General Relativity

Prof. E. Castorina, a.a. 2024-25

Leonardo Cerasi[1]

GitHub repository: LeonardoCerasi/notes

[1] leo.cerasi@pm.me

Contents

Introduction

Nello studio delle interazioni a distanza si introducono le cosiddette teorie di campo: un campo è un'entità fisica che esiste in ogni punto dello spaziotempo (es: campo elettrico, magnetico, etc...) e che viene modificata dalla presenza di portatori della carica associata al campo.
Nel caso del di una teoria di campo per descrivere la gravità, è necessario un campo gravitazionale che sia influenzato dalla massa. Nel caso Newtoniano il campo gravitazionale $\Phi(\mathbf{r}, t)$ è legato alla densità di massa $\rho(\mathbf{r}, t)$ da un'equazione di Poisson:

$$\triangle\Phi = 4\pi G\rho \tag{1}$$

dove $G \approx 6.67 \cdot 10^{-11}\,\mathrm{m^3kg^{-1}s^{-2}}$ è la costante universale di Newton.
È banale ricavare il campo gravitazionale di una massa puntiforme M: in questo caso $\rho(\mathbf{r}) = M\delta^3(\mathbf{r})$, dunque $\Phi(\mathbf{r}) = -\frac{GM}{r}$. Il caso in cui invece ρ dipende dal tempo è non banale e per essere trattato necessita di un'equazione più generale di Eq. 1: l'equazione di campo di Einstein.

Analogie e differenze con l'Elettromagnetismo Superficialmente, il problema della generalizzazione relativistica della gravitazione potrebbe sembrare analogo a quello dell'elettromagnetismo: entrambe le forze, nel caso stazionario, sono governate da una legge proporzionale a r^{-2} ed entrambi i campi sono determinati da equazioni di Poisson in cui cambia solo la costante dimensionale.
La differenza tra le due teorie di campo, però, sta proprio nella descrizione matematica delle sorgenti che subentrano nelle equazioni di Poisson: nel caso dell'Elettromagnetismo, in regime stazionario il campo elettromagnetico è determinato dalla densità di carica ρ_e e dalla densità di corrente $\mathbf{J}$, e per avere una descrizione relativistica bisogna combinarle in una densità di corrente quadrivettoriale $j^\mu = (c\rho_e, \mathbf{J})$ (si può vedere che ρ_e trasforma come una componente temporale poiché $\rho_e \sim \mathrm{Vol}_3^{-1} \sim (\mathrm{Vol}_4/ct)^{-1} \sim ct$, dato che il quadrivolume è un invariante di Lorentz): ciò risulta naturalmente in un potenziale quadrivettoriale $A^\mu = (\phi/c, \mathbf{A})$.
D'altra parte, per quanto riguarda la gravitazione, bisogna ricordare l'uguaglianza relativistica tra massa energia; inoltre, a differenza della carica elettrica, l'energia non è un invariante relativistico, ma è la componente temporale del quadrivettore impulso: in particolare, a generare il campo gravitazionale sono la densità di energia ρ e la densità di momento ρ_p, alle quali sono associate una densità di corrente di energia $\mathbf{j}$ ed una densità di corrente di momento $\mathbf{T}^i$ per ciascuna componente ρ_p^i. Risulta evidente che l'equazione relativistica che descrive il campo gravitazionale sia notevolmente più complicata di quella del campo elettromagnetico, poiché le sorgenti non sono descritte da un quadrivettore, bensì da un tensore, il tensore energia-impulso:

$$T^{\mu\nu} \sim \begin{bmatrix} \rho c & \rho_p c \\ \mathbf{j} & \mathrm{T} \end{bmatrix} \tag{2}$$

Naturalmente, anche il potenziale gravitazionale sarà un tensore $h_{\mu\nu}$, ed il potenziale Newtoniano sarà $h_{00} \sim \Phi$.

Scala della Relatività Generale Tramite le costanti fondamentali G e c è possibile associare ad una massa M una sua lunghezza caratteristica, detta raggio di Schwarzschild:

$$R_s := \frac{2GM}{c^2} \tag{3}$$

Le correzioni relativistiche alla teoria della gravitazione sono determinate dal parametro R_s/r e, nella maggior parte delle situazioni, sono trascurabili: basti calcolare che per la Terra $R_s \approx 10^{-2}$ m, mentre il suo raggio è $R_T = 6 \cdot 10^6$ m, dunque sulla superficie terrestre le correzioni relativistiche alla gravità Newtoniana sono dell'ordine di 10^{-8}.
Gli effetti relativistici diventano importanti quando si considerano oggetti compatti come stelle di neutroni e buchi neri.

Part I

Il Principio d'Equivalenza

Chapter 1

Geodetiche

Nelle teorie classiche di campo vengono considerati due oggetti distinti: le particelle e i campi. I campi determinano il moto delle particelle, mentre le particelle determinano le oscillazioni dei campi.

1.1 Particelle non-relativistiche

Per descrivere il moto di una particella tra due punti fissati $\mathbf{x}(t_1) \equiv \mathbf{x}_1$ e $\mathbf{x}(t_2) \equiv \mathbf{x}_2$ si studia l'azione S associata alla traiettoria $\mathbf{x}(t)$, definita come:

$$S\left[\mathbf{x}(t)\right] := \int_{t_1}^{t_2} dt\, L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) \tag{1.1}$$

dove L è la lagragiana che descrive la particella.
La traiettoria percorsa dalla particella, per il principio di minima azione, è quella che estremizza S, ovvero tale per cui $\delta S = 0\ \forall \delta\mathbf{x}(t) : \delta\mathbf{x}(t_1) = \delta\mathbf{x}(t_2) = 0$; esplicitando:

$$\begin{aligned} \delta S = \int_{t_1}^{t_2} dt\, \delta L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) &= \int_{t_1}^{t_2} dt\, \left(\frac{\partial L}{\partial x^i}\delta x^i + \frac{\partial L}{\partial \dot{x}^i}\delta \dot{x}^i\right) \\ &= \int_{t_1}^{t_2} dt\, \left(\frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\right)\delta x^i + \left[\frac{\partial L}{\partial \dot{x}^i}\delta x^i\right]_{t_1}^{t_2} \end{aligned} \tag{1.2}$$

dove si è usata la convenzione di somma di Einstein.
Si vede subito che il termine di bordo è nullo, dunque estremizzare l'azione equivale alle equazioni di Eulero-Lagrange:

$$\frac{\partial L}{\partial x^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i} = 0 \tag{1.3}$$

1.1.1 Equazione geodetica

In generale, il moto di una particella libera su una generica varietà differenziale è descritto dalla lagrangiana $L = \frac{1}{2}m\dot{\mathbf{x}}\cdot\dot{\mathbf{x}}$; bisogna dunque tener conto della metrica della varietà considerata:

$$L = \frac{m}{2}g_{ij}(x)\dot{x}^i\dot{x}^j \tag{1.4}$$

dove x rappresenta collettivamente tutte le coordinate x^i sulla varietà. Si ricordi che, per una varietà reale n-dimensionale, $g_{ij} \in \mathbb{R}^{n\times n}$ è una matrice reale simmetrica.

Le equazioni di Eulero-Lagrange diventano dunque:

$$\frac{m}{2}\frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j - \frac{d}{dt}\left(mg_{ik}\dot{x}^i\right) = 0 \tag{1.5}$$

Espandendo il secondo termine:

$$\frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j - \frac{\partial g_{ik}}{\partial x^j}\dot{x}^i\dot{x}^j - g_{ik}\ddot{x}^i = 0 \tag{1.6}$$

Del termine $g_{ik,j} - \frac{1}{2}g_{ij,k}$, essendo contratto con un fattore simmetrico $\dot{x}^i\dot{x}^j$, sopravvive solo la parte simmetrica rispetto agli indici i e j, ovvero:

$$g_{ik}\ddot{x}^i + \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k}\right)\dot{x}^i\dot{x}^j = 0 \tag{1.7}$$

A questo punto, si contrae per la metrica inversa g^{lk}, che per definizione soddisfa $g^{lk}g_{ik} = \delta^l_i$, così da ottenere (rinominando gli indici):

$$\ddot{x}^i + \Gamma^i_{jk}\dot{x}^j\dot{x}^k = 0 \tag{1.8}$$

dove è stato definito il simbolo di Christoffel:

$$\Gamma^i_{jk} := \frac{1}{2}g^{il}\left(\frac{\partial g_{lj}}{\partial x^k} + \frac{\partial g_{lk}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^l}\right) \tag{1.9}$$

Questa equazione del moto è nota come equazione geodetica e le sue soluzioni sono dette geodetiche.

1.2 Particelle relativistiche

È possibile estendere la meccanica lagrangiana allo spaziotempo di Minkowski $\mathbb{R}^{1,3}$, descritto dalla metrica:

$$\eta_{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1) \tag{1.10}$$

Dato che questa metrica non è definita positiva, è possibile classificare due punti deparati da una distanza infinitesima $ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu$ in base al segno di ds^2: se $ds^2 < 0$ si dicono timelike-separated, se $ds^2 > 0$ spacelike-separated e se $ds^2 = 0$ lightlike-separated (o null).
A differenza del caso classico, in cui l'orbita è parametrizzata dal tempo (che è assoluto), nel caso relativistico essa deve essere parametrizzata da un generico $\sigma \in \mathbb{R}$ monotono crescente lungo la traiettoria.
In ambito relativistico, il principio di minima azione ha un'interpretazione geometrica: la traiettoria deve estremizzare la distanza tra due punti dello spaziotempo. Di conseguenza, dato che una particella di massa m deve seguire una traiettoria timelike, si definisce l'azione come:

$$S = -mc\int_{x_1}^{x_2}\sqrt{-ds^2} = -mc\int_{\sigma_1}^{\sigma_2}\sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}} \tag{1.11}$$

Il coefficiente è necessario per rendere l'azione dimensionalmente omogenea con $\hbar$.
L'azione così definita presenta due simemtrie:

1. invarianza di Lorentz: l'azione è invariante per $x^\mu \mapsto \Lambda^\mu{}_\nu x^\nu$, con $\Lambda : \Lambda^\mu{}_\rho\eta_{\mu\nu}\Lambda^\nu{}_\sigma = \eta_{\rho\sigma}$;

2. invarianza per riparametrizzazioni: essendo σ un parametro arbitrario, è normale che l'azione non dipenda dalla sua scelta, infatti se si riparametrizza con una funzione monotona $\tilde{\sigma}(\sigma)$ si ha:

$$\tilde{S} = -mc \int_{\tilde{\sigma}_1}^{\tilde{\sigma}_2} d\tilde{\sigma} \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\tilde{\sigma}} \frac{dx^\nu}{d\tilde{\sigma}}} = -mc \int_{\sigma_1}^{\sigma_2} d\sigma \frac{d\tilde{\sigma}}{d\sigma} \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\sigma} \frac{dx^\nu}{d\sigma} \left(\frac{d\sigma}{d\tilde{\sigma}} \right)^2} = S \tag{1.12}$$

Grazie all'invarianza per riparametrizzazioni, il valore dell'azione tra due punti dello spaziotempo assume un significato ben preciso, il tempo proprio, ovvero il tempo misurato dalla particella in moto stessa:

$$\tau(\sigma) = \frac{1}{c} \int_0^\sigma d\sigma' \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\sigma'} \frac{dx^\nu}{d\sigma'}} \tag{1.13}$$

Una conseguenza dell'identificazione tra azione e tempo proprio è che il principio di minima azione richiede che la traiettoria estremizzi il tempo proprio. È anche possibile riparametrizzare l'azione col tempo proprio, essendo questo una funzione monotona crescente lungo la traiettoria.

1.2.1 Equazione geodetica

Nel caso relativistico su una varietà differenziabile generica, la lagrangiana di una particella libera è:

$$L = \sqrt{-g_{\mu\nu}(x) \dot{x}^\mu \dot{x}^\nu} \tag{1.14}$$

Dunque, le equazioni di Eulero-Lagrange diventano:

$$-\frac{1}{2L} \frac{\partial g_{\mu\nu}}{\partial x^\rho} \dot{x}^\mu \dot{x}^\nu - \frac{d}{d\sigma} \left(-\frac{1}{L} g_{\rho\nu} \dot{x}^\nu \right) = 0 \tag{1.15}$$

L'unica differenza con Eq. 1.5 è che $L = L(\sigma)$, dunque si trova un'equazione analoga all'Eq. 1.7 ma con un termine aggiuntivo:

$$g_{\mu\rho} \ddot{x}^\rho + \frac{1}{2} \left(\frac{\partial g_{\mu\rho}}{\partial x^\nu} + \frac{\partial g_{\mu\nu}}{\partial x^\rho} - \frac{\partial g_{\nu\rho}}{\partial x^\mu} \right) \dot{x}^\nu \dot{x}^\rho = \frac{1}{L} \frac{dL}{d\sigma} g_{\mu\rho} \dot{x}^\rho \tag{1.16}$$

È possibile annullare il termine $\frac{dL}{d\sigma}$ con un'opportuna scelta di parametrizzazione. Dall'Eq. 1.13 si vede che:

$$c \frac{d\tau}{d\sigma} = L(\sigma) \tag{1.17}$$

Dunque, riparametrizzando con $\tau(\sigma)$:

$$L(\tau) = \sqrt{-g_{\mu\nu}(x) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}} = \frac{d\sigma}{d\tau} L(\sigma) = c \tag{1.18}$$

In generale, qualsiasi riparametrizzazione con $\tilde{\tau} = a\tau + b$ (parametri affini della worldline) porta ad avere una lagrangiana costante.
Ricordando la definizione di connessione affine in Eq. 1.9, si trova l'*equazione geodetica*:

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\rho} \frac{dx^\nu}{d\tau} \frac{dx^\rho}{d\tau} \tag{1.19}$$

1.2.2 Momento coniugato

La differenza sostanziale tra lo spazio euclideo e lo spaziotempo di Minkowski è che, mentre nello spazio euclideo un corpo può rimanere fermo, nello spaziotempo nessun corpo può fermarsi nella direzione temporale. Questo fatto deve essere rispecchiato dal momento della particella:

$$p_\mu = \frac{dL}{dx^\mu} = \frac{d}{dx^\mu}\left(-mc\sqrt{-\eta_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}\right) = mc\frac{\eta_{\mu\nu}\dot{x}^\nu}{\sqrt{-\eta_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma}} = -\frac{m^2c^2}{L}\eta_{\mu\nu}\dot{x}^\nu \tag{1.20}$$

Non tutte le componenti del 4-momento sono indipendenti:

$$p^2 = p^\mu p_\mu = \frac{m^4c^4}{L^2}\eta_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = -m^2c^2 \tag{1.21}$$

$$\left(p^0\right)^2 = \mathbf{p}^2 + m^2c^2 \tag{1.22}$$

Di conseguenza, si ha sempre $\left(p^0\right)^2 > 0$.
Si noti che riparametrizzando la worldline col tempo proprio, dato che $\frac{d\tau}{d\sigma} = -\frac{L}{mc^2}$:

$$p^\mu = m\frac{d\sigma}{d\tau}\frac{dx^\mu}{d\sigma} = m\frac{dx^\mu}{d\tau} \tag{1.23}$$

La non-indipendenza di una delle componenti del 4-momento è naturale: da una descrizione classica del sistema risultano tre gradi di libertà $x^i(t)$, dunque passando ad una descrizione relativistica non può risultare un grado di libertà in più. Ciò è legato all'invarianza per riparametrizzazione: risolvendo le equazioni del moto si trovano le componenti della traiettoria $x^\mu = x^\mu(\sigma)$, ma il parametro σ non può rappresentare dell'informazione sul sistema, dunque una delle quattro equazioni del moto va utilizzata per eliminare la dipendenza da σ, riducendo di nuovo a tre i gradi di libertà.

1.2.3 Interazioni

Dall'invarianza per riparametrizzazione, è possibile scegliere come parametro $\sigma = t$ il tempo misurato in un qualunque RF inerziale; considerando una particella libera nello spaziotempo di Minkowski, l'azione in Eq. 1.11 diventa:

$$S = -mc^2\int_{t_0}^{t_1} dt\,\sqrt{1-\frac{\dot{\mathbf{x}}^2}{c^2}} \tag{1.24}$$

In questa forma, è chiara la presenza di soli tre gradi di libertà dovuti a $\mathbf{x}(t)$.

1.2.3.1 Elettromagnetismo

Analogamente al caso classico, per rappresentare l'interazione elettromagnetica è necessario aggiungere un termine potenziale all'azione. Il problema è che la semplice aggiunta di $\int d\sigma V(\mathbf{x})$ non soddisfa l'invarianza per riparametrizzazione: per soddisfarre questo requisito, è necessario individuare un potenziale che cancelli il fattore Jacobiano derivante dalla trasformazione della misura $d\sigma$. Un'opzione è considerare un termine lineare in $\dot{x}^\mu$, dunque per l'invarianza di Lorentz è necessario che l'indice μ sia contratto:

$$S = \int_{\sigma_1}^{\sigma_2} d\sigma\left[-mc^2\sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{x^\nu}{d\sigma}} - qA_\mu(x)\dot{x}^\mu\right] \tag{1.25}$$

dove q è la carica associata all'interazione e $A_\mu(x)$ è il suo potenziale quadrivettoriale.
Scrivendo $A_\mu(x) = (\phi(x)/c, \mathbf{A}(x))$, l'azione in Eq. 1.25 descrive l'interazione elettromagnetica; ciò diventa evidente riparametrizzando con $\sigma = t$:

$$S = \int_{t_0}^{t_1} dt \left[-mc^2 \sqrt{1 - \frac{\dot{\mathbf{x}}^2}{c^2}} - q\phi(x) - q\mathbf{A}(x) \cdot \mathbf{x} \right] \tag{1.26}$$

1.2.3.2 Gravitazione

Per descrivere l'interazione gravitazione è necessario considerare un'azione generalizzata del tipo:

$$S = \int_{t_0}^{t_1} dt \left[-mc^2 \sqrt{1 + \frac{2\Phi(\mathbf{x})}{c^2} - \frac{\dot{\mathbf{x}}^2}{c^2}} \right] \tag{1.27}$$

Nel limite non-relativistico $\dot{\mathbf{x}}^2 \ll c^2$ e $2\Phi(\mathbf{x}) \ll c^2$, dunque approssimando al prim'ordine:

$$S = \int_{t_0}^{t_1} dt \left[-mc^2 + \frac{m}{2}\dot{\mathbf{x}}^2 - m\Phi(\mathbf{x}) \right] \tag{1.28}$$

Il primo termine (l'energia a riposo della particella) non ha effetti sull'azione poiché è costante, mentre gli altri termini descrivono il moto non-relativistico di una particella in un campo gravitazionale $\Phi(\mathbf{x})$. Il termine $1 + 2\Phi(\mathbf{x})/c^2$ in Eq. 1.27 deriva dalla componente η_{00} della metrica, dunque si osserva che la metrica deve dipendere da x, ovvero la descrizione dell'interazione gravitazionale introduce uno spaziotempo curvo. La condizione che deve soddisfarre la metrica in un weak gravitational field è:

$$g_{00}(x) \approx -\left(1 + \frac{2\Phi(x)}{c^2}\right) \tag{1.29}$$

con $\Phi(x)$ il campo gravitazionale Newtoniano.

1.3 Principio di Equivalenza

Come si evince dall'Eq. 1.28, la "carica" dell'interazione gravitazionale è porprio la massa della particella: questo fatto viene definito *weak equivalence principle* (WEP) ed è solitamente espresso tramite l'uguaglianza tra la massa inerziale e la massa gravitazionale:

$$m_\mathrm{i} = m_\mathrm{g} \tag{1.30}$$

Questo è un fatto sperimentale misurato con una precisione dell'ordine di 10^{-13}.

1.3.1 Metrica di Kottler-Möller

Una conseguenza del WEP è l'indistinguibilità tra un'accellerazione costante ed un campo gravitazionale costante: ciò può essere visto in maniera analitica.
Considerando una particella di massa m con accellerazione costante $\mathbf{a} = a\hat{\mathbf{e}}_x$ in un RF inerziale $\mathcal{O}$, dalla relatività speciale si vede subito che la traiettoria non può essere $x(t) = \frac{1}{2}at^2$, poiché la velocità eccederebbe c; bisogna invece ricordare la composizione relativistica delle velocità:

$$v = \frac{v_1 + v_2}{1 + v_1 v_2/c^2} \tag{1.31}$$

È possibile definire la rapidità $\varphi : v = c \tanh \varphi$, così da poter riscrivere la composizione delle velocità come $\varphi = \varphi_1 + \varphi_2$.
Un'accelerazione costante significa che la rapidità della particella aumenta linearmente rispetto al tempo proprio, ovvero $\varphi(\tau) = a\tau/c$, quindi:

$$v(\tau) = \frac{dx}{d\tau} = c \operatorname{sech}\left(\frac{a\tau}{c}\right) \tag{1.32}$$

La relazione tra il tempo misurato nel RF inerziale ed il tempo proprio è:

$$\frac{dt}{d\tau} = \gamma(\tau) = \sqrt{\frac{1}{1 - v(\tau)^2/c^2}} = \cosh\left(\frac{a\tau}{c}\right) \quad \Longrightarrow \quad t(\tau) = \frac{c}{a} \sinh\left(\frac{a\tau}{c}\right) \tag{1.33}$$

con costante d'integrazione tale per cui $\tau = 0$ corrisponda a $t = 0$. Per quanto riguarda la traiettoria:

$$x(\tau) = \frac{c^2}{a} \cosh\left(\frac{a\tau}{c}\right) - \frac{c^2}{a} \tag{1.34}$$

con costante d'integrazione tale per cui $x(0) = 0$. Si trova dunque un'iperbole nello spaziotempo:

$$\left(x + \frac{c^2}{a}\right)^2 - c^2 t^2 = \frac{c^4}{a^2} \tag{1.35}$$

con asintoti $ct = \pm\left(x + c^2/a\right)$ per $\tau \to \pm\infty$. Si può mostrare che la trasformazione tra le coordinate (t, x) nel RF inerziale e quelle (τ, ρ) solidali alla particella è data da:

$$\begin{aligned} ct &= \left(\rho + \frac{c^2}{a}\right) \sinh\left(\frac{a\tau}{c}\right) \\ x &= \left(\rho + \frac{c^2}{a}\right) \cosh\left(\frac{a\tau}{c}\right) - \frac{c^2}{a} \end{aligned} \tag{1.36}$$

Infatti, l'orbita della particella è giustamente descritta da $\rho = 0$. Inoltre, si vede che le coordinate (τ, ρ) non ricoprono tutto lo spaziotempo di Minkowski (t, x): questo dimostra che ci sono delle regioni dello spaziotempo causalmente disconnesse dalla particella.
Ricordando che $ds^2 = -c^2 dt^2 + d\mathbf{r}^2$, sostituendo l'Eq. 1.36 si ottiene la *metrica di Kottler-Möller*:

$$ds^2 = -\left(1 + \frac{a\rho}{c^2}\right)^2 c^2 d\tau^2 + d\rho^2 + dy^2 + dz^2 \tag{1.37}$$

Mentre la parte spaziale rimane piatta, si vede che $g_{00} = g_{00}(x)$; inoltre, nel caso sub-relativistico:

$$g_{00} \approx -\left(1 + \frac{2a\rho}{c^2}\right) \tag{1.38}$$

Definendo $\Phi(\rho) = a\rho$, si trova proprio la condizione 1.29: questo è proprio l'assero del WEP, poiché un'accellerazione costante dà una metrica indistinguibile da quella di un campo gravitazionale costante (sub-relativistico).

1.3.1.1 Principio di equivalenza di Einstein

Dal WEP deriva il fatto che un campo gravitazionale costante può essere annullato dalla scelta di un particolare RF, il free-fall RF.
Il *principio di equivalenza di Einstein* è una generalizzazione del WEP: esso afferma che esiste sempre un local RF in cui gli effetti di un qualsiasi campa gravitazionali sono localmente annullati. Formalmente, ciò equivale a dire che la metrica $g_{\mu\nu}$ è sempre localmente approssimabile con la metrica di Minkowski $\eta_{\mu\nu}$.
Gli effetti di un campo gravitazionale non uniforme diventano evidenti quando è possibile svolgere misurazioni su una regione estesa di spazio. Si consideri ad esempio un osservatore confinato in un cubo chiuso in free-fall verso la Terra: non esiste alcun esperimento locale in grado di distinguere se l'osservatore stia fluttuando nello spazio oppure sia in free-fall, bensì è necessario un esperimento non-locale; un esempio di questo tipo di esperimenti consiste nel lasciare libere due masse test (quindi non influenzate vicendevolmente per interazione gravitazionale) separate da una certa distanza: se il cubo sta fluttuando nello spazio, le masse rimaranno nella loro posizione iniziale per inerzia, mentre se esso è in free-fall ciascuna massa sarà attratta verso il centro della Terra, dunque il loro spostamento avrà non solo una componente verticale (rispetto alla caduta), ma anche una orizzontale, per quanto piccola: le masse si sposteranno dunque l'una verso l'altra per effetto di una *tidal force*, uno dei principali fattori discriminanti dei campi gravitazionali non uniformi.

1.3.2 Gravitational time dilation

In condizioni di campo gravitazionale debole si ha $g_{00}(x) = 1 + \frac{2\Phi(x)}{c^2}$. Considerando il campo gravitazionale di un corpo sferico di massa M uniforme, $\Phi(r) = -\frac{GM}{r}$, si ha che un osservatore ad una distanza fissa r misurerà degli intervalli di tempo dati da:

$$d\tau^2 = g_{00}dt^2 = \left(1 - \frac{2GM}{rc^2}\right) dt^2 \tag{1.39}$$

Dunque, definendo t il tempo misurato da un osservatore a $r \to \infty$, l'osservatore a r misurerà:

$$T(r) = t\sqrt{1 - \frac{2GM}{rc^2}} \tag{1.40}$$

ovvero il tempo scorre più lentamente vicino ad un corpo massivo.
È anche possibile mettere in relazione i tempi misurati a due distanze finite r_1 ed $r_2 = r_1 + \Delta r$:

$$\begin{aligned} T_2 = t\sqrt{1 - \frac{2GM}{(r_1 + \Delta r)c^2}} &\approx t\sqrt{1 - \frac{2GM}{r_1c^2} + \frac{2GM\Delta r}{r_1^2c^2}} \\ &\approx t\sqrt{1 - \frac{2GM}{r_1c^2}}\left(1 + \frac{GM\Delta r}{r_1^2c^2}\right) = T_1\left(1 + \frac{GM\Delta r}{r_1^2c^2}\right) \end{aligned} \tag{1.41}$$

dove si è assunto che $\Delta r \ll r_1$ e $2GM \ll r_1c^2$. Ad esempio, con un per un dislivello di 10^3 m sul livello del mare ($\sim 6 \cdot 10^6$ m) si ha una differenza di 10^{-12} s, confermato sperimentalmente.

1.3.2.1 Gravitational redshift

Un'importante conseguenza della gravitational time dilation è il gravitational redshift.
Sempre in condizioni di campo debole, si consideri un segnale a distanza r_1 che si ripete ad intervalli

ΔT_1: un osservatore a r_2 misurerà:

$$\Delta T_2 = \sqrt{\frac{1 + 2\Phi(r_2)/c^2}{1 + 2\Phi(r_1)/c^2}}\Delta T_1 \approx \left(1 + \frac{\Phi(r_2) - \Phi(r_1)}{c^2}\right)\Delta T_1 \tag{1.42}$$

Dato che $\omega \sim T^{-1}$, si ha:

$$\omega_2 \approx \left(1 + \frac{\Phi(r_2) - \Phi(r_1)}{c^2}\right)^{-1}\omega_1 \tag{1.43}$$

Dato che $\Phi(r) \sim r^{-1}$ (nel caso considerato), se $r_2 > r_1$ si ha $\omega_2 < \omega_1$ (redshift), mentre se $r_2 < r_1$ si ha $\omega_2 > \omega_1$ (blueshift).

Part II

Differential Geometry

Chapter 2

Manifolds

2.1 Topological spaces

Definition 2.1.1. The *topology* $\mathcal{T}$ of a set X is a family of subsets of X, i.e. $\mathcal{T} \subseteq \mathcal{P}(X)$, defined as *open sets*, with the following properties:

1. $\emptyset, X \in \mathcal{T}$;
2. $O_\alpha, O_\beta \in \mathcal{T} \Rightarrow O_\alpha \cap O_\beta \in \mathcal{T}$;
3. $\{O_\alpha\}_{\alpha \in I} \subset \mathcal{T}$ (I arbitrary index set) $\Rightarrow \bigcup_{\alpha \in I} O_\alpha \in \mathcal{T}$.

Definition 2.1.2. A *topological space* M is a set of points, endowed with a topology $\mathcal{T}$.

Definition 2.1.3. Given a topological space $(M, \mathcal{T})$, $O \in \mathcal{T}$ is a *neighbourhood* of a point $p \in M$ if $p \in O$.

Definition 2.1.4. A topological space $(M, \mathcal{T})$ is *Hausdorff* if $\forall p, q \in M \, \exists O_1, O_2 \in \mathcal{T}$ neighbourhoods of p and q respectively such that $O_1 \cap O_2 = \emptyset$.

Definition 2.1.5. A *homeomorphism* between two topological spaces $(M_1, \mathcal{T}_1)$ and $(M_2, \mathcal{T}_2)$ is a bijective map $f : M_1 \to M_2$ which is bicontinuous, i.e. both f and f^{-1} are continuous: f is continuous if $O \in \mathcal{T}_2 \Rightarrow f^{-1}(O) \in \mathcal{T}_1$.

2.2 Differentiable Manifolds

Definition 2.2.1. An n-dimensionale *differentiable manifold* $\mathcal{M}$ is a Hausdorff topological space such that:

1. $\mathcal{M}$ is locally homeomorphic to $\mathbb{R}^n$, i.e. $\forall p \in \mathcal{M} \, \exists O \in \mathcal{T}(\mathcal{M}) : p \in O \wedge \exists \varphi : O \to U \in \mathcal{T}(\mathbb{R}^n)$ homeomorphism;
2. given $O_\alpha, O_\beta \in \mathcal{T}(\mathcal{M}) : O_\alpha \cap O_\beta \neq \emptyset$, the corresponding maps $\varphi_\alpha : O_\alpha \to U_\alpha, \varphi_\beta : O_\beta \to U_\beta$ must be *compatible*, i.e. $\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(O_\alpha \cap O_\beta) \to \varphi_\beta(O_\alpha \cap O_\beta)$ and its inverse must be smooth (of $\mathcal{C}^\infty$ class).

The maps φ_α are called *charts* and a collection of compatible charts is called an *atlas*: a *maximal atlas* $\mathcal{A}$ is an atlas such that $\bigcup_{\alpha\in I} O_\alpha = \mathcal{M}$. Two atlases are compatible if each chart of one atlas is compatible with every chart of the other: they define the same *differentiable structure* on the manifold.
Each chart φ_α provides a coordinate system on the region O_α: $\varphi_\alpha(p) = (x^1(p), \dots, x^\mu(p), \dots, x^n(p))$. The *transition functions* $\varphi_\beta \circ \varphi_\alpha^{-1}$ are therefore coordinate transformations on overlapping regions.

Example 2.2.1. The n-sphere $\mathbb{S}^n$ is a differentiable manifold.

Example 2.2.2. To define a differentiable structure on $\mathcal{S}^1$ an atlas of two charts is needed: the standard parametrization $\theta \in [0, 2\pi)$ is not a well-defined chart because $[0, 2\pi)$ is not an open set in the Euclidean topology of $\mathbb{R}$, therefore the elimination of a point is necessary; usually, the two charts of the atlas are defined by $\theta_1 \in (0, 2\pi)$, excluding $(1, 0)$ (in the embedding space $\mathbb{R}^2$), and $\theta_2 \in (-\pi, \pi)$, excluding $(-1, 0)$: they are evidently compatible, thus they form a maximal atlas.

2.2.1 Maps between manifolds

Locally mapping $\mathcal{M}$ to $\mathbb{R}^n$ allows to import concepts of Analysis from $\mathbb{R}^n$ to $\mathcal{M}$.

Definition 2.2.2. A function $f : \mathcal{M} \to \mathbb{R}$ on a differentiable manifold $(\mathcal{M}, \mathcal{A})$ is *smooth* if $f \circ \varphi_\alpha^{-1} : U_\alpha \to \mathbb{R}$ is smooth for all charts $(U_\alpha, \varphi_\alpha) \in \mathcal{A}$.

Definition 2.2.3. A map $f : \mathcal{M} \to \mathcal{N}$ between two differentiable manifolds $(\mathcal{M}, \mathcal{A}_1), (\mathcal{N}, \mathcal{A}_2)$ is *smooth* if $\psi_{\alpha_2} \circ f \circ \varphi_{\alpha_1}^{-1} : U_{\alpha_1} \to V_{\alpha_2}$ is smooth for all charts $(U_{\alpha_1}, \varphi_{\alpha_1}) \in \mathcal{A}_1, (V_{\alpha_2}, \varphi_{\alpha_2}) \in \mathcal{A}_2$.

Definition 2.2.4. A *diffeomorphism* between two differentiable manifolds $\mathcal{M}, \mathcal{N}$ is a smooth homeomorphism $f : \mathcal{M} \to \mathcal{N}$.

Proposition 2.2.1. *If $\mathcal{M}$ and $\mathcal{N}$ are diffeomorphic, then* $\dim_{\mathbb{R}} \mathcal{M} = \dim_{\mathbb{R}} \mathcal{N}$.

Example 2.2.3. $\mathbb{S}^7$ can be covered by multiple incompatible atlases: the resulting manifolds are homeomorphic but not diffeomorphic.

Example 2.2.4. $\mathbb{R}^n$ has a unique differentiable structure for all $n \in \mathbb{N}$, except for $n = 4$: $\mathbb{R}^4$ can be covered by infinitely-many incompatible atlases.

2.3 Tangent spaces

The notions of calculus can be defined on a differential manifold $(\mathcal{M}, \mathcal{A})$ via tangent spaces.

Definition 2.3.1. The derivative of a function $f : \mathcal{M} \to \mathbb{R}$ at a point $p \in \mathcal{M}$, covered by the chart (φ, U), is defined as:

$$\left.\frac{\partial f}{\partial x^\mu}\right|_p := \left.\frac{\partial (f \circ \varphi^{-1})}{\partial x^\mu}\right|_{\varphi(p)} \tag{2.1}$$

Evidently, this definition depends on the choise of coordinates x^μ, thus it depends on the chart.

2.3.1 Tangent vectors

Definition 2.3.2. The set of all smooth functions on $\mathcal{M}$ is denoted by $\mathcal{C}^\infty(\mathcal{M})$.

Definition 2.3.3. A *tangent vector* to $\mathcal{M}$ in $p \in \mathcal{M}$ is an operator $X_p : \mathcal{C}^\infty(\mathcal{M}) \to \mathbb{R}$ such that:

1. $X_p(f+g) = X_p(f) + X_p(g) \, \forall f, g \in \mathcal{C}^\infty(\mathcal{M})$;
2. $X_p(f) = 0$ for all constant functions;
3. $X_p(fg) = X_p(f)g(p) + f(p)X_p(g) \, \forall f, g \in \mathcal{C}^\infty(\mathcal{M})$.

Proposition 2.3.1. $X_p(\alpha f) = \alpha X_p(f) \, \forall \alpha \in \mathbb{R}$.

Proof. Trivial from conditions 2. and 3. of Def. 2.3.3. □

It is simple to check that $\partial_\mu|_p$ satisfies the conditions of Def. 2.3.3.

Theorem 2.3.1. *The set $T_p\mathcal{M}$ of all tangent vectors at a point $p \in \mathcal{M}$ forms an n-dimensional space, called tangent space, and $\{\partial_\mu|_p\}_{\mu=1,\dots,n}$ is a basis of such space.*

Proof. Defining $f \circ \varphi^{-1} \equiv F : U \subset \mathcal{M} \to \mathbb{R}$, with $f : \mathcal{M} \to \mathcal{M}$ and $(\varphi, U) \in \mathcal{A}$, it can be proved that, in some neighbourhood of p (not necessarily U), F cal always be written as:

$$F(x) = F(x^\mu(p)) + (x^\mu - x^\mu(p))\, F_\mu(x)$$

for some n functions F_μ (ex.: a Taylor series, or more generally $F(x) = F(0) + x\int_0^1 dt\, F(xt)$). Applying $\partial_\mu|_{x(p)}$:

$$\left.\frac{\partial F}{\partial x^\mu}\right|_{x(p)} = F_\mu(x(p))$$

Defining $f_\mu \equiv F_\mu \circ \varphi$, for any $q \in \mathcal{M}$ in an appropriate neighbourhood of p:

$$f(q) = f(p) + (x^\mu(q) - x^\mu(p))\, f_\mu(q)$$

Moreover, remembering Eq. 2.1:

$$f_\mu(p) = F_\mu \circ \varphi(p) = F_\mu(x(p)) = \left.\frac{\partial F}{\partial x^\mu}\right|_{x(p)} = \left.\frac{\partial f}{\partial x^\mu}\right|_p$$

Using these facts, the action of a tangent vector can be written explicitly:

$$\begin{aligned} X_p(f) &= X_p\left(f(p) + (x^\mu - x^\mu(p))\, f_\mu\right) \\ &= X_p\left(f(p)\right) + X_p\left((x^\mu - x^\mu(p))\right) f_\mu(p) + (x^\mu - x^\mu(p))\,(p) X_p\left(f_\mu\right) \\ &= X_p\left(x^\mu\right) f_\mu(p) \end{aligned}$$

because $f(p)$ is a constant and $(x^\mu - x^\mu(p))\,(p) = x^\mu(p) - x^\mu(p) = 0$. Therefore, remembering the expression for $f_\mu(p)$:

$$X_p = X_p(x^\mu) \left.\frac{\partial}{\partial x^\mu}\right|_p \equiv X^\mu \left.\frac{\partial}{\partial x^\mu}\right|_p$$

Thus, $T_p\mathcal{M} = \text{span}\{\partial_\mu|_p\}$. To check for linear independence, suppose $\alpha = \alpha^\mu \partial_\mu|_p \equiv 0$: acting on $f = x^\nu$, it gives $\alpha(f) = \alpha_\mu \partial_\mu(x^\nu)|_p = \alpha_\nu = 0$. This concludes the proof. □

2.3.1.1 Changing coordinates

Although $\partial_\mu|_p$ depends on the choice of coordinates (it is a *coordinate basis*), the existence of X_p is independent of that choice.
If two different charts $(\varphi, U), (\tilde{\varphi}, V)$ intersect in a neighbourhood of $p \in U \cap V$, the transition from x^μ to y^μ can be expressed as:

$$X_p(f) = X^\mu \frac{\partial f}{\partial x^\mu}\bigg|_p = X^\mu \frac{\partial y^\nu}{\partial x^\mu}\bigg|_{\varphi(p)} \frac{\partial f}{\partial y^\nu}\bigg|_p \tag{2.2}$$

This equation can have two interpretations: the alibi interpretation:

$$\frac{\partial}{\partial x^\mu}\bigg|_p = \frac{\partial y^\nu}{\partial x^\mu}\bigg|_{\varphi(p)} \frac{\partial}{\partial y^\nu}\bigg|_p \tag{2.3}$$

and the alias interpretation:

$$\tilde{X}^\nu = X^\mu \frac{\partial y^\nu}{\partial x^\mu}\bigg|_{\varphi(p)} \tag{2.4}$$

Components of vectors which transform this way are called *contravariant*.

2.3.1.2 Curves

Consider a smooth curve on $\mathcal{M}$, i.e. a smooth map $\sigma : I \in \mathcal{T}(\mathbb{R}) \to \mathcal{M}$, parametrized as $\sigma(t) : \sigma(0) = p \in \mathcal{M}$; with a given chart (φ, U), this curve becomes $\varphi \circ \sigma : I \to \mathbb{R}^n$, parametrized by $x^\mu(t)$. The *tangent vector* to the curve in p is:

$$X_p = \frac{dx^\mu(t)}{dt}\bigg|_{t=0} \frac{\partial}{\partial x^\mu}\bigg|_p \tag{2.5}$$

This operator, applied to a function $f \in \mathcal{C}^\infty(\mathcal{M})$, calculates the directional derivative of f along the curve. It can be showed that every tangent vector can be written as in Eq. 2.5, therefore the tangent space is literally the space of all possible tangents to curves passing through p.
It must be noted that tangent spaces at different points are entirely different spaces: there's no way to directly compare vectors between them.

2.3.2 Vector fields

Definition 2.3.4. A *vector field* X is a smooth map $X : p \in \mathcal{M} \mapsto X_p \in T_p\mathcal{M}$. It can also be viewed as a smooth map $X : \mathcal{C}^\infty(\mathcal{M}) \to \mathcal{C}^\infty(\mathcal{M})$, as $(X(f))(p) = X_p(f) \in \mathbb{R}$.

Definition 2.3.5. The space of all vector fields on $\mathcal{M}$ is denoted by $\mathfrak{X}(\mathcal{M})$.

Given a chart (φ, U), a vector field X can be expressed as:

$$X = X^\mu \frac{\partial}{\partial x^\mu} \tag{2.6}$$

with $X^\mu \in \mathcal{C}^\infty(\mathcal{M})$. This expression is only defined on U.

2.3.2.1 Lie brakets

Given two vector fields $X, Y \in \mathfrak{X}(\mathcal{M})$, their product is clearly not a vector field, as it does not satisfy Leibniz' rule:

$$XY(fg) = XY(f)g + Y(f)X(g) + X(f)Y(g) + fXY(g) \neq XY(f)g + fXY(g)$$

where $XY(f) \equiv X(Y(f))$.

Definition 2.3.6. Given two vector fields $X, Y \in \mathfrak{X}(\mathcal{M})$, their *commutator* (or *Lie bracket*) is defined as:

$$[X, Y](f) = XY(f) - YX(f) \tag{2.7}$$

With a given chart:

$$\begin{aligned}
[X, Y](f) &= X^\mu \frac{\partial}{\partial x^\mu}\left(Y^\nu \frac{\partial f}{\partial x^\nu}\right) - Y^\mu \frac{\partial}{\partial x^\mu}\left(X^\nu \frac{\partial f}{\partial x^\nu}\right) \\
&= \left(X^\mu \frac{\partial Y^\nu}{\partial x^\mu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu}\right)\frac{\partial f}{\partial x^\nu}
\end{aligned}$$

therefore:

$$[X, Y] = \left(X^\mu \frac{\partial Y^\nu}{\partial x^\mu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu}\right)\frac{\partial}{\partial x^\nu} \tag{2.8}$$

Theorem 2.3.2 (Jacobi)**.** *Given $X, Y, Z \in \mathfrak{X}(\mathcal{M})$, the Jacobi identity holds:*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \tag{2.9}$$

Proposition 2.3.2. *$\mathfrak{X}(\mathcal{M})$ is a Lie algebra.*

2.3.2.2 Integral curves

Definition 2.3.7. A *flow* on $\mathcal{M}$ is a one-parameter family of diffeomorphisms $\sigma_t : \mathcal{M} \to \mathcal{M}$, labelled by $t \in \mathbb{R}$, with group structure: $\sigma_0 = \mathrm{id}_\mathcal{M}$ and $\sigma_s \circ \sigma_t = \sigma_{s+t}$, thus $\sigma_{-t} = \sigma_t^{-1}$.

Such flows give rise to streamlines on the manifold: these streamlines are required to be smooth. Defining $x^\mu(\sigma_t) \equiv x^\mu(t)$, a vector field can be defined by the tangent to the streamlines at each point on the manifold:

$$X^\mu(x^\mu(t)) = \frac{dx^\mu(t)}{dt} \tag{2.10}$$

The inverse reasoning is also possible.

Definition 2.3.8. Given a vector field $X \in \mathfrak{X}(\mathcal{M})$, streamlines described by Eq. 2.10 are called *integral curves* generated by X.

Proposition 2.3.3. *The infinitesimal flow generated by $X \in \mathfrak{X}(\mathcal{M})$ is:*

$$x^\mu(t) = x^\mu(0) + tX^\mu(x(t)) + o(t) \tag{2.11}$$

Definition 2.3.9. A vector field which generates a flow defined for all $t \in \mathbb{R}$ is called *complete*.

Theorem 2.3.3. *If $\mathcal{M}$ is compact, then all $X \in \mathfrak{X}(\mathcal{M})$ are complete.*

Example 2.3.1. On $\mathbb{S}^2$, the flow generated by $X = \partial_\phi$ is described by $\dot{\phi} = 1, \dot{\theta} = 0$, thus $\theta(t) = \theta_0$ and $\phi(t) = \phi_0 + t$: the flow lines are lines of constant latitude.

2.3.3 Lie derivative

Defining calculus for vector fields requires a way to compare vectors of different tangent spaces.

Definition 2.3.10. Given a diffeomorphism between two manifolds $\varphi : \mathcal{M} \to \mathcal{N}$ and a function $f : \mathcal{N} \to \mathbb{R}$, the *pull-back* of f is the function $\varphi^* f : \mathcal{M} \to \mathbb{R}$ such that $\varphi^* f(p) = f(\varphi(p))$.

Definition 2.3.11. Given a diffeomorphism between two manifolds $\varphi : \mathcal{M} \to \mathcal{N}$ and a vector field $X \in \mathfrak{X}(\mathcal{M})$, the *push-forward* of X is the vector field $\varphi_* X \in \mathfrak{X}(\mathcal{N})$ such that $\varphi_* X(f) = X(\varphi^* f)$.

This last equality must be evaluated at the appropriate points: $[\varphi_* X(f)](\varphi(p)) = [X(\varphi^* f)](p)$.
With the appropriate charts on $\mathcal{M}$ and $\mathcal{N}$, the definitions above can be rewritten with coordinates:

$$\varphi^* f(x) = f(y(x)) \tag{2.12}$$

$$\varphi_* X(f) = X^\mu \frac{\partial f(y(x))}{\partial x^\mu} = X^\mu \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial f(y)}{\partial y^\alpha} \tag{2.13}$$

The notions of pull-back and push-forward allow to compare tangent vectors at neighbouring points and, in particular, to define the derivative along a vector field.

Definition 2.3.12. Given a function $f : \mathcal{M} \to \mathbb{R}$ and a vector field $X \in \mathfrak{X}(\mathcal{M})$, the derivative of f along X (called *Lie derivative*) is defined as:

$$\mathcal{L}_X f(x) \coloneqq \lim_{t \to 0} \frac{f(\sigma_t(x)) - f(x)}{t} = \left. \frac{df(\sigma_t(x))}{dt} \right|_{t=0} \tag{2.14}$$

where σ_t is the flow generated by X.

Proposition 2.3.4. $\mathcal{L}_X f = X(f)$.

Proof. $\mathcal{L}_X f = \frac{df(\sigma_t)}{dt} = \frac{\partial f}{\partial x^\mu} \frac{dx^\mu(t)}{dt} = X^\mu \frac{\partial f}{\partial x^\mu} = X(f)$. □

Definition 2.3.13. Given two vector fields $X, Y \in \mathfrak{X}(\mathcal{M})$, the *Lie derivative* of Y along X is defined as:

$$\mathcal{L}_X Y_p \coloneqq \lim_{t \to 0} \frac{((\sigma_{-t})_* Y)_p - Y_p}{t} \tag{2.15}$$

where σ_t is the flow generated by X.

The use of the inverse flow σ_{-t} is necessary because to evaluate the vector field $\mathcal{L}_X Y$ at the point $p \in \mathcal{M}$, the tangent vector $Y_{\sigma_t(p)} \in T_{\sigma_t(p)}\mathcal{M}$ must be "pushed-back" to $T_p\mathcal{M} = T_{\sigma_0(p)}\mathcal{M}$.
With $t \to 0$, the infinitesimal flow σ_{-t} is, according to Eq. 2.11, $x^\mu(t) = x^\mu(0) - tX^\mu + o(t)$, therefore the Lie derivative of base tangent vectors can be expressed as:

$$(\sigma_{-t})_* \partial_\mu = \frac{\partial x^\nu(t)}{\partial x^\mu} \frac{\partial}{\partial x^\nu(t)} = \left(\delta^\nu_\mu - t \frac{\partial X^\nu}{\partial x^\mu} + o(t) \right) \partial_\nu(t) \quad \Longrightarrow \quad \mathcal{L}_X \partial_\mu = -\frac{\partial X^\nu}{\partial x^\mu} \partial_\nu \tag{2.16}$$

Proposition 2.3.5. $\mathcal{L}_X Y = [X, Y]$.

Proof. $\mathcal{L}_X Y = \mathcal{L}_X (Y^\mu \partial_\mu) = (\mathcal{L}_X Y^\mu)\, \partial_\mu + Y^\mu \,(\mathcal{L}_X \partial_\mu) = X^\nu \frac{\partial Y^\mu}{\partial x^\nu} \partial_\mu - Y^\mu \frac{\partial X^\nu}{\partial x^\mu} \partial_\nu = [X, Y]$. □

Proposition 2.3.6. $\mathcal{L}_X \mathcal{L}_Y Z - \mathcal{L}_Y \mathcal{L}_X Z = \mathcal{L}_{[X,Y]} Z$.

Proof. Trivial with Jacobi identity. □

2.4 Tensors

2.4.1 Dual Spaces

Definition 2.4.1. Given a vector space V, its *dual* V^* is the space of all linear maps $f : V \to \mathbb{R}$.

Given a basis $\{\mathbf{e}_\mu\}_{\mu=1,\dots,n}$ of V, its *dual basis* $\{\mathbf{f}^\mu\}_{\mu=1,\dots,n}$ of V^* can be defined by:

$$\mathbf{f}^\nu(\mathbf{e}_\mu) = \delta^\nu_\mu \tag{2.17}$$

A general vector in V can be written as $X = X^\mu \mathbf{e}_\mu$, thus according to Eq. 2.17 $X^\mu = \mathbf{f}^\mu(X)$.

Proposition 2.4.1. *The map $f : \mathbf{e}_\mu \mapsto \mathbf{f}^\mu$ is an isomorphism between V and V^*.*

This isomorphism, however, is basis-dependent.

Proposition 2.4.2. $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} V^*$.

Proposition 2.4.3. $(V^*)^* = V$.

Proof. The natural isomorphism between $(V^*)^*$ and V is basis-independent: suppose $X \in V$ and $\omega \in V^*$, so that $\omega(X) \in \mathbb{R}$; X can be viewed as $X \in (V^*)^*$ by setting $V(\omega) \equiv \omega(V)$. □

2.4.2 Cotangent vectors

Definition 2.4.2. Given a differentiable manifold $(\mathcal{M}, \mathcal{A})$ and a point $p \in \mathcal{M}$, the *cotangent space* to $\mathcal{M}$ at p is defined as $T^*_p\mathcal{M} := (T_p\mathcal{M})^*$.

Elements of $T^*_p\mathcal{M}$ are called *cotangent vectors* (or *covectors*).

Definition 2.4.3. A *covector field* (or *1-form*) is a smooth map $\omega : p \in \mathcal{M} \mapsto \omega_p \in T^*_p\mathcal{M}$. It can also be viewed as a smooth map $\omega : \mathfrak{X}(\mathcal{M}) \to \mathcal{C}^\infty(\mathcal{M})$, as $(\omega(X))(p) = \omega_p(X_p) \in \mathbb{R}$.

Definition 2.4.4. The space of all 1-forms on $\mathcal{M}$ is denoted by $\Lambda^1(\mathcal{M})$.

Proposition 2.4.4. *$\{dx^\mu\}_{\mu=1,\dots,n}$ is a basis of $\Lambda^1(\mathcal{M})$ dual to the basis $\{\partial_\mu\}_{\mu=1,\dots,n}$ of $\mathfrak{X}(\mathcal{M})$.*

Proof. Consider $f \in \mathcal{C}^\infty(\mathcal{M})$ and define $df \in \Lambda^1(\mathcal{M})$ by $df(X) = X(f)$: taking $f = x^\mu$ and $X = \partial_\mu$, $df(X) = \partial_\nu(x^\mu) = \delta^\mu_\nu$, therefore $\{dx^\mu\}_{\mu=1,\dots,n}$ is the dual basis of $\Lambda^1(\mathcal{M})$. □

This is also confirmed by $df = \frac{\partial f}{\partial x^\mu} dx^\mu$. These are coordinate basis: in fact, given two different charts $(\varphi, U), (\tilde{\varphi}, V)$:

$$dy^\mu = \frac{dy^\mu}{dx^\nu} dx^\nu \tag{2.18}$$

which is the inverse of Eq. 2.3 (not evaluated at a specific point). This ensures that:

$$dy^\mu \left(\frac{\partial}{\partial y^\nu} \right) = \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial y^\nu} dx^\alpha \left(\frac{\partial}{\partial x^\beta} \right) = \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial x^\alpha}{\partial y^\nu} = \delta^\mu_\nu$$

A 1-form $\omega \in \Lambda^1(\mathcal{M})$ can thus be expressed both as $\omega = \omega_\mu dx^\mu = \tilde{\omega}_\mu dx^\mu$, with:

$$\tilde{\omega}_\omega = \frac{\partial x^\nu}{\partial y^\mu} \omega_\nu \tag{2.19}$$

Components of 1-forms which transform this way are called *covariant*.

Definition 2.4.5. Given a diffeomorphism between two manifolds $\varphi : \mathcal{M} \to \mathcal{N}$ and a 1-form $\omega \in \Lambda^1(\mathcal{N})$, the *pull-back* of ω is the 1-form $\varphi^*\omega \in \Lambda^1(\mathcal{M})$ such that $\varphi^*\omega(X) = \omega(\varphi_* X)$.

With the appropriate charts on $\mathcal{M}$ and $\mathcal{N}$, the definition above can be rewritten with coordiantes:

$$\varphi^*\omega = \omega_\alpha \frac{\partial y^\alpha}{\partial x^\mu} dx^\mu \tag{2.20}$$

Definition 2.4.6. Given a vector field $X \in \mathfrak{X}(\mathcal{M})$ and a 1-form $\omega \in \Lambda^1(\mathcal{M})$, the *Lie derivative* of ω along X is defined as:

$$\mathcal{L}_X \omega_p := \lim_{t \to 0} \frac{(\sigma_t^*\omega)_p - \omega_p}{t} \tag{2.21}$$

where σ_t is the flow generated by X.

In contranst with the Lie derivative of a vector field, which pushes forward with σ_{-t} (i.e. pushes back), the Lie derivative of a 1-form pulls back with σ_t: this results in the difference of a minus sign with respect to Eq. 2.16, giving:

$$\mathcal{L}_X dx^\mu = \frac{\partial X^\mu}{\partial x^\nu} dx^\nu \tag{2.22}$$

Therefore, on a general 1-form $\omega = \omega_\mu dx^\mu$:

$$\mathcal{L}_X \omega = \left(X^\nu \partial_\nu \omega_\mu + \omega_\nu \partial_\mu X^\nu\right) dx^\mu \tag{2.23}$$

2.4.3 Tensor fields

Definition 2.4.7. A *tensor of rank* (r, s) at a $p \in \mathcal{M}$ of a differentiable manifold $(\mathcal{M}, \mathcal{A})$ is a multi-linear map defined as:

$$T_p : \overbrace{T_p^*\mathcal{M} \times \cdots \times T_p^*\mathcal{M}}^{r} \times \overbrace{T_p\mathcal{M} \times \cdots \times T_p\mathcal{M}}^{s} \to \mathbb{R} \tag{2.24}$$

Example 2.4.1. A cotangent vector $\omega_p \in T_p^*\mathcal{M}$ is a tensor of rank $(1, 0)$, while a tangent vector $X_p \in T_p\mathcal{M}$ is a tensor of rank $(0, 1)$.

Definition 2.4.8. A *tensor field* of rank (r, s) is a smooth map $T : p \in \mathcal{M} \mapsto T_p$ tensor of rank (r, s) at p. It can also be viewed as a smooth map $T : [\Lambda^1(\mathcal{M})]^r \times [\mathfrak{X}(\mathcal{M})]^s \to \mathcal{C}^\infty(\mathcal{M})$.

Given appropriate basis for vector fields $\{\mathbf{e}_\mu\}_{\mu=1,\dots,n}$ and 1-forms $\{\mathbf{f}^\mu\}_{\mu=1,\dots,n}$, the components of a tensor field are defined as:

$$T^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s} := T(\mathbf{f}^{\mu_1}, \dots, \mathbf{f}^{\mu_r}, \mathbf{e}_{\nu_1}, \dots, \mathbf{e}_{\nu_s}) \tag{2.25}$$

Proposition 2.4.5. *On an n-dimensional manifold, a (r, s) tensor field has n^{r+s} components, each being element of $\mathcal{C}^\infty(\mathcal{M})$.*

Consider two general basis transformations, for vector fields and 1-forms, described by invertible matrices A, B such that $\tilde{\mathbf{e}}_\mu = A^\nu{}_\mu \mathbf{e}_\nu$ and $\tilde{\mathbf{f}}^\mu = B^\mu{}_\nu \mathbf{f}^\nu$, with necessary condition $A^\mu{}_\nu B^\rho{}_\mu = \delta^\rho_\nu$ to ensure duality: this implies $B = A^{-1}$, i.e. covectors transform inversely with respect to vectors. Thus:

$$\tilde{T}^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s} = B^{\mu_1}{}_{\rho_1} \dots B^{\mu_r}{}_{\rho_r} A^{k_1}{}_{\nu_1} \dots A^{k_s}{}_{\nu_s} T^{\rho_1 \dots \rho_r}{}_{k_1 \dots k_s} \tag{2.26}$$

If the considered basis are coordinate basis, then $A^\mu{}_\nu = \frac{\partial x^\mu}{\partial y^\nu}$ and $B^\mu{}_\nu = \frac{\partial y^\mu}{\partial x^\nu}$.

2.4.4 Operations on tensors

Algebric addition and multiplication by functions are trivially defined on tensors of the same rank.

Proposition 2.4.6. *The space of all (r, s) tensors at a point $p \in \mathcal{M}$ is denoted by $T_p^{(r,s)}\mathcal{M}$, and it is a vector space.*

Definition 2.4.9. Given two tensor fields S of rank (p, q) and T of rank (r, s), their *tensor product* is defined as:

$$\begin{aligned} S \otimes T(\omega_1, \ldots, \omega_p, \eta_1, \ldots, \eta_r, X_1, \ldots, X_q, Y_1, \ldots, Y_s) \\ = S(\omega_1, \ldots, \omega_p, X_1, \ldots, X_q)T(\eta_1, \ldots, \eta_r, Y_1, \ldots, Y_s) \end{aligned} \tag{2.27}$$

or, in components:

$$(S \otimes T)^{\mu_1 \ldots \mu_p \nu_1 \ldots \nu_r}{}_{\rho_1 \ldots \rho_q \ldots \sigma_1 \ldots \sigma_s} = S^{\mu_1 \ldots \mu_p}{}_{\rho_1 \ldots \rho_q} T^{\nu_1 \ldots \nu_r}{}_{\sigma_1 \ldots \sigma_s} \tag{2.28}$$

It is also possible to contract tensor $((r, s) \mapsto (r-1, s-1))$: for example, given a rank $(2, 1)$ tensor, a rank $(1, 0)$ tensor can be defined as $S(\omega) = T(\omega, \mathbf{f}^\mu, \mathbf{e}_\mu)$, with components $S^\mu = T^{\nu\mu}{}_\mu$; it must be noted that, in general, $T^{\nu\mu}{}_\mu \neq T^{\mu\nu}{}_\mu$.

Definition 2.4.10. Given an object $T_{\mu_1 \ldots \mu_n}$ dependent on some indices, its *symmetric* and *antisymmetric* parts are respectively defined as:

$$T_{(\mu_1 \ldots \mu_n)} := \frac{1}{n!} \sum_{\sigma \in S^n} T_{\sigma(\mu_1) \ldots \sigma(\mu_n)} \tag{2.29}$$

$$T_{[\mu_1 \ldots \mu_n]} := \frac{1}{n!} \sum_{\sigma \in S^n} \operatorname{sgn}(\sigma) T_{\sigma(\mu_1) \ldots \sigma(\mu_n)} \tag{2.30}$$

Conventionally, indices surrounded by || are not (anti-)symmetrized (ex: $T_{[\mu|\nu|\rho]} = \frac{1}{2}(T_{\mu\nu\rho} - T_{\rho\nu\mu})$). As previously seen, vector fields are pushed forward and 1-form are pulled back: tensors will thus behave in a mixed way.

Definition 2.4.11. Given a diffeomorphism between two manifolds $\varphi : \mathcal{M} \to \mathcal{N}$ and a (r, s) tensor field T on $\mathcal{M}$, the *push-forward* of T is the (r, s) tensor field $\varphi_* T$ on $\mathcal{N}$ such that, for $\omega_j \in \Lambda^1(\mathcal{N})$ and $X_j \in \mathfrak{X}(\mathcal{N})$:

$$\varphi_* T(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s) = T(\varphi^* \omega_1, \ldots, \varphi^* \omega_r, \varphi_*^{-1} X_1, \ldots, \varphi_*^{-1} X_s) \tag{2.31}$$

Definition 2.4.12. Given a vector field $X \in \mathfrak{X}(\mathcal{M})$ and a (r, s) tensor field T on $\mathcal{M}$, the *Lie derivative* of T along X is defined as:

$$\mathcal{L}_X T_p := \lim_{t \to 0} \frac{((\sigma_{-t})_* T)_p - T_p}{t} \tag{2.32}$$

where σ_t is the flow generated by X.

2.5 Differential forms

Definition 2.5.1. A totally anti-symmetric $(0, p)$ tensor is defined as a *p-form*. The set of all p-forms over a manifold $\mathcal{M}$ is denoted as $\Lambda^p(\mathcal{M})$.

Proposition 2.5.1. *A p-form has* $\binom{n}{p}$ *independent components.*

Proposition 2.5.2. *The maximum degree of differential forms is* $p = n \equiv \dim_{\mathbb{R}} \mathcal{M}$*: forms in* $\Lambda^n(\mathcal{M})$ *are called top forms.*

Definition 2.5.2. Given $\omega \in \Lambda^p(\mathcal{M}), \eta \in \Lambda^q(\mathcal{M})$, their *wedge product* is a $(p+q)$-form defined as:

$$(\omega \wedge \eta)_{\mu_1 \ldots \mu_p \nu_1 \ldots \nu_q} = \frac{(p+q)!}{p!q!} \omega_{[\mu_1 \ldots \mu_p} \eta_{\nu_1 \ldots \nu_q]} \tag{2.33}$$

Example 2.5.1. Given $\omega, \eta \in \Lambda^2(\mathcal{M})$, their wedge product is $(\omega \wedge \eta)_{\mu\nu} = \omega_\mu \eta_\nu - \omega_\nu \eta_\mu$.

Proposition 2.5.3. *Given* $\omega \in \Lambda^p(\mathcal{M}), \eta \in \Lambda^q(\mathcal{M})$*:*

$$\omega \wedge \eta = (-1)^{pq} \eta \wedge \omega \tag{2.34}$$

Corollary 2.5.3.1. $\omega \wedge \omega = 0 \ \forall \omega \in \Lambda^p(\mathcal{M}) : p$ *is odd.*

Proposition 2.5.4. *The wedge product is associative.*

Proposition 2.5.5. *If* $\{\mathbf{f}^\mu\}_{\mu=1,\ldots,n}$ *is a basis of* $\Lambda^1(\mathcal{M})$*, then* $\{\mathbf{f}^{\mu_1} \wedge \cdots \wedge \mathbf{f}^{\mu_p}\}_{\mu_1,\ldots,\mu_p=1,\ldots,n}$ *is a basis of* $\Lambda^p(\mathcal{M})$*.*

Locally $\{dx^\mu\}_{\mu=1,\ldots,n}$ is a basis of $T_p^*\mathcal{M}$, thus a general p-form can be locally written as:

$$\omega = \frac{1}{p!} \omega_{\mu_1 \ldots \mu_p} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_p} \tag{2.35}$$

Definition 2.5.3. The *exterior derivative* is a map $d : \Lambda^p(\mathcal{M}) \to \Lambda^{p+1}(\mathcal{M})$ defined as:

$$(d\omega)_{\mu_1 \ldots \mu_{p+1}} = (p+1) \partial_{[\mu_1} \omega_{\mu_2 \ldots \mu_{p+1}]} \tag{2.36}$$

In local coordinates:

$$d\omega = \frac{1}{p!} \frac{\partial \omega_{\mu_1 \ldots \mu_p}}{\partial x^\nu} dx^\nu \wedge dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_p} \tag{2.37}$$

Theorem 2.5.1 (Poincaré)**.** $d^2 = 0$.

Proof. Consequence of Schwarz lemma. □

Proposition 2.5.6. *Given* $\omega \in \Lambda^p(\mathcal{M}), \eta \in \Lambda^q(\mathcal{M})$*, then:*

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^p \omega \wedge d\eta \tag{2.38}$$

Proposition 2.5.7. *Given a diffeomorphism between two manifolds* $\varphi : \mathcal{M} \to \mathcal{N}$ *and* $\omega \in \Lambda^p(\mathcal{M})$*, then* $d(\varphi^* \omega) = \varphi^*(d\omega)$*.*

Corollary 2.5.7.1. *Given* $X \in \mathfrak{X}(\mathcal{M}), \omega \in \Lambda^p(\mathcal{M})$*, then* $d(\mathcal{L}_X \omega) = \mathcal{L}_X(d\omega)$*.*

Definition 2.5.4. $\omega \in \Lambda^p(\mathcal{M})$ is *closed* if $d\omega = 0$.

Definition 2.5.5. $\omega \in \Lambda^p(\mathcal{M})$ is *exact* if $\exists \eta \in \Lambda^{p-1}(\mathcal{M}) : \omega = d\eta$.

Theorem 2.5.2. *$\omega \in \Lambda^p(\mathcal{M})$ is exact $\Rightarrow$ it is closed.*

Definition 2.5.6. Given a vector field $X \in \mathfrak{X}(\mathcal{M})$, the *interior product* determined by X is a map $\iota_X : \Lambda^p(\mathcal{M}) \to \Lambda^{p-1}(\mathcal{M})$ defined as:

$$\iota_X\omega(Y_1, \dots, Y_{p-1}) := \omega(X, Y_1, \dots, Y_{p-1}) \tag{2.39}$$

On 0-forms (i.e. scalar functions), it is defined as $\iota_X f \equiv 0$.

Proposition 2.5.8. *Given $X, Y \in \mathfrak{X}(\mathcal{M})$, then $\iota_X\iota_Y = -\iota_Y\iota_X$.*

Proof. Consequence of the total anti-symmetry of p-forms. □

Proposition 2.5.9. *Given $X \in \mathfrak{X}(\mathcal{M}), \omega \in \Lambda^p(\mathcal{M}), \eta \in \Lambda^q(\mathcal{M})$, then:*

$$\iota_X(\omega \wedge \eta) = \iota_X\omega \wedge \eta + (-1)^p\omega \wedge \iota_X\eta \tag{2.40}$$

Theorem 2.5.3 (Cartan)**.** *Given a vector field $X \in \mathfrak{X}(\mathcal{M})$, then:*

$$\mathcal{L}_X = d \circ \iota_X + \iota_X \circ d \tag{2.41}$$

Proof. Consider $\omega \in \Lambda^1(\mathcal{M})$:

$$\iota_X(d\omega) = \iota_X\frac{1}{2}\left(\partial_\mu\omega_\nu - \partial_\nu\omega_\mu\right)dx^\mu \wedge dx^\nu = X^\mu\partial_\mu\omega_\nu dx^\nu - X^\nu\partial_\mu\omega_\nu dx^\mu$$

$$d(\iota_X\omega) = d(\omega_\mu X^\mu) = X^\mu\partial_\nu\omega_\mu dx^\nu + \omega_\mu\partial_\nu X^\mu dx^\nu$$

Thus, adding these expressions and recalling Eq. 2.23:

$$(d\iota_X + \iota_X d)\omega = \left(X^\mu\partial_\mu\omega_\nu + \omega_\mu\partial_\nu X^\mu\right)dx^\nu = \mathcal{L}_X\omega$$

□

2.5.1 de Rham cohomology

While exact $\Rightarrow$ closed, the converse is not true, in general: it depends on the topological properties of the manifold.

Lemma 2.5.1 (Poincaré)**.** *If $\mathcal{M}$ is simply connected, then $\omega \in \Lambda^p(\mathcal{M})$ closed $\Rightarrow \omega$ exact.*

In general, it is always possible to choose a simply connected neighbourhood of a point $p \in \mathcal{M}$, in which every closed form is exact, but that may not always be possible globally.
It is convenient to set the notation $d_p \equiv d : \Lambda^p(\mathcal{M}) \to \Lambda^{p+1}(\mathcal{M})$.

Definition 2.5.7. The set of all closed p-forms on $\mathcal{M}$ is denoted by $Z^p(\mathcal{M}) := \ker d_p$.

Definition 2.5.8. The set of all exact p-forms on $\mathcal{M}$ is denoted by $B^p(\mathcal{M}) := \operatorname{ran} d_{p-1}$.

Definition 2.5.9. Two closed p-forms $\omega, \omega' \in Z^P(\mathcal{M})$ are said to be *equivalent* if $\omega = \omega' + \eta$ for some $\eta \in B^p(\mathcal{M})$.

Definition 2.5.10. The p^{th} *de Rham cohomolgy group* of a manifold $\mathcal{M}$ is defined to be:

$$H^p(\mathcal{M}) \coloneqq Z^p(\mathcal{M})/B^p(\mathcal{M}) \tag{2.42}$$

Definition 2.5.11. The *Betti numbers* of a manifold $\mathcal{M}$ are defined as:

$$B_p \coloneqq \dim_{\mathbb{R}} H^p(\mathcal{M}) \tag{2.43}$$

Theorem 2.5.4. *Given a differentiable manifold, its Betti numbers are always finite.*

$B_0 = 1$ for any connected manifold: there exist constant functions, which are manifestly closed and not exact, due to the non-existence of "-1-forms". Higher Betti numbers are non-zero only if the manifold has some non-trivial topology.

Definition 2.5.12. The *Euler's character* of a manifold $\mathcal{M}$ is defined as:

$$\chi(\mathcal{M}) \coloneqq \sum_{p \in \mathbb{N}_0} (-1)^p B_p \tag{2.44}$$

Example 2.5.2. The n-sphere $\mathbb{S}^n$ has only $B_0 = B_n = 1$, thus $\chi(\mathbb{S}^n) = 1 + (-1)^n$.

Example 2.5.3. The n-torus $\mathbb{T}^n$ has $B_p = \binom{n}{p}$, thus $\chi(\mathbb{T}^n) = 0$.

2.5.2 Integration

Definition 2.5.13. A *volume form* on an n-dimensional differentiable manifold $\mathcal{M}$ is a nowhere-venishing top form v, i.e. locally $v = v(x) dx^1 \wedge \cdots \wedge dx^n : v(x) \neq 0$. If such a form exists, the manifold is said to be *orientable*.

Definition 2.5.14. Given an orientable manifold $\mathcal{M}$ with volume form v, the orientation is:

- right-handed if $v(x) > 0$ locally on every neighbourhood of $\mathcal{M}$;
- left-handed if $v(x) < 0$ locally on every neighbourhood of $\mathcal{M}$;

To ensure that the handedness of the manifold doesn't change on overlapping charts:

$$v = v(x) \frac{\partial x^1}{\partial y^{\mu_1}} dy^{\mu_1} \wedge \cdots \wedge \frac{\partial x^n}{\partial x^{\mu_n}} dx^{\mu_n} = v(x) \det \left(\frac{\partial x^\mu}{\partial y^\nu} \right) dy^1 \wedge \cdots \wedge dy^n$$

It is therefore necessary that the two sets of coordinates on the overlapping region satisfy:

$$\det \left(\frac{\partial x^\mu}{\partial y^\nu} \right) > 0 \tag{2.45}$$

Non-orientable manifolds cannot be covered by overlapping charts satisfying this condition.

Example 2.5.4. The real projective space $\mathbb{RP}^n$ is orientable for odd n and non-orientable for even n.

Example 2.5.5. The complex projective space $\mathbb{CP}^n$ is orientable for all $n \in \mathbb{N}$.

Definition 2.5.15. Given a function $f : \mathcal{M} \to \mathbb{R}$ on an orientable manifold $\mathcal{M}$ with volume form v and a chart (φ, U) on $\mathcal{M}$ with coordinates $\{x^\mu\}_{\mu=1,\dots,n}$, the *integral* of f on $O = \varphi^{-1}(U) \subset \mathcal{M}$ is defined as:

$$\int_O fv := \int_U dx_1 \dots dx_n \, f(x)v(x) \tag{2.46}$$

It is clear that the volume form acts like a measure on the manifold. To integrate over the whole manifold, it must be divided up into different regions, each covered by a single chart.

Definition 2.5.16. A k-dimensional manifold Σ is a *submanifold* of an n-dimensionale manifold $\mathcal{M}$, with $n > k$, if there exists an injective map $\varphi : \Sigma \to \mathcal{M}$ such that $\varphi_* : T_p(\Sigma) \to T_{\varphi(p)}(\mathcal{M})$ is injective.

Definition 2.5.17. Given a k-form $\omega \in \Lambda^k(\mathcal{M})$, its integral over a k-dimensional submanifold Σ of $\mathcal{M}$ is defined as:

$$\int_{\varphi(\Sigma)} \omega := \int_\Sigma \varphi^* \omega \tag{2.47}$$

Example 2.5.6. Consider a 1-form $\omega \in \Lambda^1(\mathcal{M})$ and a 1-dimensional submanifold γ of $\mathcal{M}$ described by a curve $\sigma : \gamma \to \mathcal{M} : x^\mu = \sigma^\mu(t)$: locally $\omega = \omega_\mu(x)dx^\mu$, thus the integral of ω on γ can be calculated as $\int_{\sigma(\gamma)} \omega = \int_\gamma \sigma^* \omega = \int_\gamma d\tau \, \omega_\mu(x) \frac{dx^\mu}{d\tau}$.

2.5.2.1 Stokes' theorem

Integration can be generalized beyond smooth (i.e. differentiable) manifolds.

Definition 2.5.18. An n-dimensional *manifold with boundary* is a Hausdorff topological space, equipped with a compatible maximal atlas, which is locally homeomorphic to $\mathbb{R}^{n-1} \times [a, \infty) : a \in \mathbb{R}$. The *boundary* $\partial\mathcal{M}$ is the 1-dimensional submanifold determined by $x^n = a$.

Theorem 2.5.5 (Stokes)**.** *Given an n-dimensional manifold $\mathcal{M}$ with boundary $\partial\mathcal{M}$, then for any $\omega \in \Lambda^{n-1}(\mathcal{M})$:*

$$\int_\mathcal{M} d\omega = \int_{\partial\mathcal{M}} \omega \tag{2.48}$$

This important theorem unifies many different results.
Given the 1-dimensional manifold $I = [a, b] \subset \mathbb{R}$, then for any 0-form (i.e. scalar function) $\omega = \omega(x)$:

$$\int_I d\omega = \int_a^b \frac{d\omega}{dx} dx = \int_{\partial I} \omega = \omega(b) - \omega(a)$$

which is the fundamental theorem of calculus.
Given a 2-dimensional manifold with boundary $S \subset \mathbb{R}^2$ and a 1-form $\omega = \omega_1 dx^1 + \omega_2 dx^2$, then $d\omega = (\partial_1\omega_2 - \partial_2\omega_1)dx^1 \wedge dx^2$ and:

$$\int_S d\omega = \int_S \left(\frac{\partial\omega_2}{\partial x^1} - \frac{\partial\omega_1}{\partial x^2} \right) dx^1 dx^2 = \int_{\partial S} \omega = \int_{\partial S} \omega_1 dx^1 + \omega_2 dx^2$$

which is Green's theorem.
GIven a 3-dimensional manifold with boundary $V \subset \mathbb{R}^3$ and a 2-form $\omega = \omega_1 dx^2 \wedge dx^3 + \omega_2 dx^3 \wedge dx^1 + \omega_3 dx^1 \wedge dx^2$, then $d\omega = (\partial_1\omega_1 + \partial_2\omega_2 + \partial_3\omega_3)dx^1 \wedge dx^2 \wedge dx^3$ and:

$$\int_V d\omega = \int_V \left(\frac{\partial\omega_1}{\partial x^1} + \frac{\partial\omega_2}{\partial x^2} + \frac{\partial\omega_3}{\partial x^3} \right) = \int_{\partial V} \omega = \int_{\partial V} \omega_1 dx^2 dx^3 + \omega_2 dx^3 dx^1 + \omega_3 dx^1 dx^2$$

which is Gauss' theorem.

Chapter 3

Riemannian Geometry

3.1 Metric manifolds

Definition 3.1.1. A *metric* g is a (0,2) tensor field on a manifold $\mathcal{M}$ that is:

1. symmetric: $g(X, Y) = g(Y, X)$;
2. non-degenerate: $\exists p \in \mathcal{M} : g(X, Y)|_p = 0 \, \forall Y \in T_p\mathcal{M} \Rightarrow X_p = 0$.

Definition 3.1.2. A *metric manifold* $(\mathcal{M}, g)$ is a manifold equipped with a metric.

With a choice of coordinates, the metric can be written as:

$$g = g_{\mu\nu}(x) dx^\mu \otimes dx^\nu \tag{3.1}$$

where:

$$g_{\mu\nu} = g\left(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right) \tag{3.2}$$

It is often written also as $ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu$. The matrix $g_{\mu\nu}(x) \in \mathbb{R}^{n\times n}$ is symmetric, and there's always a choice of basis on each tangent space such that this matrix is diagonal: the non-degeneracy condition implies that none of the diagonal elements vanish.

Proposition 3.1.1. *The signature of a metric, i.e. the number of negative entries when diagonalized, is independent on the choice of basis.*

Proof. From Sylvester's theorem of inertia. □

Riemannian manifolds

Definition 3.1.3. A *Riemannian manifold* $(\mathcal{M}, g)$ is a manifold equipped with a metric with totally-positive signature.

Example 3.1.1. The Euclidean space $\mathbb{R}^n$, equipped with the metric $g_{\mu\nu} = \delta_{\mu\nu}$ (in Cartesian coordinates), is a Riemannian manifold.

Definition 3.1.4. Given a Riemannian manifold $(\mathcal{M}, g)$ and $X \in \mathfrak{X}(\mathcal{M})$, the *length* of X at $p \in \mathcal{M}$ is:

$$|X_p| := \sqrt{g(X, X)|_p} \tag{3.3}$$

Given $Y \in \mathfrak{X}(\mathcal{M})$, the *angle* between X and Y at $p \in \mathcal{M}$ is:

$$\cos\theta := \frac{g(X, Y)|_p}{|X_p|\,|Y_p|} \tag{3.4}$$

This can be generalized to distances between points on a curve $\sigma : \mathbb{R} \to \mathcal{M}$:

$$d(p,q) = \int_a^b dt \sqrt{g(X,X)|_{\sigma(t)}} \tag{3.5}$$

where $\sigma(a) = p$, $\sigma(b) = q$ and X is the tangent vector field of the curve. With parametrization $x^\mu(t)$, the tangent vector has components $X^\mu = \frac{dx^\mu}{dt}$, thus:

$$d(p,q) = \int_a^b dt \sqrt{g_{\mu\nu}(x)\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}} \tag{3.6}$$

It is important to note that this distance is independent of the parametrization.

Lorentzian manifolds

Definition 3.1.5. A *Lorentzian manifold* $(\mathcal{M}, g)$ is a manifold equipped with a metric which has a signature with a single negative sign.

Example 3.1.2. The simplest Lorentzian manifold is $\mathbb{R}^n$ with the *Minkowski metric*:

$$\eta = -dx^0 \otimes dx^0 + dx^1 \otimes dx^1 + \cdots + dx^{n-1} \otimes dx^{n-1} \tag{3.7}$$

Its components are $\eta_{\mu\nu} = \mathrm{diag}(-1, +1, \ldots, +1)$, thus this is a Lorentzian manifold.

On a general Lorentzian manifold, at any point $p \in \mathcal{M}$ it is always possible to choose an orthonormal basis $\{e_\mu\}_{\mu=0,\ldots,n-1}$ of $T_p(\mathcal{M})$ such that $g_{\mu\nu}|_p = \eta_{\mu\nu}$: this fact is closely related to the equivalence principle. Consider a different basis $\tilde{e}_\mu = \Lambda^\nu{}_\mu e_\nu$: the condition for it to leave the Minkowski metric unchanged is:

$$\eta_{\mu\nu} = \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu \eta_{\rho\sigma} \tag{3.8}$$

This is the defining equation of a Lorentz transformation: on a Lorentzian manifold, the basic features of special relativity are locally recovered. Thus, other ideas from special relativity can be imported.

Definition 3.1.6. Given a Lorentzian manifold $(\mathcal{M}, g)$ and $X \in \mathfrak{X}(\mathcal{M})$, at $p \in \mathcal{M}$ the vector field is said to be:

- *timelike* if $g(X_p, X_p) < 0$;
- *null* if $g(X_p, X_p) = 0$;
- *spacelike* if $g(X_p, X_p) > 0$.

At each point $p \in \mathcal{M}$ it is possible to draw *lightcones*, i.e. the null tangent vectors at that point, which are past-directed or future-directed: these lightcones vary smoothly as the point is varied smoothly on the manifold, elucidating the causal structure of spacetime.

The distance between two points on a curve depends on the nature of the tangent vector field of the curve: a *timelike curve* is a curve whose tangent vector field is everywhere timelike, and analogously for the other cases. The distance on a spacelike curve is defined as in Eq. 3.5, while that on a timelike curve gets a negative sign in the square root. With parametrization $x^\mu(t)$, it is possible to define the *proper time* on a timelike curve as:

$$\tau = \int_a^b dt \sqrt{-g_{\mu\nu}(x)\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}} \tag{3.9}$$

This is precisely the action of a free particle moving in spacetime.

3.1.1 Metric properties

The metric defines a natural isomorphism between vectors and covectors.

Proposition 3.1.2. *Given a metric manifold* $(\mathcal{M}, g)$, *the metric defines for each* $p \in \mathcal{M}$ *a natural isomorphism* $g : X_p \in T_p(\mathcal{M}) \to \omega_p \in T_p^*(\mathcal{M}) : \omega_p(Y_p) = g(X_p, Y_p)\, \forall Y_p \in T_p(\mathcal{M})$.

In a chosen coordinate basis, the vector $X = X^\mu \partial_\mu$ is mapped to the one-form $X = X_\mu dx^\mu$, thus the following identity holds:

$$X_\mu = g_{\mu\nu} X^\nu \tag{3.10}$$

Being g non-degenerate, the matrix $g_{\mu\nu}$ is invertible, with inverse $g^{\mu\nu}$ such that:

$$g^{\mu\nu} g_{\nu\rho} = \delta^\mu_\rho \tag{3.11}$$

Its elements are the components of a (2,0) symmetric tensor $\hat{g} := g^{\mu\nu} \partial_\mu \otimes \partial_\nu$, which defines the inverse of the natural isomorphism in Prop. 3.1.2:

$$X^\mu = g^{\mu\nu} X_\nu \tag{3.12}$$

The metric also defines a natural volume form on the manifold.

Definition 3.1.7. Given an n-dimensional metric manifold $(\mathcal{M}, g)$, the *volume form* is the top-form:

$$v := \sqrt{\mathrm{g}}\, dx^1 \wedge \cdots \wedge dx^n \tag{3.13}$$

where $\mathrm{g} := |\det g_{\mu\nu}|$.

Proposition 3.1.3. *The volume form is basis-independent.*

Proof. Consider a new set of coordinates y^μ such that $dx^\mu = A^\mu{}_\nu dy^\nu$, where $A^\mu{}_\nu = \frac{\partial x^\mu}{\partial y^\nu}$. In general:

$$dx^1 \wedge \cdots \wedge dx^n = A^1{}_{\mu_1} \dots A^n{}_{\mu_n} dy^{\mu_1} \wedge \cdots \wedge dy^{\mu_n}$$

Recalling the anti-symmetry of the wedge product and the definition of determinant, this can be rewritten as:

$$dx^1 \wedge \cdots \wedge dx^n = \sum_{\pi \in S^n} \operatorname{sgn} \pi\, A^1{}_{\pi(1)} \dots A^n{}_{\pi(n)} dy^1 \wedge \cdots \wedge dy^n = \det A\, dy^1 \wedge \cdots \wedge dy^n$$

Note the Jacobian factor which arises when changing the measure. On the other hand:

$$g_{\mu\nu} = \frac{\partial y^\rho}{\partial x^\mu} \frac{\partial y^\sigma}{\partial x^\nu} \tilde{g}_{\rho\sigma} = (A^{-1})^\rho{}_\mu (A^{-1})^\sigma{}_\nu \tilde{g}_{\rho\sigma} \quad \Rightarrow \quad \det g_{\mu\nu} = \frac{\det \tilde{g}_{\mu\nu}}{(\det A)^2}$$

The factors $\det A$ and $(\det A)^{-1}$ cancel, thus yielding the thesis. □

The volume form can be rewritten as:

$$v = \frac{1}{n!} v_{\mu_1 \dots \mu_n} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_n} \equiv \frac{1}{n!} \sqrt{\mathrm{g}}\, \epsilon_{\mu_1 \dots \mu_n} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_n} \tag{3.14}$$

where $\epsilon_{\mu_1 \dots \mu_n}$ is the totally-antisymmetric n-dimensional symbol (generalization of the Levi-Civita symbol). $\epsilon_{\mu_1 \dots \mu_n}$ cannot be considered a proper tensor, as its components are always $+1, -1, 0$

indipendently if the indices are covariant or contravariant: it is, in fact, a *tensor density*, i.e. a tensor divided by $\sqrt{\mathrm{g}}$. It can be shown that:

$$v^{\mu_1 \dots \mu_n} = g^{\mu_1 \nu_1} \dots g^{\mu_n \nu_n} v_{\mu_1 \dots \mu_n} = \sigma \frac{1}{\sqrt{\mathrm{g}}} \epsilon^{\mu_1 \dots \mu_n} \tag{3.15}$$

where σ is the sign of the signature (ex.: $\sigma = +1$ for Riemannian manifolds and $\sigma = -1$ for Lorentzian manifolds). As notation, the integral of a generic function f on $\mathcal{M}$ is denoted as:

$$\int_{\mathcal{M}} f v \equiv \int_{\mathcal{M}} d^n x \sqrt{\mathrm{g}} f \tag{3.16}$$

3.1.1.1 Hodge theory

Definition 3.1.8. Given an n-dimensional oriented metric manifold $(\mathcal{M}, g)$, the *Hodge dual* is defined as the map $\star : \Lambda^p(\mathcal{M}) \to \Lambda^{n-p}(\mathcal{M}) : \omega \mapsto \star \omega$ such that:

$$\star \omega_{\mu_1 \dots \mu_{n-p}} := \frac{1}{(n-p)!} \sqrt{\mathrm{g}}\, \epsilon_{\mu_1 \dots \mu_{n-p} \nu_1 \dots \nu_p} \omega^{\nu_1 \dots \nu_p} \tag{3.17}$$

In this section, the orientedness and n-dimensionality of the manifold are implied.

Proposition 3.1.4. *The Hodge dual is basis-independent.*

It is useful to state a lemma for future calculations.

Lemma 3.1.1. $v^{\mu_1 \dots \mu_p \rho_1 \dots \rho_{n-p}} v_{\nu_1 \dots \nu_p \rho_1 \dots \rho_{n-p}} = \sigma p!(n-p)! \delta^{\mu_1}_{[\nu_1} \dots \delta^{\mu_p}_{\nu_p]}$.

Proposition 3.1.5. $\star (\star \omega) = \sigma \, (-1)^{p(n-p)} \, \omega$.

The Hodge dual defines an inner product on each $\Lambda^p(\mathcal{M})$:

$$\langle \omega, \eta \rangle := \int_{\mathcal{M}} \omega \wedge \star \eta \tag{3.18}$$

This allows to define operators and their adjoints on the form spaces.

Proposition 3.1.6. *Given a metric manifold $(\mathcal{M}, g)$ and two forms $\omega \in \Lambda^p(\mathcal{M})$, $\alpha \in \Lambda^{p-1}(\mathcal{M})$, then:*

$$\langle d\alpha, \omega \rangle = \langle \alpha, d^\dagger \omega \rangle \tag{3.19}$$

where the adjoint of the exterior derivative $d^\dagger : \Lambda^p(\mathcal{M}) \to \Lambda^{p-1}(\mathcal{M})$ is defined as:

$$d^\dagger := \sigma \, (-1)^{np+n-1} \star d \star \tag{3.20}$$

Proof. To simplify the proof, consider a closed manifold; then, from Stokes' theorem and Eq. 2.38:

$$0 = \int_{\mathcal{M}} d(\alpha \wedge \star \omega) = \langle d\alpha, \omega \rangle + \int_{\mathcal{M}} (-1)^{p-1} \, \alpha \wedge d \star \omega$$

The second term is proportional to $\langle \alpha, \star d \star \omega \rangle$: to determine the relative sign, note that $d \star \omega \in \Lambda^{n-p+1}(\mathcal{M})$, thus, from Prop. 3.1.5, $\star \star d \star \omega = \sigma \, (-1)^{(n-p+1)(p-1)} \, d \star \omega$. In conclusion:

$$\langle \alpha, \star d \star \omega \rangle = \sigma \, (-1)^{(n-p)(p-1)} \int_{\mathcal{M}} (-1)^{p-1} \, \alpha \wedge d \star \omega \quad \Rightarrow \quad \langle d\alpha, \omega \rangle = \sigma \, (-1)^{(n-p)(p-1)+1} \langle \alpha, \star d \star \omega \rangle$$

Noting that $(-1)^{(n-p)(p-1)+1} = (-1)^{np+n-1}$, as in general $(-1)^{-n} = (-1)^n$ and $(-1)^{-p^2+p+1} = (-1)^{-1}$ due to $p(p-1)$ being always even, concludes the proof. $\square$

Definition 3.1.9. Given a metric manifold $(\mathcal{M}, g)$, the *Laplacian* $\triangle : \Lambda^p(\mathcal{M}) \to \Lambda^p(\mathcal{M})$ is defined as the operator:

$$\triangle := (d + d^\dagger)^2 \tag{3.21}$$

Proposition 3.1.7. $\triangle = dd^\dagger + d^\dagger d = \{d, d^\dagger\}$.

Proof. Trivial, given $d^2 = d^{\dagger 2} = 0$. □

It is possible to calculate an explicit expression for the Laplacian of functions.

Lemma 3.1.2. *Given $f \in \mathcal{C}^\infty(\mathcal{M})$, then $d^\dagger f = 0$.*

Proof. Trivial noting that $\star f$ is a top-form. □

Proposition 3.1.8. *Given $f \in \mathcal{C}^\infty(\mathcal{M})$, then:*

$$\triangle f = -\frac{\sigma}{\sqrt{\mathrm{g}}}\partial_\nu\left(\sqrt{\mathrm{g}}\, g^{\mu\nu}\partial_\mu f\right) \tag{3.22}$$

Proof. Via direct calculation, using Lemma 3.1.2:

$$\begin{aligned}
\triangle f &= \sigma\,(-1)^{n^2+n-1} \star d \star (\partial_\mu f dx^\mu) = -\sigma \star d\,(\partial_\mu f \star dx^\mu) \\
&= -\frac{\sigma}{(n-1)!} \star d\left(\partial_\mu f g^{\mu\nu} \sqrt{\mathrm{g}}\,\epsilon_{\nu\rho_1\ldots\rho_{n-1}} dx^{\rho_1} \wedge \cdots \wedge dx^{\rho_{n-1}}\right) \\
&= -\frac{\sigma}{(n-1)!} \star \partial_\alpha\left(\sqrt{\mathrm{g}}\, g^{\mu\nu}\partial_\mu f\right) \epsilon_{\nu\rho_1\ldots\rho_{n-1}} dx^\alpha \wedge dx^{\rho_1} \wedge \cdots \wedge dx^{\rho_{n-1}} \\
&= -\sigma \star \partial_\nu\left(\sqrt{\mathrm{g}}\, g^{\mu\nu}\partial_\mu f\right) dx^1 \wedge \ldots dx^n = -\frac{\sigma}{\sqrt{\mathrm{g}}}\partial_\nu\left(\sqrt{\mathrm{g}}\, g^{\mu\nu}\partial_\mu f\right)
\end{aligned}$$

□

The Laplacian operator is linked to the de Rham cohomology.

Definition 3.1.10. Given $\omega \in \Lambda^p(\mathcal{M})$, it is said to be *harmonic* if $\triangle\omega = 0$.

Definition 3.1.11. The space of harmonic p-forms on $(\mathcal{M}, g)$ is denoted as $\mathrm{Harm}^p(\mathcal{M})$.

Proposition 3.1.9. *A harmonic form is both closed and co-closed.*

Proof. $0 = \langle\omega, \triangle\omega\rangle = \langle d\omega, d\omega\rangle + \langle d^\dagger\omega, d^\dagger\omega\rangle$, thus $d\omega = 0$ and $d^\dagger\omega = 0$, for the inner product is positive-defined. □

Theorem 3.1.1. *Given a compact Riemannian manifold $(\mathcal{M}, g)$, any $\omega \in \Lambda^p(\mathcal{M})$ can be uniquely decomposed as $\omega = d\alpha + d^\dagger\beta + \gamma$, with $\alpha \in \Lambda^{p-1}(\mathcal{M})$, $\beta \in \Lambda^{p+1}(\mathcal{M})$ and $\gamma \in \mathrm{Harm}^p(\mathcal{M})$.*

Theorem 3.1.2 (Hodge)**.** *Given a compact Riemannian manifold $(\mathcal{M}, g)$, there is an isomorphism:*

$$\mathrm{Harm}^p(\mathcal{M}) \cong H^p(\mathcal{M}) \tag{3.23}$$

Proof. From Prop. 3.1.9 $\mathrm{Harm}^p(\mathcal{M}) \subset Z^p(\mathcal{M})$, but the uniqueness of decomposition in Th. 3.1.1 implies $\forall\gamma \in \mathrm{Harm}^p(\mathcal{M})\, \exists\eta_\gamma \in \Lambda^{p-1}(\mathcal{M}) : \gamma \neq d\eta_\gamma$, thus $\mathrm{Harm}^p(\mathcal{M}) \subset H^p(\mathcal{M})$.
WTS that any equivalence class $[\omega] \in H^p(\mathcal{M})$ can be represented by a harmonic form. By Th, 3.1.1 $\omega = d\alpha + d^\dagger\beta + \gamma$, but $\omega \in H^p(\mathcal{M})$ implies $d\omega = 0$ by definition, so:

$$0 = \langle d\omega, \beta\rangle = \langle\omega, d^\dagger\beta\rangle = \langle d\alpha + d^\dagger\beta + \gamma, d^\dagger\beta\rangle = \langle d^\dagger\beta, d^\dagger\beta\rangle$$

The inner product is positive-definite, thus $d^\dagger\beta = 0$, hence $\omega = \gamma + d\alpha$. By definition $H^p(\mathcal{M}) := Z^p(\mathcal{M})/B^p(\mathcal{M})$, so $[\omega] = \gamma$. □

Corollary 3.1.2.1. $B_p = \dim_{\mathbb{R}} \mathrm{Harm}^p(\mathcal{M})$.

3.2 Connections

There's a different way to differentiate tensor fields distinct from the Lie derivative, associated to a different way to map different vector spaces at different points: the covariant derivative.
From now on, $\mathcal{M}$ is implied to be an n-dimensional metric manifold with metric g.

3.2.1 Covariant derivative

Definition 3.2.1. The *connection* is a map $\nabla : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$, usually written as $\nabla(X, Y) \equiv \nabla_X Y$, where ∇_X is called the *covariant derivative*, satisfying the following properties for all $X, Y, Z \in \mathfrak{X}(\mathcal{M})$:

1. $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$;
2. $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z \, \forall f, g \in \mathcal{C}^\infty(\mathcal{M})$;
3. $\nabla_X(fY) = f\nabla_X Y + X(f)Y \, \forall f \in \mathcal{C}^\infty(\mathcal{M})$.

Usually $X(f) \equiv \nabla_X f$. The covariant derivative endows the manifold with more structure: in particular, given a basis $\{e_\mu\}$ of $\mathfrak{X}(\mathcal{M})$, its covariant derivative is expressed as:

$$\nabla_{e_\rho} e_\mu \equiv \Gamma^\mu_{\rho\nu} e_\mu \tag{3.24}$$

The $\Gamma^\mu_{\rho\nu}$ are the components of the connection on that basis. Usually $\nabla_{e_\mu} \equiv \nabla_\mu$, thus resembling a partial derivative. To elucidate how the covariant derivative acts on vector fields:

$$\begin{aligned}
\nabla_X Y &= \nabla_X(Y^\mu e_\mu) \\
&= X(Y^\mu) e_\mu + Y^\mu \nabla_X e_\mu \\
&= X^\nu e_\nu(Y^\mu) e_\mu + Y^\mu X^\nu \nabla_\nu e_\mu \\
&= X^\nu \left[e_\nu(Y^\mu) + \Gamma^\mu_{\nu\rho} Y^\rho \right] e_\mu \\
&= X^\nu \nabla_\nu Y = X^\nu (\nabla_\nu Y)^\mu e_\mu
\end{aligned}$$

The dependency on X can therefore be eliminated, and in components:

$$(\nabla_\nu Y)^\mu = e_\nu(Y^\mu) + \Gamma^\mu_{\nu\rho} Y^\rho \tag{3.25}$$

A sloppy notation is often used: $(\nabla_\nu Y)^\mu \equiv \nabla_\nu Y^\mu$. This must not be confused as the covariant derivative of Y^μ. Moreover $\nabla_\nu Y^\mu \equiv Y^\mu_{\ ;\nu}$, while $\partial_\mu f \equiv f_{,\mu}$. On the coordinate basis $e_\mu = \partial_\mu$, then:

$$Y^\mu_{\ ;\nu} = Y^\mu_{\ ,\nu} + \Gamma^\mu_{\nu\rho} Y^\rho \tag{3.26}$$

Note that $Y^\mu_{\ ;\nu}$ is the μ^{th} component of $\nabla_\nu Y$, while $Y^\mu_{\ ,\nu}$ is the partial derivative of Y^μ along ∂_ν.
The covariant derivative coincides with other derivatives on $\mathcal{C}^\infty(\mathcal{M})$: it can be shown that $\nabla_X f = \mathcal{L}_X f = X(f)$ and $\nabla_\mu f = \partial_\mu f$. On $\mathfrak{X}(\mathcal{M})$, however, ∇_X and $\mathcal{L}_X$ are distinct: while $\nabla_X = X^\mu \nabla_\mu$, there's no way to write the same relation for $\mathcal{L}_X$, for it depends not only on X but on its first derivative too. The covariant derivative is thus the natural generalization of the partial derivative to curved manifolds.

Proposition 3.2.1. *$\Gamma^\mu_{\rho\nu}$ are not components of a tensor.*

Proof. Given the basis transformation $\tilde{e}_\nu = A^\mu{}_\nu e_\mu$, with A and invertible matrix (if they're both coordinate basis, then $A^\mu{}_\nu = \frac{\partial x^\mu}{\partial x^\nu}$), the components of a (1,2) tensor must transform as:

$$\tilde{T}^\mu{}_{\rho\nu} = (A^{-1})^\mu{}_\tau A^\sigma{}_\rho A^\lambda{}_\nu T^\tau{}_{\sigma\lambda}$$

In the new basis:

$$\begin{aligned}\tilde{\Gamma}^\mu_{\rho\nu}\tilde{e}_\mu &= \nabla_{\tilde{e}_\rho}\tilde{e}_\nu = \nabla_{A^\sigma{}_\rho e_\sigma}(A^\lambda{}_\nu e_\nu) = A^\sigma{}_\rho \nabla_{e_\sigma}(A^\lambda{}_\nu e_\lambda) \\ &= A^\sigma{}_\rho A^\lambda{}_\nu \Gamma^\tau_{\sigma\lambda} e_\tau + A^\sigma{}_\rho e_\lambda \partial_\sigma A^\lambda{}_\nu = \left[A^\sigma{}_\rho A^\lambda{}_\nu \Gamma^\tau_{\sigma\lambda} + A^\sigma{}_\rho \partial_\sigma A^\tau{}_\nu\right] e_\tau \\ &= \left[A^\sigma{}_\rho A^\lambda{}_\nu \Gamma^\tau_{\sigma\lambda} + A^\sigma{}_\rho \partial_\sigma A^\tau{}_\nu\right](A^{-1})^\mu{}_\tau \tilde{e}_\mu\end{aligned}$$

Thus, there's a second term proportional to ∂A which deviates from the transformation law:

$$\tilde{\Gamma}^\mu_{\rho\nu} = (A^{-1})^\mu{}_\tau A^\sigma{}_\rho \left[A^\lambda{}_\nu \Gamma^\tau_{\sigma\lambda} \partial_\sigma A^\tau{}_\nu\right]$$

□

3.2.2 Covariant derivative of tensors

First of all, it is necessary to elucidate how the covariant derivative acts on one-forms. Given a one-form ω, the one-form $\nabla_X \omega$ is defined by its action on vector fields. By Leibniz rule:

$$\nabla_X(\omega(Y)) = (\nabla_X \omega)(Y) + \omega(\nabla_X Y)$$

Recalling that $\omega(Y)$ is a function, $\nabla_X(\omega(Y)) = X(\omega(Y))$, therefore:

$$(\nabla_X \omega)(Y) = X(\omega(Y)) - \omega(\nabla_X Y) \tag{3.27}$$

Expressing it in coordinates:

$$\begin{aligned}X^\mu(\nabla_\mu \omega)_\nu Y^\nu &= X^\mu \partial_\mu(\omega_\nu Y^\nu) - \omega_\nu X^\mu \left[\partial_\mu Y^\nu + \Gamma^\nu_{\mu\rho} Y^\rho\right] \\ &= X^\mu \left[\partial_\mu \omega_\rho - \Gamma^\nu_{\mu\rho}\omega_\nu\right] Y^\rho\end{aligned}$$

Crucially, the ∂Y terms cancel out, allowing to define $\nabla_X \omega$ without referencing Y:

$$(\nabla_\mu \omega)_\rho = \partial_\mu \omega_\rho - \Gamma^\nu_{\mu\rho}\omega_\nu \tag{3.28}$$

Using the same notation as for vector fields $(\nabla_\mu \omega)_\rho \equiv \nabla_\mu \omega_\rho \equiv \omega_{\rho;\mu}$:

$$\omega_{\rho;\mu} = \omega_{\rho,\mu} - \Gamma^\nu_{\mu\rho}\omega_\nu \tag{3.29}$$

This kind of argument can be extended to a general (p, q) tensor field:

$$\begin{aligned}\nabla_\rho T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\mu_q} = \partial_\rho T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q} &+ \Gamma^{\mu_1}_{\rho\sigma} T^{\sigma\mu_2\ldots\mu_p}{}_{\nu_1\ldots\nu_q} + \cdots + \Gamma^{\mu_p}_{\rho\sigma} T^{\mu_1\ldots\mu_{p-1}\sigma}{}_{\nu_1\ldots\nu_q} \\ &- \Gamma^\sigma_{\rho\nu_1} T^{\mu_1\ldots\mu_p}{}_{\sigma\nu_2\ldots\nu_q} - \cdots - \Gamma^\sigma_{\rho\nu_q} T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_{q-1}\sigma}\end{aligned} \tag{3.30}$$

The pattern is clear: for each upper index μ there's a $+\Gamma^\mu_{\rho\sigma} T^\sigma$ term, while for each lower index ν there's $-\Gamma^\sigma_{\rho\nu} T_\sigma$ term. Furthermore, it is necessary to generalize the comma-notation: for example, $X^\mu{}_{;\nu\rho} \equiv \nabla_\rho \nabla_\nu X^\mu$, so the rightmost index is the one whose covariant derivative acts first.

3.2.2.1 Torsion and curvature

Even though the connection is not a tensor, it is used to construct two important tensors.

Definition 3.2.2. The *torsion* is a (1,2) tensor defined on $\Lambda^1(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M})$ as:

$$T(\omega, X, Y) \coloneqq \omega(\nabla_X Y - \nabla_Y X - [X, Y]) \tag{3.31}$$

Alternatively, the torsion can be viewed as a map $T : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$ sucht that:

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] \tag{3.32}$$

Definition 3.2.3. The *curvature* is a (1,3) tensor defined on $\Lambda^1(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M})$ as:

$$R(\omega, X, Y, Z) \coloneqq \omega(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z) \tag{3.33}$$

Alternatively, the curvature can be viewed as a map from $\mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M})$ to the space of differential operators on $\mathfrak{X}(\mathcal{M})$ such that:

$$R(X, Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]} \tag{3.34}$$

The fact that these are indeed tensors, i.e. they are linear in each argument, can be shown by direct calculation, recalling that $[fX, Y] = f[X, Y] - Y(f)X$.

Proposition 3.2.2. *On the coordinate basis $\{\partial_\mu\}$ and $\{dx^\mu\}$ the torsion components are:*

$$T^\rho{}_{\mu\nu} = \Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\nu\mu} \tag{3.35}$$

Proof. By direct calculation:

$$\begin{aligned}
T^\rho{}_{\mu\nu} &= T(dx^\rho, \partial_\mu, \partial_\nu) = dx^\mu(\nabla_\mu \partial_\nu - \nabla_\nu \partial_\mu - [\partial_\mu, \partial_\nu]) \\
&= dx^\mu(\partial_\mu \partial_\nu - \Gamma^\sigma_{\mu\nu} \partial_\sigma - \partial_\nu \partial_\mu + \Gamma^\sigma_{\nu\mu} \partial_\sigma) \\
&= \left[\Gamma^\sigma_{\mu\nu} - \Gamma^\sigma_{\nu\mu}\right] \delta^\mu_\sigma = \Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\nu\mu}
\end{aligned}$$

□

Interestingly, even though $\Gamma^\rho_{\mu\nu}$ is not a tensor, its anti-symmetric part $\Gamma^\rho_{[\mu\nu]} = \frac{1}{2} T^\sigma{}_{\mu\nu}$ is. Clearly, the torsion tensor is anti-symmetric in its lower indices, thus for connections which are symmetric in their lower indices the torsion is null: such connections are said to be *torsion-free.*

Proposition 3.2.3. *On the coordiante basis $\{\partial_\mu\}$ and $\{dx^\mu\}$ the curvature components are:*

$$R^\sigma{}_{\rho\mu\nu} = \partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho} + \Gamma^\lambda_{\nu\rho} \Gamma^\sigma_{\mu\lambda} - \Gamma^\lambda_{\mu\rho} \Gamma^\sigma_{\nu\lambda} \tag{3.36}$$

Proposition 3.2.4. *By direct calculation:*

$$\begin{aligned}
R(dx^\sigma, \partial_\mu, \partial_\nu, \partial_\rho) &= dx^\sigma(\nabla_\mu \nabla_\nu \partial_\rho - \nabla_\nu \nabla_\mu \partial_\rho - \nabla_{[\partial_\mu, \partial_\nu]} \partial_\rho) \\
&= dx^\sigma(\nabla_\mu \nabla_\nu \partial_\rho - \nabla_\nu \nabla_\mu \partial_\rho) = dx^\sigma(\nabla_\mu(\Gamma^\lambda_{\nu\rho} \partial_\lambda) - \nabla_\nu(\Gamma^\lambda_{\mu\rho} \partial_\lambda)) \\
&= dx^\sigma((\partial_\mu \Gamma^\lambda_{\nu\rho})\partial_\lambda + \Gamma^\lambda_{\nu\rho} \Gamma^\tau_{\mu\lambda} \partial_\tau - (\partial_\nu \Gamma^\lambda_{\mu\rho})\partial_\lambda - \Gamma^\lambda_{\mu\rho} \Gamma^\tau_{\nu\lambda} \partial_\tau) \\
&= \partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho} + \Gamma^\lambda_{\nu\rho} \Gamma^\sigma_{\mu\lambda} - \Gamma^\lambda_{\mu\rho} \Gamma^\sigma_{\nu\lambda}
\end{aligned}$$

Clearly, the curvature tensor is anti-symmetric in its last two lower indices, i.e. $R^{\sigma}{}_{\rho\mu\nu} = R^{\sigma}{}_{\rho[\mu\nu]}$. It's also easy to show that:

$$R^{\sigma}{}_{\rho\mu\nu} = 2\partial_{[\mu}\Gamma^{\sigma}_{\nu]\rho} + 2\Gamma^{\sigma}_{[\mu|\lambda|}\Gamma^{\lambda}_{\nu]\rho} \tag{3.37}$$

Theorem 3.2.1. *The following idendity, known as the Ricci identity, holds:*

$$2\nabla_{[\mu}\nabla_{\nu]}Z^{\sigma} = R^{\sigma}{}_{\rho\mu\nu}Z^{\rho} - T^{\rho}{}_{\mu\nu}\nabla_{\rho}Z^{\sigma} \tag{3.38}$$

Proof. By direct calculation:

$$\begin{aligned}\nabla_{[\mu}\nabla_{\nu]}Z^{\sigma} &= \partial_{[\mu}(\nabla_{\nu]}Z^{\sigma}) + \Gamma^{\sigma}_{[\mu|\lambda|}\nabla_{\nu]}Z^{\lambda} - \Gamma^{\rho}_{[\mu\nu]}\nabla_{\rho}Z^{\sigma} \\ &= \partial_{[\mu}\partial_{\nu]}Z^{\sigma} + (\partial_{[\mu}\Gamma^{\sigma}_{\nu]\rho})Z^{\rho} + (\partial_{[\mu}Z^{\rho})\Gamma^{\sigma}_{\nu]\rho} + \Gamma^{\sigma}_{[\mu|\lambda|}\partial_{\nu]}Z\lambda + \Gamma^{\sigma}_{[\mu|\lambda|}\Gamma^{\lambda}_{\nu]\rho}Z^{\rho} - \frac{1}{2}T^{\rho}{}_{\mu\nu}\nabla_{\rho}Z^{\sigma} \\ &= \left(\partial_{[\mu}\Gamma^{\sigma}_{\nu]\rho} + \Gamma^{\sigma}_{[\mu|\lambda|}\Gamma^{\lambda}_{\nu]\rho}\right)Z^{\rho} - \frac{1}{2}T^{\rho}{}_{\mu\nu}\nabla_{\rho}Z^{\sigma} = \frac{1}{2}R^{\sigma}{}_{\rho\mu\nu}Z^{\rho} - \frac{1}{2}T^{\sigma}{}_{\mu\nu}\nabla_{\rho}Z^{\sigma}\end{aligned}$$

□

3.2.2.2 Levi-Civita connection

The discussion on the connection has so far been independent of the metric. Starting to consider it, an important result is the *fundamental theorem of Riemannian geometry.*

Theorem 3.2.2 (Riemann)**.** *On a metric manifold $(\mathcal{M}, g)$, there exists a unique torsion-free connection that is compatible with the metric, i.e. for all $X \in \mathfrak{X}(\mathcal{M})$:*

$$\nabla_X g = 0 \tag{3.39}$$

This is called the Levi-Civita connection.

Proof. WTS uniqueness: suppose such a connection exists. Then, by Leibniz:

$$X(g(Y,Z)) = \nabla_X(g(Y,Z)) = (\nabla_X g)(Y,Z) + g(\nabla_X Y, Z) + g(Y, \nabla_X, Z)$$

Since $\nabla_X g = 0$, by cyclic permutations of X, Y and Z:

$$\begin{aligned}X(g(Y,Z)) &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \\ Y(g(Z,X)) &= g(\nabla_Y Z, X) + g(Z, \nabla_Y X) \\ Z(g(X,Y)) &= g(\nabla_Z X, Y) + g(X, \nabla_Z Y)\end{aligned}$$

Since the connection is torsion-free, $\nabla_X Y - \nabla_Y X = [X,Y]$, thus these equations become:

$$\begin{aligned}X(g(Y,Z)) &= g(\nabla_Y X, Z) + g(\nabla_X Z, Y) + g([X,Y], Z) \\ Y(g(Z,X)) &= g(\nabla_Z Y, X) + g(\nabla_Y X, Z) + g([Y,Z], X) \\ Z(g(X,Y)) &= g(\nabla_X Z, Y) + g(\nabla_Z Y, X) + g([Z,X], y)\end{aligned}$$

Adding the first two and subtracting the third:

$$\begin{aligned}g(\nabla_Y X, Z) = \frac{1}{2}[&X(g(Y,Z)) + Y(g(Z,X)) + Z(g(X,Y)) \\ &- g([X,Y],Z) - g([Y,Z],X) + g([Z,X],Y)]\end{aligned}$$

The metric is non-degenerate, thus this uniquely specifies the connection. By direct calculation it can be shown that it indeed satisfies all the properties of a connection. □

Proposition 3.2.5. *On the coordinate basis $\{\partial_\mu\}$ and $\{dx^\mu\}$ the Levi-Civita connection's components, called Christoffel symbols, are:*

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}) \tag{3.40}$$

Proof. Recalling that $[\partial_\mu, \partial_\nu] = 0$:

$$\Gamma^\lambda_{\mu\nu} g_{\lambda\rho} = g(\nabla_\mu \partial_\nu, \partial_\rho) = \frac{1}{2}(\partial_\nu g_{\mu\rho} + \partial_\mu g_{\nu\rho} - \partial_\rho g_{\mu\nu})$$

□

Example 3.2.1. In flat space $\mathbb{R}^n$, endowed with either Euclidean or Minkowski metric, it is always possibile to choose Cartesian coordinates, in which case the Christoffel symbols vanish. Being the Riemann tensor a genuine tensor, it therefore will vanish in all possibile coordinate systems on $\mathbb{R}^n$, even in those with $\Gamma^\rho_{\mu\nu} \neq 0$: this expresses the flatness of $\mathbb{R}^n$.

3.2.2.3 Gauss' theorem

The divergence theorem (or Gauss' theorem) states that the integral of a total derivative is a boundary term. It is possible to express this theorem on curved manifolds in a convenient way.

Lemma 3.2.1. $\Gamma^\mu_{\mu\nu} = \frac{1}{\sqrt{\mathrm{g}}}\partial_\nu \sqrt{\mathrm{g}}$.

Proof. A useful identity for diagonalizable matrices: $\operatorname{tr}\log A = \log\det A$. Thus (WLOG $\det g > 0$):

$$\Gamma^\mu_{\mu\nu} = \frac{1}{2} g^{\mu\rho}\partial_\nu g_{\mu\rho} = \frac{1}{2}\operatorname{tr}(g^{-1}\partial_\nu g) = \frac{1}{2}\operatorname{tr}(\partial_\nu \log g) = \frac{1}{2}\partial_\nu \log\det g = \frac{1}{\sqrt{\det g}}\partial_\nu \sqrt{\det g}$$

□

Theorem 3.2.3 (Gauss)**.** *Given a Riemannian manifold $(\mathcal{M}, g)$, consider a region $M \subseteq \mathcal{M}$ with boundary ∂M and let n^μ be an outward-pointing unit vector orthogonal to ∂M. Then, for any vector field X^μ on M:*

$$\int_M d^n x \sqrt{\mathrm{g}}\, \nabla_\mu X^\mu = \int_{\partial M} d^{n-1}x \sqrt{\gamma}\, n_\mu X^\mu \tag{3.41}$$

where γ_{ij} is the pull-back of the metric to ∂M and $\gamma \equiv \det \gamma_{ij}$.

Proof. From Lemma 3.2.1:

$$\sqrt{\mathrm{g}}\, \nabla_\mu X^\mu = \sqrt{\mathrm{g}}\left(\partial_\mu X^\mu + \Gamma^\mu_{\mu\nu} X^\nu\right) = \sqrt{\mathrm{g}}\left(\partial_\mu X^\mu + X^\nu \frac{1}{\sqrt{\mathrm{g}}}\partial_\nu \sqrt{\mathrm{g}}\right) = \partial_\mu(\sqrt{\mathrm{g}} X^\mu)$$

The integral becomes:

$$\int_M d^n x \sqrt{\mathrm{g}}\, \nabla_\mu X^\mu = \int_M d^n x\, \partial_\mu(\sqrt{\mathrm{g}} X^\mu)$$

This is the integral of an ordinary partial derivative, so the ordinary divergence theorem applies. To evaluate the integral on the boundary, it is convenient to pick coordinates so that ∂M is a surface at constant x^n. Moreover, to simplify the proof, the possible metrics will be restricted to $g_{\mu\nu} = \mathrm{diag}\,(\gamma_{ij}, N^2)$. By usual integration rules:

$$\int_M d^n x\, \partial_\mu(\sqrt{\mathrm{g}}X^\mu) = \int_{\partial M} d^{n-1}x\sqrt{\gamma N^2}X^n$$

The unit normal vector is $n^\mu = (0, \dots, 0, \frac{1}{N})$, so that $g_{\mu\nu}n^\mu n^\nu = 1$, therefore $n_\mu = g_{\mu\nu}n^\nu = (0, \dots, 0, N)$. The proof is then concluded because:

$$\int_{\partial M} d^{n-1}x\sqrt{\gamma N^2}X^n = \int_{\partial M} d^{n-1}x\sqrt{\gamma}n_\mu X^\mu$$

□

Note that this theorem holds on Lorentzian manifolds too, with the condition that ∂M must be purely timelike or purely spacelike, ensuring that $\gamma \neq 0$ at any point.

3.2.2.4 Maxwell action

Consider spacetime as a manifold $\mathcal{M}$. The electromagnetic field can be described by a form on this manifold: indeed, the electromagnetic gauge field $A_\mu = (\phi, \mathbf{A})$ is to be thought as the components of a one-form $A = A_\mu(x)dx^\mu$. The exterior derivative of this form is a 2-form $F = dA$:

$$F = \frac{1}{2}F_{\mu\nu}dx^\mu \wedge dx^\nu = \frac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu)dx^\mu \wedge dx^\nu$$

The components $F_{\mu\nu}$ are in reality the components of a tensor, the *Faraday tensor*. By construction, a useful identity holds, sometimes called the *Bianchi identity*:

$$dF = 0 \tag{3.42}$$

From this identity derive two Maxwell equations: $\nabla\cdot\mathbf{B} = 0$ and $\nabla\times\mathbf{E}+\partial_t\mathbf{B} = 0$. Moreover, note that the gauge field is not unique: the gauge transformation $A \mapsto A + d\alpha$, which equals $A_\mu \mapsto A_\mu + \partial_\mu\alpha$, leaves F unchanged.

To study the dynamics of these fields, an action is needed: Differential Geometry allows very few actions to be written down.

For example, suppose that on the considered manifold no metric is defined. To integrate over $\mathcal{M}$ a 4-form is needed, but F is a 2-form, thus the only possible action is:

$$S_{\mathrm{top}} = -\frac{1}{2}\int F \wedge F \tag{3.43}$$

The integrand become $dx^0dx^1dx^2dx^3\mathbf{E}\cdot\mathbf{B}$. Actions of this kind, independent of the metric, are called *topological actions* and are of no interest in classical physics: in fact, $F \wedge F = d(A \wedge F)$, so the action is a total derivative and doesn0't affect the equations of motion.

To construct an action of classical interest, a metric is needed. This allows to introduce a second 2-form, $\star F$, so to construct the *Maxwell action*:

$$S_{\mathrm{M}} = -\frac{1}{2}\int F \wedge \star F \tag{3.44}$$

The integrand can then be expanded as:

$$S_{\mathrm{M}} = -\frac{1}{4}\int d^4x\sqrt{\mathrm{g}}\, g^{\mu\nu}g^{\rho\sigma}F_{\mu\rho}F_{\nu\sigma} = -\frac{1}{4}\int d^4x\sqrt{\mathrm{g}}F^{\mu\nu}F_{\mu\nu}$$

In flat spacetime $F^{\mu\nu}F_{\mu\nu} = 2(\mathbf{B}^2 - \mathbf{E}^2)$. In a general curved spacetime, the equation of motion resulting from the variation of the Maxwell action is $d\star F = 0$.
To complete the theory, consider a gauge field coupled to a current, described by a one-form J. The Maxwell action then becomes:

$$S_{\mathrm{M}} = \int -\frac{1}{2}F\wedge\star F + A\wedge\star J \tag{3.45}$$

This action must retain its gauge invariance, but under $A\mapsto A+d\alpha$ it transforms as $S_{\mathrm{M}}\mapsto S_{\mathrm{M}} + \int d\alpha\wedge\star J$, therefore, after integrating by parts, the condition of gauge invariance translates to:

$$d\star J = 0 \tag{3.46}$$

This is current conservation in the language of forms. Varying the action in Eq. 3.44 now leads to the Maxwell equations with source terms:

$$d\star F = \star J \tag{3.47}$$

To define electric and magnetic charges, integrate over submanifolds. Consider a three-dimensional spatial submanifold Σ: the electric charge in Σ is defined as:

$$Q_e(\Sigma) := \int_\Sigma \star J \tag{3.48}$$

This agrees with the usual definition in flat spacetime $Q_e = \int_\Sigma d^3x J^0$. Using the equations of motion and Stokes' theorem, a general form of Gauss' law is obtained:

$$Q_e(\Sigma) = \int_{\partial\Sigma}\star F \tag{3.49}$$

Simiilarly, the magnetic charge in Σ is defined as:

$$Q_m(\Sigma) := \int_{\partial\Sigma} F \tag{3.50}$$

The non-existence of magnetic charges, following from Bianchi identity, can be evaded in topologically interesting manifolds.
From charge conservation in Eq. 3.46, it follows that the electric charge in a region cannot change, unless current flows in or out of that region. Consider a cylindrical region of spacetime V, ending in two spatial hypersurfaces Σ_1 and Σ_2: its boundary is $\partial V = \Sigma_1\cup\Sigma_2\cup B$, where B is a cylindrical timelike hypersurface. The statement that no current flows in or out of V means that $J|_B = 0$. Then:

$$Q_e(\Sigma_1) - Q_e(\Sigma_2) = \int_{\Sigma_1}\star J - \int_{\Sigma_2}\star J = \int_{\partial V}\star J - \int_B \star J = \int_{\partial V}\star J = \int_V d\star J = 0$$

Thus, electric charge in remains constant in time.

Maxwell equations from connections First note that, given the gauge field $A \in \Lambda^1(\mathcal{M})$, the field strength can be expressed via covariant derivatives:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \nabla_\mu A_\nu - \nabla_\nu A_\mu$$

The Christoffel symbols cancel out due to anti-symmetry: this is what allows to define the exterior derivative without introducing connections first.

Proposition 3.2.6. *Current conservation can be written as:* $d \star J = 0 \Leftrightarrow \nabla_\mu J^\mu = 0$.

Proof. Recalling Lemma 3.2.1:

$$\nabla_\mu J^\mu = \partial_\mu J^\mu + \Gamma^\mu_{\mu\rho} J^\rho = \partial_\mu J^\mu + \partial_\rho(\log\sqrt{\mathrm{g}}) J^\rho = \frac{1}{\sqrt{\mathrm{g}}}\partial_\mu(\sqrt{\mathrm{g}} J^\mu) \propto d \star J$$

□

As an aside, in general the divergence in different coordinate systems can be computed using the formula $\nabla_\mu J^\mu = \frac{1}{\sqrt{\mathrm{g}}}\partial_\mu(\sqrt{\mathrm{g}} J^\mu)$.

Proposition 3.2.7. $d \star F = \star J \Leftrightarrow \nabla_\mu F^{\mu\nu} = J^\nu$.

Proof. Recalling Lemma 3.2.1:

$$\nabla_\mu F^{\mu\nu} = \partial_\mu F^{\mu\nu} + \Gamma^\mu_{\mu\rho} F^{\rho\nu} + \Gamma^\nu_{\mu\rho} F^{\mu\rho} = \frac{1}{\sqrt{\mathrm{g}}}\partial_\mu(\sqrt{\mathrm{g}} F^{\mu\nu})$$

where $\Gamma^\nu_{\mu\rho} F^{\mu\rho} = 0$ because $\Gamma^\nu_{\mu\rho}$ is symmetric in μ and ρ, while $F^{\mu\rho}$ is anti-symmetric. The proof follows recalling the definition of the Hodge dual in Eq. 3.17. □

3.3 Parallel transport

The connection connects tangent spaces, or more generally any tensor vector space, at different points of the manifold: this map is called *parallel transport* and it's necessary for the definition of differentiation.

Definition 3.3.1. Consider a vector field X and some associated integral curve γ, with coordinates $x^\mu(\tau)$ such that:

$$X^\mu\Big|_\gamma = \frac{dx^\mu(\tau)}{d\tau}$$

A tensor field T is said to be *parallely transported* along γ if:

$$\nabla_X T = 0 \tag{3.51}$$

Suppose that γ connects two points $p, q \in \mathcal{M}$: Eq. 3.51 provides a map from the tensor vector space defined at p to that defined at q. To illustrate this, consider the parallel transport of a vector field Y:

$$X^\nu(\partial_\nu Y^\mu + \Gamma^\mu_{\nu\rho} Y^\rho) = 0$$

Evaluating this equation on γ, considering $Y^\mu = Y^\mu(x(\tau))$:

$$\frac{dY^\mu}{d\tau} + X^\nu \Gamma^\mu_{\nu\rho} Y^\rho = 0$$

These are a set of coupled ODEs, thus, given an initial condition (ex.: at $\tau = 0$, i.e. at p), these equations can be solved to find a unique vector at each point along the curve.
Note that the parallel transport depends both on the path (characterized by the vector field X) and on the connection.

Definition 3.3.2. Given a vector field X, a *geodesic* is a curve tangent to X such that:

$$\nabla_X X = 0 \tag{3.52}$$

Proposition 3.3.1. *A geodesic is described by the geodesic equation:*

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\rho} \frac{dx^\nu}{d\tau} \frac{dx^\rho}{d\tau} = 0 \tag{3.53}$$

Proof. From the above calculations, along γ:

$$0 = \frac{dX^\mu}{d\tau} + \Gamma^\mu_{\nu\rho} X^\nu X^\rho = \frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\rho} \frac{dx^\nu}{d\tau} \frac{dx^\rho}{d\tau}$$

$\square$

For the Levi-Civita connection $\nabla_X g = 0$. If $Y \in \mathfrak{X}(\mathcal{M})$ is parallely transported along a geodesic associated to $X \in \mathfrak{X}(\mathcal{M})$, then $\nabla_X Y = \nabla_X X = 0$, therefore $\frac{d}{d\tau} g(X, Y) = 0$: this ensures that the two tangent vectors always make the same angles along the geodesic.

3.3.1 Normal coordinates