Torsion free. For all $f \in C^\infty(M, \mathbb{R})$, we have:*

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f$$

Remark: As in Definition 2.0.1, even though we denote a derivative operator on a manifold M as ∇_a, the operator ∇_a itself is not a tensor, nor a dual vector.

Note: It is not hard to verify that such a derivative operator ∇_a defined on a manifold M exists. Given a chart ψ on M, in the region $U \subseteq M$ covered by ψ we can define a derivative operator ∂_a, called the ordinary derivative, by letting the tensor $\partial_a T^{b_1\cdots b_k}{}_{c_1\cdots c_l}$ have components given by $\partial(T^{\mu_1\cdots \mu_k}{}_{\nu_1\cdots \nu_l})/\partial x^\sigma$ in the given coordinate basis. Note further that the ordinary derivative operator is coordinate dependent.

Lemma 2.0.1
Let ∇_a and $\widetilde{\nabla}_a$ be derivative operators on a manifold M, for $T \in \mathcal{T}(k,l)$, we have:

$$\nabla_a T^{b_1\cdots b_k}{}_{c_1\cdots c_l}$$
$$= \widetilde{\nabla}_a T^{b_1\cdots b_k}{}_{c_1\cdots c_l} + \sum_i C^{b_i}{}_{ad} T^{b_1\cdots d\cdots b_k}{}_{c_1}\cdots c_l - \sum_j C^d{}_{ac_j} T^{b_1\cdots b_k}{}_{c_1\cdots d\cdots c_l} \qquad (1.8)$$

where $C^a{}_{bc}$ is an $(1,2)$ tensor field on M. Note here the i-th dual-vector component of $T^{b_1\cdots d\cdots b_k}{}_{c_1\cdots c_l}$ is labeled as d instead of b_i, and the j-th component of $T^{b_1\cdots b_k}{}_{c_1\cdots d\cdots c_l}$ is labeled as d instead of c_j.

Lemma 2.0.1 suggests a relation between two derivative operators on a manifold M. For a vector field t^a on M, Eq. (1.8) reduces to the following:

$$\nabla_a t^b = \widetilde{\nabla}_a t^b + C^b{}_{ac} t^c \tag{1.9}$$

where $C^b{}_{ac}$ is a (1,2) tensor field on M. In particular, in the case where $\widetilde{\nabla}_a$ in Eq. (1.9) is taken to be the ordinary derivative operator ∂_a associated to a given chart, the tensor field $C^b{}_{ac}$ is denoted as $\Gamma^b{}_{ac}$ instead, and $\Gamma^b{}_{ac}$ is called the Christoffel symbol, satisfying the following:

$$\nabla_a t^b = \partial_a t^b + \Gamma^b{}_{ac} t^c \tag{1.10}$$

Lemma 2.0.2
Let ∇_a be a derivative operator on a manifold M, let v^a and w^b be vector fields on M, then we have $[v, w]^b := [v, b]$ being a vector field on M satisfying the following:

$$[v, w]^b = v^a \nabla_a w^b - w^a \nabla_a v^b$$

Definition 2.0.2
Given a derivative operator ∇_a on a manifold M and a curve $c : I \to M$ on M, where I is a connected subset of $\mathbb{R}$, a vector v^a given at each point p on the curve c is said to be parallelly transported along c whose tangent is t^a provided that we have $t^a \nabla_a v^b = 0$ holds along the curve. A tensor $T^{b_1 \cdots b_k}{}_{c_1 \cdots c_k}$ is said to be parallelly transported along c provided that we have $t^a \nabla_a T^{b_1 \cdots b_k}{}_{c_1 \cdots c_k}$ holds along the curve.

Note: By specifying a coordinate system, $t^a \nabla_a v^b = 0$ as in Definition 2.0.2 has a form:

$$t^a \partial_a v^b + t^a \Gamma^b{}_{ac} v^c = 0$$

in terms of components, we have:

$$\frac{dv^\nu}{dt} + \sum_{\mu, \lambda} t^\mu \Gamma^\nu{}_{\mu\lambda} v^\lambda = 0$$

The definition of parallel transport allows us to *connect* tangent vector spaces at different points on a manifold, hence the derivative operator ∇_a is sometimes called the *connection* on the manifold. Practically, one would like to have a derivative operator that is compatible with the metric defined on a manifold M. That is, when vectors v^b and w^c are parallelly transported along curves on the manifold, $g_{bc} v^b w^c$ should be invariant, that is $t^a \nabla_a (g_{bc} v^b w^c) = 0$, which leads to $\nabla_a g_{bc} = 0$.

Theorem 2.1 (Theorem 3.1.1 on Wald's *General Relativity*)
Let g_{bc} be a metric on a manifold M, there exists a unique derivative operator ∇_a on M that is compatible with g_{ab}, that is $\nabla_a g_{bc} = 0$ everywhere on M. In which case the tensor field C in Eq. (1.8) satisfies the following for any given $\widetilde{\nabla}_a$:[5]

$$C^c{}_{ab} = \frac{1}{2} g^{cd} \left(\widetilde{\nabla}_a g_{bd} + \widetilde{\nabla}_b g_{ad} - \widetilde{\nabla}_d g_{ad} \right) \tag{1.11}$$

---

[5]A torsion free derivative operator ∇_a that is compatible with the given metric on the manifold is called the Levi-Civita Connection on M.

Corollary 2.1.1

From Theorem 2.1, in terms of a given ordinary derivative operator, the Christoffel symbol in Eq. (1.10) that makes ∇_a being compatible with the metric on M has the following form:

$$\Gamma^c_{\ ab} = \frac{1}{2} g^{cd} \left(\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab} \right)$$

The components of the Christoffel symbol is given by the following:

$$\Gamma^\sigma_{\ \mu\nu} = \frac{1}{2} \sum_\sigma g^{\rho\sigma} \left(\frac{\partial g_{\nu\sigma}}{\partial x^\mu} + \frac{\partial g_{\mu\sigma}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\sigma} \right)$$

Curvatures and Geodesics

Let M be a manifold and let $p, q \in M$, with tangent spaces V_p and V_q. In general, parallelly transporting a vector from V_p to V_q depends on the path connecting p and q. We can use the path dependence of parallel transport to define an intrinsic notion of curvature.

Let ∇_a be a derivative operator on a manifold M, and let w_c be a dual vector field on M, it is not hard to show that $\nabla_a \nabla_b w_c - \nabla_b \nabla_a w_c$ can be viewed as a $(1, 3)$ tensor acting on w_d, such that the following (Eq. (3.2.3) on Wald's *General Relativity*) holds:

$$\nabla_a \nabla_b w_c - \nabla_b \nabla_a w_c = R_{abc}{}^d w_d$$

where the operator $R_{abc}{}^d$ is called the Riemannian curvature tensor. One also finds that the action of $R_{abc}{}^d$ on a vector field t^c gives the following:

$$\nabla_a \nabla_b t^c - \nabla_b \nabla_a t^c = -R_{abd}{}^c t^d$$

from which induction gives:

$$(\nabla_a \nabla_b - \nabla_b \nabla_a) T^{c_1 \cdots c_k}{}_{d_1 \cdots d_l} = -\sum_{i=1}^{k} R_{abe}{}^{c_i} T^{c_1 \cdots e \cdots c_k}{}_{d_1 \cdots d_l} + \sum_{j=1}^{l} R_{abd_j}{}^{e} T^{c_1 \cdots c_k}{}_{d_1 \cdots e \cdots d_l}$$

Lemma 3.0.1
The Riemannian curvature tensor $R_{abc}{}^d$ on a manifold satisfies the followings: [6] [7]

1. $R_{abc}{}^d = -R_{bac}{}^d$
2. $R_{[abc]}{}^d = 0$
3. $R_{abcd} = -R_{abdc}$ *for the ∇_a that satisfies $\nabla_a g_{bc} = 0$*
4. *The Bianchi identity holds, that is $\nabla_{[a} R_{bc]d}{}^e = 0$*

From Lemma 3.0.1, we see here the trace of $R_{abc}{}^d$ over its first two or last two indices vanishes, while its trace over the second and fourth, or the first and third, indices defines the Ricci curvature tensor R_{ac}, given by:

$$R_{ac} := R_{abc}{}^b$$

where we have an obvious property $R_{ac} = R_{ca}$. The scalar curvature R is defined by the following:

$$R := R_a{}^a$$

Bianchi identity leads to an important equation satisfied by R_{ab}, that is we have:

$$\nabla_a R_{bcd}{}^a + \nabla_b R_{cd} - \nabla_c R_{bd} = 0 \tag{1.12}$$

Defining the tensor G_{ab} by the following:

$$G_{ab} := R_{ab} - \frac{1}{2} R \cdot g_{ab}$$

---

[6] $R_{[abc]}{}^d = (R_{abc}{}^d + R_{bca}{}^d + R_{cab}{}^d - R_{bac}{}^d - R_{acb}{}^d - R_{cba}{}^d)/6$

[7] $\nabla_{[a} R_{bc]d}{}^e = (\nabla_a R_{bcd}{}^e + \nabla_b R_{cad}{}^e + \nabla_c R_{abd}{}^e - \nabla_b R_{acd}{}^e - \nabla_a R_{cbd}{}^e - \nabla_c R_{bad}{}^e)/6$

we obtain the following from (1.12):

$$\nabla_a R_c{}^a + \nabla_b R_c{}^d - \nabla_c R = 0 \qquad \Rightarrow \qquad \nabla^a G_{ab} = 0$$

where the tensor G_{ab} is called the Einstein tensor.

Definition 3.0.1
Let ∇_a be a derivative operator on a manifold M, a geodesic on M is a curve $c : I \to M$, where I is a connected subset of $\mathbb{R}$, whose tangent vector t^a satisfies $t^a \nabla_a t^b = 0$ along c.

Given a derivative operator ∇_a on a manifold M, one can in fact reparametrize any curve whose tangent vector t^a satisfies $t^a \nabla_a t^b = \alpha t^b$ for some $\alpha \in C^\infty(M, \mathbb{R})$ to obtain a geodesic on M. Hence we only consider curves which are affinely parametrized, that is $t^a \nabla_a t^b = 0$. Written in coordinate basis components, the geodesic equation $t^a \nabla_a t^b = 0$ reads the following:

$$\frac{d^2 x^\mu}{dt^2} + \sum_{\sigma, \nu} \Gamma^\mu{}_{\sigma\nu} \frac{dx^\sigma}{dt} \frac{dx^\nu}{dt} = 0\,. \tag{1.13}$$

which is a coupled system of n second order ODE for n functions x^μ. Given initial value of x^μ and dx^μ/dt, unique solution to (1.13) always exists by the Cauchy-Lipschitz Theorem.

Theorem 3.1
Let M be a manifold equipped with the Levi-Civita connection ∇_a. Given $p \in M$ and and tangent vector $t^a \in V_p$, there exists a unique geodesic through p with tangent vector t^a at p.

Note: From the notion of existence and uniqueness of geodesic at a given point and given velocity on the manifold, one can define the exponential map, normal coordinates, and polar coordinates locally on the manifold.

Note: For computations of curvatures in coordinate components, one makes use of Section 3.4 - *Methods for Computing Curvature* on Wald's *General Relativity*.

1.3.1   The Lengths of the Geodesics in Lorentzian Manifolds

Definition 3.1.1
Denoting the tangent vector of a geodesic $c : I \to M$ on a Lorentzian manifold M as T^a, c is said to be timelike provided that $g_{ab}T^a T^b < 0$ everywhere, c is said to be null provided that $g_{ab}T^a T^b = 0$ everywhere, and C is said to be spacelike if $g_{ab}T^a T^b > 0$ everywhere. The length l of spacelike c is defined by:

$$l = \int_{t \in I} (g_{ab}T^a T^b)^{1/2}\, dt\,. \tag{1.14}$$

The length of the null geodesic is 0. The proper time of timelike c is defined by:

$$\tau = \int_{t \in I} (-g_{ab}T^a T^b)^{1/2}\, dt\,. \tag{1.15}$$

Since the tangent vector of a geodesic is parallel transported, it has a constant norm, which ensures that geodesics in the Lorentzian manifold cannot change from timelike to spacelike or null.

Note: Length of a curve is invariant under reparametrization.

Remark: The geodesics in Riemannian manifolds are paths that extremize the length functional. A similar argument can be employed to show that the paths in Lorentzian manifolds extremize the length or the proper time. That is, curves that extremize proper time between two points are precisely the timelike geodesics, and the curves that extremize the length between two points are precisely the spacelike geodesic. In general relativity, *gravity* is treated as a built-in feature of the spacetime structure, reflected by the metric g_{ab}. A test particle in spacetime is said to be *free-falling* provided that the particle does not experience any non-gravitational force. As a result, timelike geodesics are the trajectories of massive free-falling test particles, and the proper time of the timelike geodesic is a quantity describing the time measured by a clock moving along the geodesic. Null geodesics are the trajectories of the massless free-falling test photons, and it has a proper time 0. Two independent events in spacetime are spacelike related.

Theorem 3.2
Given a Lorentzian manifold with its compatible derivative operator ∇_a, the curves which extremize propertime between two points are precisely the timelike geodesics, and the curves which extremize length between two points are precisely the spacelike or null geodesics.

Note: The proof of Theorem 3.2 is a direct application of variation calculation.

In Riemannian manifolds, one can find curves of arbitrarily long length connecting two points, while the length between two points are bounded below. In Lorentzian manifolds, in contrast, one can find timelike curves of arbitrarily small proper time connecting points. In some spacetime, the proper time of timelike curves connecting the two given points need not be bounded above, but if a curve of greatest proper time exists, it must be a timelike geodesic.

1.3.2 Deviation of Geodesics

Let $\mathcal{G} = \{\gamma_s \mid s \in \mathbb{R}\}$ denote a smooth one-parameter family of geodesics parametrized by the affine parameter t. Consider the map $F : U \to M \quad (t, s) \mapsto \gamma_s(t)$ with $U \subseteq \mathbb{R}$, which is smooth, injective, and has a local smooth inverse. Let Y denote the 2-dimensional submanifold of M spanned by the curves $\mathcal{G}$. Here Y has natural coordinates (s, t). The vector field $T^a := (\partial/\partial t)^a$ is tangent to the family of geodesics, and hence we have $T^a \nabla_a T^b = 0$, whereas the vector field $X^a = (\partial/\partial s)^a$ represents the displacement to an infinitesimally nearby geodesics, and here X^a is called the vector field of *deviation vector*. Given the Levi-Civita connection ∇_a on the manifold of interest, by re-scaling parameter t by an s-dependent factor, one can ensure that $g_{ab}T^a T^b$ does not vary with s. Here X^a and T^a are coordinate vector fields, hence they commute, and thus is not difficult to show that $X^a T_a$ is constant along each geodesic (pp. 43, 46 on Wald's *General Relativity*). Furthermore, by reparametrizing each γ_s by adding

a constant dependent on s to the affine parameter t, we may ensure that the curve $C : \mathbb{R} \to M \quad s \mapsto \gamma_s(0)$ is orthogonal to all the geodesics. Thus we have $X_a T^a = 0$ along C, and hence $X^a T_a = 0$ everywhere as $X^a T_a$ is constant along each geodesic.

With the constructions above, the quantity $v^a = T^b \nabla_b X^a$ gives the rate of change along a geodesic of the displacement to an infinitesimally nearby geodesic, so v^a may be interpreted as the relative velocity of an infinitesimally nearby geodesic. Similarly, we may interpret $a^d = T^c \nabla_c v^d = T^c \nabla_c (T^b \nabla_b X^d)$ as the relative acceleration of an infinitesimally nearby geodesic in the family $\mathscr{G}$. Simply derivation (Eq. (3.3.18) on Wald's *General Relativity*) shows that we have the following:

$$a^d = -R_{cbe}{}^d X^b T^c T^e \tag{1.16}$$

Eq. (1.16) is referred as the *geodesic deviation equation*.

Note: One has $a^d = 0$ for all families of geodesics if and only if $R_{abc}{}^d = 0$. Thus some geodesics will accelerator toward or away from each other if and only if the Riemannian curvature tensor of the manifold does not vanish everywhere.

2 | The Theory of Relativity

Prerelativity and Motivations

Before the theory of relativity, it is assumed that the space manifold M has the Euclidean structure . Each neighborhood of a point $p \in M$ can be covered by a chart ϕ whose inverse $\phi^{-1} : U \to M$ can be viewed as a composition of rotation and translation of the subset U of $\mathbb{R}^3$, the so-obtained coordinate system is called the Cartesian coordinate system. The metric equipped on M, written in the Cartesian coordinate system, has the following form:

$$h_{ab} = \sum_{\mu,\nu} h_{\mu\nu} (dx^\mu)_a (dx^\nu)_b \tag{2.1}$$

with $h_{\mu\nu} = 1$ when $\mu = \nu$, and $h_{\mu\nu} = 0$ when $\mu \neq \nu$. Note that the definition of h_{ab} is independent of the choice of Cartesian coordinate system used to cover a neighborhood in M. Furthermore, as each $h_{\mu\nu}$ is constant, it is observed that we have $\partial_a h_{bc} = 0$, and consequently $\Gamma_{bc}{}^a$ vanishes for the given Cartesian coordinate system. Thus, the curvature of M vanishes, in which case M equipped with h_{ab} is said to be *flat*. From the geodesic equations (1.13), we deduce that the geodesics on M are the *straight lines* in Cartesian coordinates, in other words, straight lines in $\mathbb{R}^3$. The distance between two points on M, in accordance with Eq. (1.14), has the simple form given by the following:

$$l = \left((\Delta x^1)^2 + (\Delta x^2)^2 + (\Delta x^3)^2 \right)^{1/2}$$

where $\Delta x^i = x^i_{\text{final}} - x^i_{\text{initial}}$ is the i-th component displacement in the Cartesian coordinate system. Thus, we can conclude that, in prerelativity physics, the space can be viewed as the manifold $\mathbb{R}^3$ equipped with a flat Riemannian metric.

Quantities of physical interest in space must eventually be reducible to numbers as all experiments in physics measure numbers. However, many quantities of interest, such as magnetic field or the stress tensor, require the additional specification of a basis of vectors in order to produce numbers. Thus a general class of quantities of interests are maps of vectors and dual vectors into numbers. Any analytic map can be Taylor-expanded as a sum of multilinear maps, hence tensor fields encompass an extremely wide class of quantities. Indeed, the notion of tensor fields appears to be great enough so that essentially all quantities which one considers in physics can be viewed as tensor fields, and the laws of physics governing these quantities can be expressed as tensor equations.

2.4.1 Principles of Covariance

The following ideas, principles of covariance, played important roles in motivating the formulation of the theory general relativity and special relativity.

The Principle of General Covariance
The metric of space is the only quantity pertaining to space that can appear in the laws of physics. The form of the laws of physics should be the same in all inertial and accelerating frames.

More specifically, the principle of general covariance states that there are no preferred vector fields or preferred bases of vector fields pertaining only to the structure of space which appear in any law of physics. The implication of the principle of general covariance is that a Christoffel symbol cannot appear by itself in any law of physics, as it is equivalent to specifying an ordinary derivative which is an additional geometric quantity pertaining to the space and is not derivable from the metric. Furthermore, the Christoffel symbol cannot appear in the law of physics as it coordinate components do not transform according to tensor transformation under general coordinate transformation.

The Principle of Special Covariance
Any physically possible set of measurements obtained by a family $\mathscr{O}$ of observers stationed in space also is physically possible set of measurements for another family $\mathscr{O}'$ of observers obtained by acting on $\mathscr{O}$ with an isometry ϕ.

The principle of special covariance implies the existence of an action of the isometry group on the states of physical fields being measured. One can use this group action to derive, or at least motivate, the possible laws governing the physical fields.

Suppose an object of physical interest in described by a tensor field T. If general covariance holds, then the equations governing T should involve only T, the metric of space h_{ab}, and the quantities determined by h_{ab}, such as its derivative operator. Furthermore, all quantities measurable by $\mathscr{O}$ should be expressible as scalars resulting from contracting T and its derivatives with the basis vector fields associated with $\mathscr{O}$. Let ϕ be an isometry, if special covariance holds, the action of ϕ on T with produce a tensor field $\phi^* T$ which will satisfy the equations for the physical field and will yield the same measurable quantities for $\mathscr{O}'$ as T yields for $\mathscr{O}$.

Remark: Special covariance of the laws of physics for tensor fields implies that if we pick a coordinate system and write out the coordinate component equations without explicitly incorporating the metric into the equation but rather substituting everywhere its coordinate components $h_{\mu\nu}$, then the form of the equations will be preserved under the coordinate transformations corresponding to the isometries, for the simple reason that the numerical values of the components $h_{\mu\nu}$ remain unchanged under these transformations.

Historically, the formulations of special and general covariance played a large role in discussions of relativity theory, and indeed, the name of special and general relativity derive from these formulations of special and general covariance.

Special Relativity

In the theory of special relativity, it is assumed that the spacetime has the manifold structure of $\mathbb{R}^4$. It is assumed that there exist preferred families of motion in spacetime, referred to as *inertial* or *nonaccelerating* motions.

The inertial observers is further assumed can set up a rigid grid of metersticks, synchronize clocks placed at the gridpoints, [1] and label each event in spacetime by the grid coordinates (x^1, x^2, x^3) and the clock reading t at an event. The map of spacetime into $\mathbb{R}^4$ is called a *global inertial coordinate system*. Indeed, many different global inertial coordinate systems are possible, hence the coordinate components of events do not have intrinsic meaning. While the spacetime interval between two events $\bar{x}$ and x, as defined by Eq. (1.1), or in units where $c = 1$ given by the following has the same value for all global inertial coordinate systems and thus can be viewed as representing an intrinsic property of spacetime:

$$I = -(t - \bar{t})^2 + (x^1 - \bar{x}^1)^2 + (x^2 - \bar{x}^2)^2 + (x^3 - \bar{x}^3)^2 \qquad (2.2)$$

Definition 5.0.1
In special relativity, the metric of spacetime is defined by the following:

$$\eta_{ab} := \sum_{\mu,\nu=0}^{3} \eta_{\mu\nu} (dx^\mu)_a (dx^\nu)_b$$

with $\eta_{\mu\nu} = diag(-1, 1, 1, 1)$, where x^μ denotes any global inertial coordinate system.

From Definition 5.0.1, we observe that the ordinary derivative operator of the global inertial coordinates satisfies $\partial_a \eta_{bc} = 0$, and thus is the Levi-Civita connection on the spacetime equipped with η_{ab}. The curvature of spacetime equipped with η_{ab} thus vanishes everywhere. The geodesics in spacetime equipped with η_{ab} are the straight lines when expressed in global inertial coordinates. In particular, the timelike geodesics are precisely the world lines of inertial observers in spacetime.
Similar to the way that Maxwell's equations relate the distribution of charges and currents in space to the electromagnetic field, Einstein's field equations (EFE) allow us to correlate the distribution of matter in spacetime to the geometry of spacetime. Once we obtain the geometry of spacetime, the geodesics equation (3) can be used to compute the paths of free-falling particles in spacetime.

The laws of physics in general relativity are governed by two principles: (1) the principle of general covariance, that is the laws of physics should be the same in all inertial and accelerating frames; (2) we require that all equations must reduce to the equations that hold in the theory of special relativity in the case where $g_{\mu\nu}$ is flat. As we have mentioned, we will model the spacetime as a manifold M equipped with a Lorentzian metric $g_{\mu\nu}$. The derivation of the EFE is based on modifying the equations that hold in special relativity under a flat metric, the full derivation is given in Ref. [**?**] (pp. 68-72), here we will just state the result,

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \kappa T_{\mu\nu} \,, \qquad (2.3)$$

---

[1]By clock synchronization, the inertial observer makes sure two given clocks give the same reading when they receive a signal sent out in a symmetrical manner by an observer stationed halfway between the two.

where $g_{\mu\nu}$ is the Lorentzian metric tensor, $R_{\mu\nu}$ is the Ricci curvature tensor, R is the scalar curvature, κ is the Einstein gravitation constant, and $T_{\mu\nu}$ is the stress-energy-momentum tensor. Another cosmological constant term is usually added to (9) to incorporate vacuum energy in space into our models of gravity and cosmology, in which case (9) becomes

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu}\,, \tag{2.4}$$

where Λ is a cosmological constant. We note that the LHS of (10) is a pure description of the geometry of the spacetime, reflected by the curvatures and the metric tensor, while the RHS of (10) provides a description of the stress-energy-momentum content of spacetime. Eq. (3) and (10) together form the basics of the mathematical formulation of general relativity.

Remark: The LHS of (10) is completely determined by the metric tensor $g_{\mu\nu}$ as the Ricci curvature tensor and the scalar curvature are both determined by the metric. The EFE is a set of nonlinear partial differential equations. The *solution* of the EFE is usually referred to as the metric that satisfies (10) for some given $T_{\mu\nu}$.

2.5.1   Conservation of Stress-energy-momentum

The stress-energy-momentum tensor $T_{\mu\nu}$ in (10) describes the energy and momentum carried by fields as well as particles. For an observer with 4-velocity v^μ, the component $T_{\mu\nu}v^\mu v^\nu$ is interpreted as the mass-energy density as measured by the observer. If x^μ is orthogonal to v^μ, the component $T_{\mu\nu}v^\mu x^\nu$ is interpreted as the momentum density of the matter in the x^μ direction. If y^μ is also orthogonal to v^μ, then $T_{\mu\nu}x^\mu y^\nu$ represents the $x^\mu - y^\nu$ component of the material stress tensor.

One key feature that is related to (10) is the local conservation of non-gravitation stress-energy-momentum. In other words, non-gravitational energy and momentum are conserved, mathematically, the conservation law is expressed as

$$\nabla_\beta T^{\alpha\beta} = T^{\alpha\beta}_{;\beta} = 0\,.$$

Proof. By choosing a normal coordinate, the contracted second Bianchi identity reads

$$R^\gamma_{\beta\gamma\delta;\epsilon} + R^\gamma_{\beta\epsilon\gamma;\delta} + R^\gamma_{\beta\delta\epsilon;\gamma} = 0\,.$$

Rewriting using antisymmetric of the Riemannian curvature tensor, we have

$$R^\gamma_{\beta\gamma\delta;\epsilon} - R^\gamma_{\beta\gamma\epsilon;\delta} + R^\gamma_{\beta\delta\epsilon;\gamma} = 0 \qquad \Rightarrow \qquad R_{\beta\delta;\epsilon} - R_{\beta\epsilon;\delta} + R^\gamma_{\beta\delta\epsilon;\gamma} = 0\,.$$

Then we contract again using the metric tensor and obtain

$$g^{\beta\delta}(R_{\beta\delta;\epsilon} - R_{\beta\epsilon;\delta} + R^\gamma_{\beta\delta\epsilon;\gamma}) = R^\delta_{\delta;\epsilon} - R^\delta_{\epsilon;\delta} + R^{\gamma\delta}_{\delta\epsilon;\gamma} = R_{;\epsilon} - 2R^\gamma_{\epsilon;\gamma} = 0\,.$$

Rewriting, and contracting with $g^{\epsilon\delta}$, we have

$$\left(R^\gamma_\epsilon - \frac{1}{2}g^\gamma_\epsilon R\right)_{;\gamma} = 0 \qquad \Rightarrow \qquad \left(R^{\gamma\delta} - \frac{1}{2}g^{\gamma\delta}R\right)_{;\gamma} = 0\,.$$

Relabeling and comparing with (9), we obtain $\nabla_\beta T^{\alpha\beta} = T^{\alpha\beta}_{;\beta} = 0$. $\square$

The Vacuum Field Solutions

Most of our Universe appears to be empty space, interspersed with clumps of matter such as stars and planets. Ignoring negligible contributions from hydrogen gas, cosmic rays, the electromagnetic fields, and so on, the stress-energy-momentum tensor, and hence the RHS of (10), vanishes in the empty regions in spaces, but the gravitational effects described by g need not necessarily vanish, as otherwise the Earth would not orbit the Sun. Hence we are interested in the Vacuum solutions of the EFE.

To solve the EFE in vacuum (with $T_{\mu\nu} = 0$), it is helpful to rewrite (10) as the contracted form

$$R - g^{\mu\nu}g_{\mu\nu}\left(\frac{1}{2}R + \Lambda\right) = \kappa T.$$

where $T = g^{\mu\nu}T_{\mu\nu}$. Note that one has $g^{\mu}\nu g_{\mu\nu} = \delta^{\nu}_{\nu} = 4$, hence we have:

$$R = -4\Lambda + -\kappa T.$$

Combining with (10), we obtain

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\left(-4\Lambda - 2\kappa T\right) - \Lambda g_{\mu\nu} = \kappa T_{\mu\nu},$$

that is, we have obtained the vacuum field equations by setting $T_{\mu\nu} = 0$ (hence $T = 0$),

$$R_{\mu\nu} + \Lambda g_{\mu\nu} = 0. \tag{2.5}$$

For the case of zero cosmological constant $\Lambda = 0$, (11) reduces to $R_{\mu\nu} = 0$. Note that the Riemannian curvature tensor need not vanish in order for (11) to hold, so spacetime can still be curved in vacuum regions, allowing nontrivial gravitational effects.

It is obvious that the Minkowski metric introduced in the previous section is flat and hence it satisfies (11). In fact, the Schwarzschild metric introduced in the previous section is also a solution to (11) in the limit of vanishing cosmological constant, and such a solution is non-trivial.

2.6.1 The Schwarzschild Metric

For the derivation of the Schwarzschild metric, we first note that static spherically symmetric metrics have the standard form

$$g = \begin{bmatrix} c^2 A(r) & 0 & 0 & 0 \\ 0 & -B(r) & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2(\theta) \end{bmatrix}$$

with respect to the spherical coordinates (t, r, θ, ϕ) for spacetime manifold M, and here A and B are some functions of r. Computing the Ricci curvature of g, we have

$$R_{rr} = \Lambda B, \qquad R_{tt} = -c^2\Lambda A, \qquad R_{\theta\theta} = \Lambda r^2, \qquad R_{\phi\phi} = \Lambda r^2\sin^2(\theta),$$

and $R_{\mu\nu} = 0$ for $\mu \neq \nu$. Utilizing (11), by setting $\kappa = 8\pi G/c^4$, we obtain

$$A(r) = 1 - \frac{2Gm}{c^2 r} - \frac{1}{3}\Lambda r^2 \,, \qquad\qquad B(r) = \frac{1}{A(r)} \,,$$

where an integration constant $2Gm/c^2$ has been imposed. Combining we see that

$$ds^2 = -\left(1 - \frac{2Gm}{rc^2} - \frac{1}{3}\Lambda r^2\right) c^2 \, dt^2 + \left(1 - \frac{2Gm}{rc^2} - \frac{1}{3}\Lambda r^2\right)^{-1} dr^2 + r^2 d\Omega^2 \,, \quad (2.6)$$

which recovers the Schwarzschild metric in the limit of $\Lambda = 0$.

In general, (12) describes the curvature of spacetime in a region outside a static spherical body, such as a nonrotating star, black hole, or ball of dust. As a note Λ is approximately given by $10^{-52}\, m^2$, and that $A(r)$ consequently vanishes at two values

$$r_* \approx \frac{2GM}{c^2} \,, \qquad\qquad r_\infty \approx \left(\frac{3}{\Lambda}\right)^{1/2} \,, \qquad\qquad (2.7)$$

corresponding to a Schwawrzchild-type event horizon and a cosmological event horizon, respectively. This is of special notice as the radial speed of light $c|A(r)|$ vanishes at each horizon characterized by (13), thus an intermediate observer cannot see light crossing from either horizon.

In Neotonian gravity, the general equation of motion for a particle in static gravitational potential $\Phi(\vec{x}) = -Gm/|\vec{x} - \vec{X}|$ is given by

$$\vec{a} = \frac{d^2\vec{x}}{dt^2} = -\nabla\Phi \,.$$

The choice of Φ depends on the motion of a large mass m at position $\vec{X}$. If the geodesic equation (3) correctly describes motion under gravity, then we expect it to reduce to Newtonian gravity in the limit of slowly moving particles and a weakly curved spacetime, that is in the limit of $|d\vec{x}/dt| \ll c$, and $g = \eta + h$ approaching the flat metric η in suitable coordinates. In such approximation, one can obtain [?]

$$\frac{d^2\vec{x}}{dt^2} \approx -\frac{1}{2}c^2 \nabla h_{00} \,.$$

That is we have, in the limit of h and Φ vanishing at infinity,

$$\Phi \approx \frac{1}{2}c^2 h_{00} \approx \frac{1}{2}c^2 \, g_{00} - 1 \,.$$

Then (12) gives the following relation:

$$\Phi = \frac{c^2}{2}(g_{00} - 1) = \frac{c^2}{2}(A - 1) = -\frac{Gm}{r} - \frac{1}{6}\Lambda c^2 r^2 \,.$$

for regions with $\Phi/c^2 \ll 1$. Therefore, in the Newtonian viewpoint, a positive cosmological constant acts as a repulsive force that increases linearly with distance.

Bibliography

[1] R. M. Wald, *General Relativity* (Univ. of Chicago Press, Chicago, 1984).

[2] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (John Wiley and Sons, Cambridge, 1972).

[3] M. J. Hall, *General Relativity: An Introduction to Black Holes, Gravitational Waves, and Cosmology* (Morgan & Claypool Publishers, San Rafael, CA, 2018).