Appendix C

Mathematics

C.1 Introduction

This appendix contains a miscellaneous assortment of mathematical ideas and results. those topics of relevance to the subject of the report.

C.2 Action Principle

C.2.1 Variational Derivative and the Euler-Lagrange Equation

The differential derivative in Cartesian coordinates in the direction of the vector $\hat{\mathbf{a}}$ (whose components a_i with respect to a linearly independent basis of vectors $\hat{\mathbf{e}}_\mathbf{i}$) is evidently

$$\lim_{\epsilon \to 0} \frac{f(x_i + \epsilon a_i) - f(x_i)}{\epsilon} \tag{C.1}$$

or as $a_i \partial_i f(x)$. The usual condition that a function by extremal is

$$\frac{\partial f(x)}{\partial x_i} = 0 \quad \text{for all } i = 1, 2, \ldots, N \tag{C.2}$$

that is $a_j \partial_j f(x) = 0$ where the components a_i are given as δ_{ij}. Due to linear independence the condition is equivalent to

$$a_j \frac{\partial f(x)}{\partial x_j} = 0, \quad \text{for all } a_j. \tag{C.3}$$

C.2.2 Action Principle for a Particle

$$\delta x(t_1) = \delta x(t_2) = 0 \tag{C.4}$$

Let us denote the path which minimizes the action as $x^{class}(t)$. Consider a variation away from this path $x(t) = x^{class}(t) + \alpha\eta(t)$ where α is a number which one varies. The action is then thought of as a function of α,

$$S(\alpha) = \int_{t_1}^{t_2} \mathcal{L}(x^{class} + \alpha\eta, \dot{x}^{class} + \alpha\dot{\eta})dt \tag{C.5}$$

This action has a minimum at $\alpha = 0$, so that

$$\frac{\partial S}{\partial \alpha}(\alpha = 0) = 0. \tag{C.6}$$

Let's take a particular example $\mathcal{L} = \frac{1}{2}m\dot{x}^2 - V(x)$. Expanding the potential and energy

$$\begin{aligned}
V(x(t) + \alpha\eta(t)) &= V(x(t)) + \alpha\eta(t)\frac{\partial V}{\partial x}(x(t)) + \ldots \\
\frac{m}{2}(\dot{x} + \alpha\dot{\eta})^2 &= \frac{m}{2}\dot{x}(t) + \alpha\ m\dot{x}\dot{\eta} + \alpha^2\ \frac{m}{2}\dot{\eta}^2(t)
\end{aligned}$$

Substituting this into the action we obtain

$$\begin{aligned}
S(\alpha) &= \int_{t_1}^{t_2}\left[\frac{1}{2}m(\dot{x} + \alpha\dot{\eta})^2 - V(x + \alpha\eta)\right]dt \\
&= \int_{t_1}^{t_2}\left[\frac{1}{2}m\dot{x}^2 - V(x)\right]dt + \int_{t_1}^{t_2}\left[m\dot{x}\dot{\eta} - \frac{\partial V}{\partial x}\eta\right]\alpha dt + \mathcal{O}(\alpha^2) \\
&= \int_{t_1}^{t_2}\left[\frac{1}{2}m\dot{x}^2 - V(x)\right]dt + \alpha\Big[m\dot{x}\eta(t)\Big]_{t_1}^{t_2} + \alpha\int_{t_1}^{t_2}\left[-m\ddot{x} - \frac{\partial V}{\partial x}\right]\eta dt + \mathcal{O}(\alpha^2) \\
&= \int_{t_1}^{t_2}\left[\frac{1}{2}m\dot{x}^2 - V(x)\right]dt + \alpha\int_{t_1}^{t_2}\left[-m\ddot{x} - \frac{\partial V}{\partial x}\right]\eta dt + \mathcal{O}(\alpha^2)
\end{aligned}$$

Then using condition (C.6) we obtain

$$m\ddot{x} = -\frac{\partial V}{\partial x}$$

which are obviously constitute Newton's equations of motion.

Sometimes you will see the notation

$$
\begin{aligned}
\delta S = \alpha \frac{\partial S}{\partial \alpha}(\alpha) &= \int_{t_1}^{t_2} \left[\frac{1}{2} m (\dot{x} + \alpha \dot{\eta})^2 - V(x + \alpha \eta) \right] dt - \int_{t_1}^{t_2} \left[\frac{1}{2} m \dot{x}^2 - V(x) \right] dt \\
&= \int_{t_1}^{t_2} \left[m \dot{x} \dot{\eta} - \frac{\partial V}{\partial x} \eta \right] \alpha dt + \mathcal{O}(\alpha^2) \\
&= \alpha \Big[m \dot{x} \eta \Big]_{t_1}^{t_2} + \int_{t_1}^{t_2} \left[-m \ddot{x} - \frac{\partial V}{\partial x} \right] \alpha \eta dt + \mathcal{O}(\alpha^2) \\
&= \int_{t_1}^{t_2} \left[-m \ddot{x} - \frac{\partial V}{\partial x} \right] \delta x(t) dt + \mathcal{O}(\alpha^2)
\end{aligned}
$$

and the condition of minimisation is written

$$
0 = \delta S = \alpha \frac{\partial S}{\partial \alpha}(\alpha = 0) = 0. \tag{C.7}
$$

C.2.3 Action Principle for one Independent Variable

We wish to minimise (or maximise) the integral

$$
S = \int_{x_1}^{x_2} dx \mathcal{L}(y, y_x, x) \tag{C.8}
$$

The path $y(x, \alpha)$ is described by

$$
y(x, \alpha) = y(x, 0) + \alpha \eta(x) \tag{C.9}
$$

The function $\eta(t)$ is arbitrary except for satifying the conditions

$$
\eta(x_1) = \eta(x_2) = 0 \tag{C.10}
$$

which is so that ever varied path passes through the fixed end points.

$$
\frac{\partial S}{\partial \alpha} = \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial y} \frac{\partial y}{\partial \alpha} + \frac{\partial \mathcal{L}}{\partial y_x} \frac{\partial y_x}{\partial \alpha} \right] dx \tag{C.11}
$$

From (C.9)

$$
\frac{\partial y(x, \alpha)}{\partial \alpha} = \eta(x) \tag{C.12}
$$

$$\frac{\partial y_x(x,\alpha)}{\partial \alpha} = \frac{d\eta(x)}{dx}. \tag{C.13}$$

Equation (C.11) becomes

$$\frac{\partial S(\alpha)}{\partial \alpha} = \int_{x_1}^{x_2} \left[\frac{\partial \mathcal{L}}{\partial y}\eta(x) + \frac{\partial \mathcal{L}}{\partial y_x}\frac{d\eta(x)}{dx}\right] dx \tag{C.14}$$

Integrating the second term by parts

$$\int_{x_1}^{x_2} \frac{\partial \mathcal{L}}{\partial y_x}\frac{d\eta(x)}{dx} dx = \left[\eta(x)\frac{\partial \mathcal{L}}{\partial y_x}\right]_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta(x)\frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_x}\right) dx \tag{C.15}$$

The integrated part vanishes due to the boundary conditions (C.10). Equation (C.14) becomes

$$\frac{\partial S(\alpha)}{\partial \alpha} = \int_{x_1}^{x_2} \left[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_x}\right)\right] \eta(x) dx \tag{C.16}$$

Putting α to zero this is equal to zero,

$$\int_{x_1}^{x_2} \left[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_x}\right)\right] \eta(x) dx = 0 \tag{C.17}$$

Since $\eta(x)$ is arbitrary the integrand vanishes,

$$\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_x}\right) = 0. \tag{C.18}$$

This is known as the Euler-Lagrange equation.

Had the function $\mathcal{L}$ also been a functional of second order derivatives of y, i.e., $\mathcal{L} = \mathcal{L}(y, y_x, y_{xx}, x)$ then there would be an extra term to (C.14):

$$\int_{x_1}^{x_2} \frac{\partial \mathcal{L}}{\partial y_{xx}}\frac{d^2\eta(x)}{dx^2} dx \tag{C.19}$$

Integrating by parts twice

$$\begin{aligned}
\int_{x_1}^{x_2} \frac{\partial \mathcal{L}}{\partial y_{xx}} d^2\eta(x) dx^2 dx &= \left[\frac{d\eta(x)}{dx}\frac{\partial \mathcal{L}}{\partial y_{xx}}\right]_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d\eta(x)}{dx}\frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_{xx}}\right) dx \\
&= \left[\frac{d\eta(x)}{dx}\frac{\partial \mathcal{L}}{\partial y_{xx}}\right]_{x_1}^{x_2} - \left[\eta(x)\frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_{xx}}\right)\right]_{x_1}^{x_2} + \int_{x_1}^{x_2} \eta(x)\frac{d^2}{dx^2}\left(\frac{\partial \mathcal{L}}{\partial y_{xx}}\right) dx \\
&= \left[\frac{d\eta(x)}{dx}\frac{\partial \mathcal{L}}{\partial y_{xx}}\right]_{x_1}^{x_2} + \int_{x_1}^{x_2} \eta(x)\frac{d^2}{dx^2}\left(\frac{\partial \mathcal{L}}{\partial y_{xx}}\right) dx
\end{aligned} \tag{C.20}$$

If the boundary conditions

$$\frac{\eta(x_1)}{dx} = \frac{\eta(x_2)}{dx} = 0 \tag{C.21}$$

are imposed, the Euler-Lagrange equation (C.18) is modified to

$$\frac{\partial\mathcal{L}}{\partial y} - \frac{d}{dx}\left(\frac{\partial\mathcal{L}}{\partial y_x}\right) + \frac{d^2}{dx^2}\left(\frac{\partial\mathcal{L}}{\partial y_{xx}}\right) = 0. \tag{C.22}$$

This particular extension, in the field variational version, is important for the variational equation for the metric or gravitational field.

C.2.4 Action Principle for Several Dependent variables

We now let $\mathcal{L}$ be a function of several dependent variables $y_1(x), y_2(x), y_3(x), \ldots$ all of which depend on x, the independent variable.

We compare neighbouring paths by writing

$$y_i(x,\alpha) = y(x,0)_i + \alpha\eta_i(x), \qquad i = 1, 2, \ldots, n, \tag{C.23}$$

where the η_i are independent, but subject them to the condition

$$\eta_i(x_1) = \eta_i(x_2) = 0, \qquad i = 1, 2, \ldots, n. \tag{C.24}$$

Differentiating with respect to α and setting $\alpha = 0$ we obtain

$$\int_{x_1}^{x_2} \sum_i \left[\frac{\partial\mathcal{L}}{\partial y_i}\eta_i(x) + \frac{\partial\mathcal{L}}{\partial y_{ix}}\frac{d\eta_i(x)}{dx}\right] dx = 0. \tag{C.25}$$

Integrating by parts

$$\int_{x_1}^{x_2} \sum_i \left[\frac{\partial\mathcal{L}}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial\mathcal{L}}{\partial y_{ix}}\right)\right] \eta_i(x)dx = 0. \tag{C.26}$$

Since the η_i are arbitrary and independent of each other, each of the terms vanishes independently

$$\frac{\partial\mathcal{L}}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial\mathcal{L}}{\partial y_{ix}}\right) = 0, \qquad i = 1, 2, \ldots, n. \tag{C.27}$$

C.2.5 Action Principle for Several Independent variables

$$\int\int\int \mathcal{L}(u, u_x, u_y, u_z, x, y, z)dxdydz \tag{C.28}$$

Let

$$u(x, y, z, \alpha) = u(x, y, z, 0) + \alpha\eta(x, y, z) \tag{C.29}$$

$$\begin{aligned} u_x(x, y, z, \alpha) &= \frac{\partial}{\partial x}u(x, y, z, 0) + \alpha\frac{\partial}{\partial x}\eta(x, y, z) \\ u_y(x, y, z, \alpha) &= \frac{\partial}{\partial y}u(x, y, z, 0) + \alpha\frac{\partial}{\partial y}\eta(x, y, z) \\ u_z(x, y, z, \alpha) &= \frac{\partial}{\partial z}u(x, y, z, 0) + \alpha\frac{\partial}{\partial z}\eta(x, y, z) \end{aligned} \tag{C.30}$$

Differentiating the integral (C.28) with respect to α and then setting $\alpha = 0$, we obtain

$$\frac{\partial S}{\partial \alpha}(\alpha = 0) = \int\int\int \left[\frac{\partial \mathcal{L}}{\partial u}\eta + \frac{\partial \mathcal{L}}{\partial u_x}\eta_x + \frac{\partial \mathcal{L}}{\partial u_y}\eta_y + \frac{\partial \mathcal{L}}{\partial u_z}\eta_z\right] dxdydz = 0 \tag{C.31}$$

Integrating by parts and using the boundary condditions

$$\int\int\int \left[\frac{\partial \mathcal{L}}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial u_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial \mathcal{L}}{\partial u_y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial \mathcal{L}}{\partial u_z}\right)\right] \eta(x, y, z)dxdydz = 0 \tag{C.32}$$

Since $\eta(x, y, z)$ is arbitrary, the integrand must be zero,

$$\frac{\partial \mathcal{L}}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial u_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial \mathcal{L}}{\partial u_y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial \mathcal{L}}{\partial u_z}\right) = 0 \tag{C.33}$$

C.2.6 Action Principle for Several Dependent variables and Several Independent variables

$$\int\int\int \mathcal{L}(y_1, y_{1x_1}, \ldots, y_{1x_m}, \ldots, y_n, y_{nx_1}, \ldots y_{nx_m} \ldots, x_1, x_2, \ldots, x_m)dxdx_2 \ldots dx_m \tag{C.34}$$

We proceed as before

$$y_i(x_1, \ldots, x_m, \alpha) = y_i(x_1, \ldots, x_m, 0) + \alpha\eta_i(x_1, \ldots, x_m), \qquad i = 1, 2, \ldots, n. \tag{C.35}$$

and

$$\frac{\partial y_i}{\partial x_j}(x_1,\ldots,x_m,\alpha) = \frac{\partial y_i}{\partial x_j}(x_1,\ldots,x_m,0)+\alpha\frac{\partial \eta_i}{\partial x_j}(x_1,\ldots,x_m), \quad i=1,2,\ldots,n \text{ and } j=1,2,\ldots,m. \tag{C.36}$$

Differentiating the integral (C.34) with respect to α and then setting $\alpha = 0$, we obtain

$$\frac{\partial S}{\partial \alpha}(\alpha=0) = \int_{x_{int}}^{x_{fin}} \sum_i \left[\frac{\partial \mathcal{L}}{\partial y_i}\eta_i(x) + \sum_j \frac{\partial \mathcal{L}}{\partial y_{ix_j}}\frac{\partial \eta_i(x)}{\partial x_j}\right] dx_1 \ldots dx_m = 0. \tag{C.37}$$

Integrating by parts results in

$$\int_{x_{int}}^{x_{fin}} \sum_i \left[\frac{\partial \mathcal{L}}{\partial y_i} - \sum_j \frac{\partial}{\partial x_j}\left(\frac{\partial \mathcal{L}}{\partial y_{ix_j}}\right)\right] \eta_i(x_1,\ldots,x_m) dx_1 \ldots dx_m = 0. \tag{C.38}$$

As the $\eta_i(x_1,\ldots,x_m)$ are arbitrary and independent of each other the integrand vanishes

$$\frac{\partial \mathcal{L}}{\partial y_i} - \sum_j \frac{\partial}{\partial x_j}\left(\frac{\partial \mathcal{L}}{\partial(\partial y_i/\partial x_j)}\right) = 0, \qquad i=1,2,\ldots,n. \tag{C.39}$$

C.2.7 Complex dependent variables

We now let $\mathcal{L}$ be a function of a complex valued dependent variables $y(x) = y_R(x) + iy_I(x)$ and $\overline{y}(x) = y_R(x) - iy_I(x)$ which depends on x, the independent variable. The action is of the form

$$S = \int_{x_1}^{x_2} \mathcal{L}(\overline{y}, y, \overline{y}_x, y_x)dx. \tag{C.40}$$

We compare neighbouring paths by writing

$$y_R(x,\alpha) = y(x,0)_R + \alpha\eta_R(x), \qquad y_I(x,\alpha) = y(x,0)_I + \alpha\eta_I(x) \tag{C.41}$$

where the η_I and η_R are independent, but subject them to the condition

$$\eta_{R,I}(x_1) = \eta_{R,I}(x_2) = 0. \tag{C.42}$$

$$
\begin{aligned}
S(\alpha) \quad = \quad & \int_{x_1}^{x_2} \mathcal{L}(\overline{y}, y)dx + \int_{x_1}^{x_2} \left[\frac{\partial \mathcal{L}}{\partial \overline{y}}(\alpha\eta_R - i\alpha\eta_I) + \frac{\partial \mathcal{L}}{\partial \overline{y}_x}(\alpha\frac{d\eta_R}{dx} - i\alpha\frac{d\eta_R}{d\eta_I}) \right. \\
& \left. + \frac{\partial \mathcal{L}}{\partial y}(\alpha\eta_R + i\alpha\eta_I) + \frac{\partial \mathcal{L}}{\partial y_x}(\alpha\frac{d\eta_R}{dx} + i\alpha\frac{d\eta_R}{d\eta_I}) \right] + \mathcal{O}(\alpha^2)
\end{aligned} \tag{C.43}
$$

Differentiating with respect to α and setting $\alpha = 0$ we obtain

$$
\begin{aligned}
& \int_{x_1}^{x_2} \left[\frac{\partial \mathcal{L}}{\partial \overline{y}}(\eta_R(x) - i\eta_I(x)) + \frac{\partial \mathcal{L}}{\partial \overline{y}_x}(\frac{d\eta_R(x)}{dx} - i\frac{d\eta_R(x)}{d\eta_I}) \right. \\
& \quad \left. + \frac{\partial \mathcal{L}}{\partial y}(\eta_R(x) + i\eta_I(x)) + \frac{\partial \mathcal{L}}{\partial y_x}(\frac{d\eta_R(x)}{dx} + i\frac{d\eta_R(x)}{d\eta_I}) \right] dx = 0.
\end{aligned} \tag{C.44}
$$

Integrating by parts

$$
\int_{x_1}^{x_2} \left\{ \left[\frac{\partial \mathcal{L}}{\partial \overline{y}} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial \overline{y}_x} \right) \right] \big(\eta_R(x) - i\eta_I(x)\big) + \left[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_x} \right) \right] \big(\eta_R(x) + i\eta_I(x)\big) \right\} dx \tag{C.45}
$$

Which can be written

$$
\begin{aligned}
& \int_{x_1}^{x_2} \left\{ \left(\left[\frac{\partial \mathcal{L}}{\partial \overline{y}} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial \overline{y}_x} \right) \right] + \left[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_x} \right) \right] \right) \eta_R(x) \right. \\
& \left. +i \left(- \left[\frac{\partial \mathcal{L}}{\partial \overline{y}} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial \overline{y}_x} \right) \right] + \left[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_x} \right) \right] \right) \eta_I(x) \right\} dx = 0.
\end{aligned} \tag{C.46}
$$

Since the η_R and η_I are arbitrary and independent of each other, each of the terms vanishes independently and we get

$$
\begin{aligned}
\left[\frac{\partial \mathcal{L}}{\partial \overline{y}} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial \overline{y}_x} \right) \right] + \left[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_x} \right) \right] &= 0 \\
- \left[\frac{\partial \mathcal{L}}{\partial \overline{y}} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial \overline{y}_x} \right) \right] + \left[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_x} \right) \right] &= 0
\end{aligned} \tag{C.47}
$$

or, finally,

$$
\frac{\partial \mathcal{L}}{\partial \overline{y}} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial \overline{y}_x} \right) = 0 \qquad \text{and} \qquad \frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y_x} \right) = 0. \tag{C.48}
$$

Action for Dirac's equation

This is an exaple of an action with several dependent complex variables and several independent variables. Dirac's equation can be written

$$\left(i\hbar\gamma^{\mu}\frac{\partial}{\partial x^{\mu}} - m_0\right)\psi(x,y,z,t) = 0 \tag{C.49}$$

where

$$\psi = \begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix} \tag{C.50}$$

and γ^{μ} are the 4×4 matrices:

$$\gamma^0 = \begin{pmatrix} I & 0\\ 0 & -I\end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \hat{\sigma}_i\\ -\hat{\sigma}_i & 0\end{pmatrix} \tag{C.51}$$

where $\hat{\sigma}_i$ are Pauli's 2×2 matrices and I is the 2×2 unit matrix. With the explicit form of the Pauli matrices we have in detail.

$$\begin{aligned}
\gamma^0 &= \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}, \qquad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\\ 0 & -1 & 0 & 0\\ -1 & 0 & 0 & 0\end{pmatrix}\\
\gamma^2 &= \begin{pmatrix} 0 & 0 & 0 & -i\\ 0 & 0 & i & 0\\ 0 & i & 0 & 0\\ -i & 0 & 0 & 0\end{pmatrix}, \qquad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\\ 1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\end{pmatrix}
\end{aligned} \tag{C.52}$$

The action can be written

$$S = \int \overline{\psi}\left(i\hbar\gamma^{\mu}\frac{\partial}{\partial x^{\mu}} - m_0 I\right)\psi d^4x \tag{C.53}$$

where $\overline{\psi} = \psi^{\dagger}\gamma^0$. The action written more explicitly is

$$
\begin{aligned}
S &= \int (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*) \\
&\quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \left\{ i\hbar \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \frac{\partial}{\partial ct} + i\hbar \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \frac{\partial}{\partial x^1} \right. \\
&+ \left. i\hbar \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \frac{\partial}{\partial x^2} + i\hbar \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \frac{\partial}{\partial x^3} - m_0 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right\} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} d^4x \\
&= \int (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*) \begin{pmatrix} \tilde{\psi}_1 \\ \tilde{\psi}_2 \\ \tilde{\psi}_3 \\ \tilde{\psi}_4 \end{pmatrix} d^4x \\
&= \int (\psi_1^* \tilde{\psi}_1 + \psi_2^* \tilde{\psi}_2 + \psi_3^* \tilde{\psi}_3 + \psi_4^* \tilde{\psi}_4) d^4x \qquad \text{(C.54)}
\end{aligned}
$$

where

$$
\begin{pmatrix} \tilde{\psi}_1 \\ \tilde{\psi}_2 \\ \tilde{\psi}_3 \\ \tilde{\psi}_4 \end{pmatrix}
$$

denotes the components of

$$
\gamma^0 \left(i\hbar \gamma^\mu \frac{\partial}{\partial x^\mu} - m_0 I \right) \psi.
$$

The Eurler-Lagrange equations we use are

$$
\frac{\partial \mathcal{L}}{\partial \psi_i^*} - \sum_{\mu=0}^{3} \frac{\partial}{\partial x^\mu} \left(\frac{\partial \mathcal{L}}{\partial (\partial \psi_i^* / \partial x^\mu)} \right) = 0, \qquad i = 1, 2, 3, 4. \qquad \text{(C.55)}
$$

from which we easily obtain

$$
\tilde{\psi}_1 = 0, \quad \tilde{\psi}_2 = 0, \quad \tilde{\psi}_3 = 0, \quad \tilde{\psi}_4 = 0
$$

or

$$\begin{pmatrix} \tilde{\psi}_1 \\ \tilde{\psi}_2 \\ \tilde{\psi}_3 \\ \tilde{\psi}_4 \end{pmatrix} = 0 \tag{C.56}$$

in other words

$$\gamma^0 \left(i\hbar\gamma^\mu \frac{\partial}{\partial x^\mu} - m_0 I \right) \psi = 0 \tag{C.57}$$

and upon using $(\gamma^0)^2 = I$ we obtain Dirac's equation (C.49).

C.2.8 Fucntional Derivatives and the Fundamental Lemma of Calculus of Variations

First approach:

The variational principle for ϕ_σ

$$\delta \int \mathcal{L} \left(\phi_\sigma, \frac{\partial \phi_\sigma}{\partial x_\mu} \right) d^d x = 0 \tag{C.58}$$

$$\delta \int \mathcal{L} \left(\phi_\sigma, \frac{\partial \phi_\sigma}{\partial x^\mu} \right) d^d x = \int \left[\frac{\partial \mathcal{L}}{\partial \phi_\sigma} \delta \phi_\sigma + \frac{\partial \mathcal{L}}{\partial (\partial \phi_\sigma / \partial x^\mu)} \delta \left(\frac{\partial \phi_\sigma}{\partial x^\mu} \right) \right] d^d x = 0 \tag{C.59}$$

using

$$\delta \left(\frac{\partial \phi_\sigma}{\partial x^\mu} \right) = \frac{\partial}{\partial x^\mu} (\phi_\sigma + \delta \phi_\sigma) - \frac{\partial}{\partial x^\mu} \phi_\sigma = \frac{\partial}{\partial x^\mu} (\delta \phi_\sigma).$$

and integrating by parts the second term in (C.59),

$$\left[\frac{\partial \mathcal{L}}{\partial (\partial \phi_\sigma / \partial x^\mu)} \delta \phi_\sigma \right] + \int \left[\frac{\partial \mathcal{L}}{\partial \phi_\sigma} - \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial (\partial \phi_\sigma / \partial x^\mu)} \right] \delta \phi_\sigma d^d x = 0$$

Taking into account that $\delta \phi_\sigma$ vanishes at the boundariesleads to,

$$\int \left[\frac{\partial \mathcal{L}}{\partial \phi_\sigma} - \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial (\partial \phi_\sigma / \partial x^\mu)} \right] \delta \phi_\sigma d^d x = 0 \tag{C.60}$$

As these equations are valid for arbitrary variations $\delta \phi_\sigma$, we can use the fundamental lemma of calculus of variations (proven below) to obtain the Euler-Lagrange equations or field equations:

$$\frac{\partial \mathcal{L}}{\partial \phi_\sigma} - \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial(\partial \phi_\sigma / \partial x^\mu)} = 0. \tag{C.61}$$

Second approach:

We define the functional derivative in the direction of $\eta(x)$ as

$$\frac{\delta_\eta F[\phi]}{\delta_\eta \phi} = \lim_{\epsilon \to 0} \frac{F[\phi + \epsilon\eta] - F[\phi]]}{\epsilon}. \tag{C.62}$$

$$\frac{\delta_\eta \phi(y)}{\delta_\eta \phi(x)} = \lim_{\epsilon \to 0} \frac{\phi(y) + \epsilon\eta(x) - \phi(y)}{\epsilon} = \eta(x - y) \tag{C.63}$$

$$F[\phi] = \int dx \mathcal{L}[\phi(x), \partial_a \phi(x)] \tag{C.64}$$

$$\begin{aligned} \frac{\delta_\eta F[\phi(y)]}{\delta_\eta \phi} &= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int dy \mathcal{L}\left[\phi(y) + \epsilon\eta(y), \partial_a(\phi(y) + \epsilon\eta(y))\right] - \mathcal{L}\left[\phi(y), \partial_a \phi(y)\right] \\ &= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int dy \left[\frac{\partial \mathcal{L}[\phi, \partial_y \phi]}{\partial \phi(x)} - \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}[\phi, \partial_y \phi]}{\partial(\partial_x \phi(x))}\right)\right] \eta(y) \end{aligned} \tag{C.65}$$

$$\frac{\delta}{\delta \phi(x)} \int dy \mathcal{L}[\phi(y), \partial_y \phi(y)] = \frac{\partial \mathcal{L}[\phi, \partial_y \phi]}{\partial \phi(x)} - \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}[\phi, \partial_y \phi]}{\partial(\partial_x \phi(x))}\right) \tag{C.66}$$

the rather singular equation $\delta(x - y)$ can be replaced by a multiplicity of conditions involving well behaved "test" functions $\eta(x - y)$.

Proofs of the fundamental lemma of calculus of variations

In mathematics, specfically in the calculus of variations, the variation δF of function F can be concentrated on arbitrarily small interval, but not a single point. Accordingly, the necessary condition of extremum (functional derivative equal to zero) apears in a weak formulation (variational form) integrated with an arbitrary function δF. The fundamental lemma of calculus of variations is typically used to transform this weak forulation into the strong forumaltion (differential equation), free of the integration with arbitrary function. The proof usually exploits the possibility to choose δF concentrated on an interval on which F keeps sign (positive or negative). Several versions of the lemma are in use. Here we provide the proof of a few basic versions.

Lemma C.2.1 *Fundamental lemma of calculus of variations.*

If $F(x)$ is continuous in $[a, b]$ and if $\int_a^b F(x)\eta(x)dx = 0$ for every function $\eta(x) \in c^0(a, b)$ such that $\eta(a) = \eta(b) = 0$, then $F(x) = 0$ for all $x \in [a, b]$.

Proof:

Suppose there exists a point $c \in (a, b)$ such that $F(c) \neq 0$. We may ussume without loss of generality that $F(c) > 0$. By continuity of the function $F(x)$ there exists an open inetrval $(x_1, x_2) \subset (a, b)$ containing the point c such that $F(x) > 0$ on (x_1, x_2). Let $\eta(x) = (x-x_1)(x_2-x)$ for $x \in [x_1, x_2]$ and zero outside of $[x_1, x_2]$. Note that $(x-x_1)(x_2-x)$ is positive for $x \in (x_1, x_2)$. Now, consider

$$\begin{aligned}
\int_a^b F(x)\eta(x)dx &= \int_a^{x_1} F(x)\eta(x)dx + \int_{x_1}^{x_2} F(x)\eta(x)dx + \int_{x_2}^{b} F(x)\eta(x)dx \\
&= 0 + \int_{x_1}^{x_2} F(x)\eta(x)dx + 0 \\
&= \int_{x_1}^{x_2} F(x)(x - x_1)(x_2 - x)dx > 0. \qquad \text{(C.67)}
\end{aligned}$$

Thus, we get a contradiction to what is given in the lemma. So, we can conclude that $F(x)$ cannot be non-zero anywhere in (a, b). By continuity of F we obtain $F(x) = 0$ on $[a, b]$.

□

Lemma C.2.2 *Fundamental lemma of calculus of variations.*

If

$$\int_a^b F(x)\eta(x)dx = 0 \qquad \text{(C.68)}$$

for all η with continuous second derivatives, and $\eta(a) = \eta(b) = 0$ then

$$F(x) = 0, \quad for\ x \in (a, b). \qquad \text{(C.69)}$$

Proof: *Let*

$$\eta(x) = F(x)(x - a)(b - x) \qquad \text{(C.70)}$$

Note η is continuous and $F(x)\eta(x) > 0$ for $x \in (a, b)$ with $F(a)\eta(a) = F(b)\eta(b) = 0$. If the integral of a non-negative function is zero then the function must be zero at all point, i.e., $F(x)\eta(x) = 0$ for $x \in [a, b]$. Thus

$$F^2(x)(x-a)(b-x) = 0, \quad \text{for } x \in [a,b] \tag{C.71}$$

and since $(x-a)(b-x) > 0$ *for* $x \in (a,b)$ *then* $F^2(x) = 0$ *for* $x \in (a,b)$*, thus* $F(x) = 0$ *for* $x \in (a,b)$*.*

□

Fundamental lemma of calculus of variations via a Molifier

We consider the case where $\eta(x)$ is smooth (infinitely differentiable) and positive but with compact support - i.e. non-zero (positive) in a finite range, say between values x_1 and x_2.

First we introduce the function defined by

$$f(x) = \begin{cases} \exp(-1/x) & \text{if } x > 0, \\ 0 & \text{if } x \le 0, \end{cases} \tag{C.72}$$

defined for every real number x. We wish to prove the function is smooth. To this end...

By the power series representation of the exponential function, we have for every natural number m (including zero)

$$\begin{aligned} \frac{1}{x^m} &= x\left(\frac{1}{x^{m+1}}\right) \\ &\le (m+1)!x\left(\sum_{n=0}^{m}\frac{1}{n!}\left(\frac{1}{x}\right)^n + \frac{1}{(m+1)!}\frac{1}{x^{m+1}} + \sum_{n=m+2}^{\infty}\frac{1}{n!}\left(\frac{1}{x}\right)^n\right) \\ &= (m+1)!x\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{1}{x}\right)^n \\ &= (m+1)!x\exp\left(\frac{1}{x}\right), \qquad x > 0. \end{aligned} \tag{C.73}$$

Therefore,

$$\lim_{x\to 0_+}\frac{e^{-1/x}}{x^m} = (m+1)!\lim_{x\to 0_+} x = 0. \tag{C.74}$$

We now prove that f has continuous dreivatives to all orders in all points x of the real line, given by

$$f^{(n)}(x) = \begin{cases} \dfrac{p_n(x)}{x^{2n}}f(x) & \text{if } x > 0, \\ 0 & \text{if } x \le 0, \end{cases} \tag{C.75}$$

where $p_n(x)$ is a polynomial of degree $n-1$ given recursively by $p_1(x) = 1$ and

$$p_{n+1}(x) = x^2 p_n'(x) - (2nx - 1)p_n(x), \qquad n \in \mathbb{N}. \tag{C.76}$$

We prove the formula for the nth derivative of f by mathematical induction. First note that $p_1(x) = 1$ means $p_1(x)$ is of degree 0. Of course the derivative of f is zero for $x < 0$. It remains to show that the right-hand side derivative of f at $x = 0$ at zero. Using the above limit, we see that

$$f'(0) = \lim_{x \to 0_+} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0_+} \frac{e^{-1/x}}{x} = 0. \tag{C.77}$$

The inductive step from n to $n+1$ is similar. For $x > 0$ we get for the derivative

$$\begin{aligned} f^{(n+1)}(x) &= \frac{d}{dx}\left(\frac{p_n(x)}{x^{2n}} f(x)\right) \\ &= \left(\frac{p_n'(x)}{x^{2n}} - 2n\frac{p_n(x)}{x^{2n+1}} + \frac{p_n(x)}{x^{2n+2}}\right) f(x) \\ &= \frac{x^2 p_n'(x) - (2nx-1)p_n(x)}{x^{2n+2}} f(x) \\ &= \frac{p_{n+1}(x)}{x^{2(n+1)}} \end{aligned} \tag{C.78}$$

where $p_{n+1}(x)$ is a polynomial of degree $n = (n+1) - 1$. Of course, the $(n+1)$st derivative of f is zero for $x < 0$. For the right-hand side derivative of $f^{(n)}$ at $x = 0$ we obtain with the above limit

$$\lim_{x \to 0_+} \frac{f^{(n)}(x) - f^{(n)}(0)}{x - 0} = \lim_{x \to 0_+} \frac{p_n(x)}{x^{2n+1}} e^{-1/x} = 0. \tag{C.79}$$

We can now introduce the "bump-function" using the function f described above,

$$\begin{aligned} \tilde{f}(x) &= f(c+x)f(c-x) \\ &= (\exp(-1/(c+x)))(\exp(-1/(c-x))) \\ &\equiv \exp(-1/(c^2 - x^2))) \\ &= f(c^2 - x^2) \end{aligned} \tag{C.80}$$

which is smooth and vanishes for $|x| \geq c > 0$. This gives a "bump" centred at the origin, with width $2c$, symmetric about the origin. We can shift the function, the "bump", to the right by the amount Δ via $x \to x - \Delta$, viz

$$\tilde{\tilde{f}}(x) = \tilde{f}(x - \Delta) = f(c + x - \Delta) f(c - x + \Delta). \tag{C.81}$$

If we require that our function $\eta(x)$ to have the above properties and be zero for $x \leq x_1$ and $x \geq x_2$, then we require

$$-c + \Delta = x_1 \quad \text{and} \quad c + \Delta = x_2$$

or

$$\Delta = \frac{x_1 + x_2}{2} \quad \text{and} \quad c = \frac{x_2 - x_1}{2},$$

thus

$$\eta(x) = f\left(\frac{x_2 - x_1}{2} + \left(x - \frac{x_1 + x_2}{2}\right)\right) f\left(\frac{x_2 - x_1}{2} - \left(x - \frac{x_1 + x_2}{2}\right)\right) \tag{C.82}$$

has all the desired properties specified at the beginning. It simplifies to

$$\eta(x) = f\left((x - x_1)(x_2 - x)\right) \tag{C.83}$$

Lemma C.2.3 *Fundamental lemma of calculus of variations.*

If

$$\int_a^b F(x)\eta(x)dx = 0 \tag{C.84}$$

for F continuous and all smooth η such that $\eta(a) = \eta(b) = 0$.

Proof:

Suppose there exists a point $c \in (a, b)$ such that $F(c) \neq 0$. We may ussume without loss of generality that $F(c) > 0$. By continuity of the function $F(x)$ there exists an open inetrval $(x_1, x_2) \subset (a, b)$ containing the point c such that $F(x) > 0$ on (x_1, x_2). Let $\eta(x)$ be given as above (i.e. (C.83)) so that it is positive in $x \in (x_1, x_2)$ and zero outside. Now, consider

$$\begin{aligned} \int_a^b F(x)\eta(x)dx &= \int_a^{x_1} F(x)\eta(x)dx + \int_{x_1}^{x_2} F(x)\eta(x)dx + \int_{x_2}^b F(x)\eta(x)dx \\ &= 0 + \int_{x_1}^{x_2} F(x)\eta(x)dx + 0 \\ &= \int_{x_1}^{x_2} F(x)\eta(x)dx > 0. \end{aligned} \tag{C.85}$$

Thus, we get a contradiction to what is given in the lemma. So, we can conclude that $F(x)$ cannot be non-zero anywhere in (a, b). By continuity of F we obtain $F(x) = 0$ on $[a, b]$.

$\square$

Lemmas C.2.1 and C.2.3 generalise to the case where F is complex valued. By linearity of the integral, it is easy to see that one only needs to prove the claim for real F.

C.2.9 Noether's Theorem

In classical mechanics and classical field theory there is a basic result relates symmetries to conserved quantities (often called conserved charges). Noether's theorem is a general proof of this, which states that for every continuous symmetry of the system there is a conserved current. Its simplest form in classical mechanics is easily proved.

First the time translation symmetry. We explain what symmetry we are talking about - we need to explain the difference between an active transformation and a passive one. Say we start off a particle from a certain position at a certain time. We measure the position of the particle and time with a clock. Say after some time we perform the same - if the trajectory is the same as in the first experiment we say that there is a **time-translation symmetry**. This is an active transformation. Whereas, had we used a different clock in the first experiment. Symmetries are only revealed by active transformations.

$$\mathcal{L} = \mathcal{L}(q_i, \dot{q}_i) \tag{C.86}$$

Then

$$\begin{aligned} \frac{d\mathcal{L}}{dt} &= \sum_i \frac{\partial \mathcal{L}}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\ddot{q}_i \\ &= \sum_i \dot{q}_i \left(\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} \right) + \sum_i \frac{d}{dt}\left(\dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \end{aligned} \tag{C.87}$$

$$\frac{d}{dt}\left(\dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L} \right) = 0. \tag{C.88}$$

$$E = \frac{d}{dt}\left(\dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L} \right) = \sum_i \dot{q}_i p_i - \mathcal{L} = H, \tag{C.89}$$

The total energy.

Let

$$
\begin{aligned}
0 &= \frac{\partial \mathcal{L}}{\partial q_i}\delta_\epsilon q_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\delta_\epsilon \dot{q}_i \\
&= \left[\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right]\delta_\epsilon q_i + \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\delta_\epsilon q_i\right)
\end{aligned} \tag{C.90}
$$

When $\delta_\epsilon q_i$ is applied to a solution to the equations of motion the term inside the square brackets vanishes and we conclude that there is a conserve quantity

$$
\dot{Q} = 0 \qquad \text{with} \qquad Q := \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\delta_\epsilon q_i. \tag{C.91}
$$

$$
\partial_\mu K^\mu = \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi}\partial_\mu \delta_\epsilon \phi + \frac{\partial \mathcal{L}}{\partial \phi}\delta_\epsilon \phi = \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi}\partial_\mu \delta_\epsilon\right) + \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right)\right]\delta_\epsilon \phi. \tag{C.92}
$$

If ϕ is a solution to the equations of motion the term inside the square brackets vanishes, and we find that there is a conserve current

$$
\partial_\mu J^\mu = 0 \qquad \text{with} \qquad Q := J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta_\epsilon \phi - K^\mu. \tag{C.93}
$$

A conserved current impliers the existence of a charge

$$
Q := \int d^3x J^0(t, \vec{x}) \tag{C.94}
$$

which is conserved as

$$
\begin{aligned}
\frac{dQ}{dt} &= \int d^3x\, \partial_0 J^0(t, \vec{x}) = -\int d^3x\, \partial_i J^i(t, \vec{x}) \\
&= -\int dS_i J^i(t, \vec{x}) \qquad \text{by Gauss' theorem} \\
&= 0,
\end{aligned} \tag{C.95}
$$

provided the fields vanish fast enough at infinity.

As an example we consider a scalar field $\phi(x)$ which is invaraint under a space-time translation $x^\mu \to x'^\mu = x^\mu + a^\mu$. We have

$$
\phi'(x) - \phi(x) = -a^\mu \partial_\mu \phi + \mathcal{O}(a^2) \tag{C.96}
$$

This implies

$$\delta\phi = -a^\mu \partial_\mu \phi. \tag{C.97}$$

$$\delta\mathcal{L} = -a^\mu \partial_\mu \mathcal{L}. \tag{C.98}$$

Therefore the corresponding conserved charge is

$$J^\mu = -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} a^\mu \partial_\mu \phi + a^\mu \mathcal{L} := -a_\nu T^{\mu\nu}. \tag{C.99}$$

where we have introduced quantity which turns out to be a general formula for the energy-momnetum tensor

$$T^{\mu\nu} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \partial^\mu \phi - \eta^{\mu\nu}\mathcal{L}. \tag{C.100}$$

C.2.10 Action Principle with Variantion of Boundary

Would like to consider also changes in spacelike final and initial hypersurfaces corresponding to intial times t_1 and final time t_2 which form part of the boundary ∂R. Later we would like to consider arbitrary hypersurfaces of the region R.

$$\begin{aligned} x^\mu \to x'^\mu &= x^\mu + \delta x^\mu && \text{(C.101)} \\ \phi \to \phi'(x) &= \phi(x) + \delta\phi(x). && \text{(C.102)} \end{aligned}$$

It is important to note that in (C.102) is the functional variance in ϕ, that is, ϕ' is compared to ϕ at the same point in spacetime. There will also be a variation in ϕ coming from (C.101). So there will be a total variation in ϕ, $\Delta\phi$,

$$\phi'(x') = \phi(x) + \Delta\phi(x) \tag{C.103}$$

to first order in δx^μ

$$\begin{aligned} \Delta\phi(x) &= \phi'(x') - \phi(x) \\ &= \left[\phi'(x') - \phi(x')\right] + \left[\phi(x') - \phi(x)\right] \\ &= \delta\phi + (\partial_\mu\phi)\delta x^\mu. && \text{(C.104)} \end{aligned}$$

The variation in the action is

$$\delta S = \int \mathcal{L}(\phi', \partial_\mu \phi', x'^\mu) d^4x' - \int \mathcal{L}(\phi, \partial_\mu \phi, x^\mu) d^4x \tag{C.105}$$

where $d^4x' = J(x'/x)d^4x$, $J(x'/x)$ being the Jacobian of the transformation $x \to x'$.

$$\frac{\partial x'^\mu}{\partial x^\nu} = \delta^\nu_\mu + \partial_\nu \delta x^\mu$$

so

$$J\left(\frac{x'}{x}\right) = \det\left(\frac{\partial x'^\mu}{\partial x^\nu}\right) = 1 + \partial_\mu(\delta x^\mu)$$

$$\delta S = \int \tag{C.106}$$

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi}\delta(\partial_\mu \phi) + \frac{\partial \mathcal{L}}{\partial x^\mu}\delta x^\mu \tag{C.107}$$

$$\delta S = \int_R \left\{ \frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) \right\} \delta\phi + \int_{\partial R} \left\{ \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta\phi + \mathcal{L}\delta x^\mu \right\} d\sigma_\mu. \tag{C.108}$$

C.2.11 Lagrange multipliers

Suppose we are given the problem to find the extremum of a function of two variables $f(x, y)$ subject to a constraint $g(x, y) = 0$.

First derivation

First we need a result of level surfaces of a function $g(x, y)$. A level surface is the curve defined by $g(x, y) = C$ where C is a constant. For a vector

$$d\vec{r} = \vec{i}dx + \vec{j}dy$$

.

we have

$$(\nabla g) \cdot d\vec{r} = \frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy = dg,$$

the change in the function g corresponding to a change in position $d\vec{r}$. Consider two points $\vec{OQ}$ and $\vec{OP}$ on the curve $g(x, y) = C$, such that $\vec{OQ} = \vec{OP} + d\vec{r}$. Then

$$(\nabla g) \cdot d\vec{r} = 0,$$

since we stay on the curve $g(x, y) = C$. This shows that ∇g is perpendicular to $d\vec{r}$ and hence ∇g is normal to the curve $g(x, y) = C$.

Let us return to the original problem. The graph of $g(x, y) = 0$ is a curve $\mathcal{C}$ in the $xy-$plane. This curve has a parametrization

$$x = h(t), \quad y = k(t)$$

for t in some interval I. Let

$$\vec{r}(t) = x\hat{i} + y\hat{j} = h(t)\hat{i} + k(t)\hat{j}$$

be the position vector of the point $P(x, y)$ on $\mathcal{C}$. If we define a function F of one variable t by

$$F(t) = f(h(t), k(t)),$$

then as t varies, we obtain function values $f(x, y)$ that correpsond to (x, y) on $\mathcal{C}$. We look for the extremum of $F(t)$. $F'(t_0)$. By the chain rule

$$\frac{d}{dt}F(t) = \frac{\partial}{\partial x}f(x, y)\frac{dx}{dt} + \frac{\partial}{\partial y}f(x, y)\frac{dy}{dt} = \frac{\partial}{\partial x}f(x, y)h'(t) + \frac{\partial}{\partial y}f(x, y)k'(t)$$

Say we have extremum at $t = t_0$, then

$$\begin{aligned} 0 = \frac{d}{dt}F(t) &= \frac{\partial}{\partial x}f(x_0, y_0)h'(t_0) + \frac{\partial}{\partial y}f(x_0, y_0)k'(t_0) \\ &= \nabla f(x_0, y_0) \cdot \vec{r}'(t_0) \end{aligned} \tag{C.109}$$

This shows that the vector $\vec{\nabla} f(x_0, y_0)$ is orthogonal to the tangent vector $\vec{r}'(t_0)$ to $\mathcal{C}$. However, $\vec{\nabla} g(x_0, y_0)$ is also orthogonal to $\vec{r}'(t_0)$ because $\mathcal{C}$ is a level curve for g. And hence we can write

$$\vec{\nabla} f(x_0, y_0) = \lambda \vec{\nabla} g(x_0, y_0)$$

for some λ. This is the condition to have an extremum of $f(x, y)$ on the level surface defined by $g(x, y) = 0$.

Extremum subject to the constraint are found by solving the system of equations:

$$\begin{aligned}
\frac{\partial}{\partial x}f(x,y) &= \lambda\frac{\partial}{\partial x}g(x,y)\\
\frac{\partial}{\partial y}f(x,y) &= \lambda\frac{\partial}{\partial y}g(x,y)\\
g(x,y) &= 0.
\end{aligned}\tag{C.110}$$

This can be generalised to any number of constraints.

Second derivation

In the absence of constraints the function $f(x,y,z)$ has an extremum if and only if

$$\frac{\partial f}{\partial x}=0,\qquad \frac{\partial f}{\partial y}=0,\qquad \frac{\partial f}{\partial z}=0 \tag{C.111}$$

$$df=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy+\frac{\partial f}{\partial z}dz=0 \tag{C.112}$$

However if we impose $g(x,y)=0$ the variables x,y are subject to constraints and so are no longer independent.

Langrangian multipliers

From $g(x,y,z)=0$ we have

$$dg=\frac{\partial g}{\partial x}dx+\frac{\partial g}{\partial y}dy+\frac{\partial g}{\partial z}dz=0 \tag{C.113}$$

If we take x and y to be the independent variables, dz is no longer arbitrary.

$$dz=-\left(\frac{\partial g}{\partial x}/\frac{\partial g}{\partial z}\right)dx-\left(\frac{\partial g}{\partial y}/\frac{\partial g}{\partial z}\right)dy \tag{C.114}$$

Denote by $(df)_g$ the change in f while remaining on the level surface $g(x,y,z)=0$. Then

$$\begin{aligned}
(df)_g &= \frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy+\frac{\partial f}{\partial z}\left[-\left(\frac{\partial g}{\partial x}/\frac{\partial g}{\partial z}\right)dx-\left(\frac{\partial g}{\partial y}/\frac{\partial g}{\partial z}\right)dy\right]\\
&= \left(\frac{\partial f}{\partial x}-(\frac{\partial f}{\partial z}/\frac{\partial g}{\partial z})\frac{\partial g}{\partial x}\right)dx+\left(\frac{\partial f}{\partial y}-(\frac{\partial f}{\partial z}/\frac{\partial g}{\partial z})\frac{\partial g}{\partial y}\right)dy.
\end{aligned}\tag{C.115}$$

At an extremum $(\frac{\partial f}{\partial z}/\frac{\partial g}{\partial z})$ has a definite value. Our Lagrangian multiplier λ is chosen as

$$\lambda = -\left(\frac{\partial f}{\partial z} / \frac{\partial g}{\partial z}\right) \tag{C.116}$$

assuming $\partial g/\partial y \neq 0$. At an extremum we have

$$(df)_g = \left(\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y}\right) dy = 0. \tag{C.117}$$

However, as we took dx and dy to be arbitrary the quantity in the paranthses must vanish,

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0 \tag{C.118}$$

Equantion (C.116) rearranged is

$$\frac{\partial f}{\partial z} + \lambda \frac{\partial g}{\partial z} = 0 \tag{C.119}$$

When (C.118) and (C.119) are satisfied f is an extremum (again we have derived that at an extremum $\nabla f \propto \nabla g$). Note that there are now four unknowns: x, y, z, and λ. The fourth equation is, of course, the constraint $g(x, y, z) = 0$.

Treating $f + \lambda g$ as a new function for arbitary dx, dy and dz we have an extremum when

$$df + \lambda dg = \left(\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y}\right) dy + \left(\frac{\partial f}{\partial z} + \lambda \frac{\partial g}{\partial z}\right) dz = 0. \tag{C.120}$$

Solving for $f + \lambda g$ for its unconstrained extremum points generate exactly the same extremum points as solving for the constrained extremum points of f under the constraint g.

C.2.12 Action principle subject to constraints

As before, we seek the 'path' that make the integral

$$S = \int \mathcal{L}\left(y_i, \frac{\partial yi}{\partial x_j}, x_j\right) d^d x \tag{C.121}$$

stationary, but now where there are also one or more constraints. This means that the y_i's are no longer independent of each other. Not all the η_i's may be varied arbitrarily and () and (C.39) would not apply.

Sometimes we deal with constraints of the form

$$g(y_i, x_j)_k = 0. \tag{C.122}$$

In this case we multiply (C.122) by a test function $\lambda_k(x_j)$ and integrate this function over the same range as in the action to obtain

$$\int dx \lambda_k(x_j) g_k(y_i, x_j) = 0. \tag{C.123}$$

Sometimes we deal with constraints that appear in the form of an integral

$$\int dx_j g_k(y_i, x_j) = const. \tag{C.124}$$

We may introduce any constant Lagrangian multiplier.

In either case we now extremise

$$\delta \int \left[\mathcal{L}\left(y_i, \frac{\partial yi}{\partial x_j}, x_j \right) + \sum_k \lambda_k g_k(y_i, x_j) \right] dx_j = 0. \tag{C.125}$$

The Lagrangian multiplier λ_k may depend on x_j when the constraint is of the form (C.122). This is what gets added, possibly with more than one constraint, to the action.

Treating the entire integrand as a new function

$$\tilde{\mathcal{L}}\left(y_i, \frac{\partial yi}{\partial x_j}, x_j \right) = \mathcal{L} + \int dx_j \sum_k \lambda_k(x_j) g_k(y_i, x_j) \tag{C.126}$$

This new function $\tilde{\mathcal{L}}$ satisfies the usual Euler-Lagrange equations

$$\frac{\partial \tilde{\mathcal{L}}}{\partial y_i} - \sum_j \frac{\partial}{\partial x_j} \frac{\partial \tilde{\mathcal{L}}}{\partial(\partial y_i / \partial x_j)} = 0 \tag{C.127}$$

for each dependent variable y_i. These equations and the constraint equations are then solved simultaneously.

C.3 Some Inequalities

Cauchy-Swartz

$$\left(\sum_i |x_i + y_i|^2\right)^{1/2} \leq \left(\sum_i |x_i^2\right)^{1/2} \left(\sum_i |y_i|^2\right)^{1/2} \tag{C.128}$$

Minkowski's inequality:

$$\left(\sum_i |x_i + y_i|^2\right)^{1/p} \leq \left(\sum_i |x_i^2\right)^{1/p} \left(\sum_i |y_i|^2\right)^{1/p} \tag{C.129}$$

Hölder inequality:

$$a\alpha + b\beta \geq a^\alpha b^\beta \tag{C.130}$$

Proof: Consider the function

$$f(x) = a\alpha e^{x/\alpha} + b\beta e^{-x/\beta}. \tag{C.131}$$

Taking the first derivative and setting the result to zero,

$$\frac{df}{dx}(x) = ae^{x/\alpha} - be^{-x/\beta} = 0. \tag{C.132}$$

This has the solution

$$e^x = \left(\frac{b}{a}\right)^{\alpha\beta}. \tag{C.133}$$

Taking the second derivative of C.131,

$$\frac{d^2 f}{dx^2}(x) = f(x) = \frac{a}{\alpha}e^{x/\alpha} + \frac{b}{\beta}e^{-x/\beta} > 0 \tag{C.134}$$

we see that we have a minimum, therefore

$$f(x=0) \geq f(x_{min}) = a\alpha\left(\frac{b}{a}\right)^\beta + b\beta\left(\frac{a}{b}\right)^\alpha = a^\alpha b^\beta(\alpha+\beta) = a^\alpha b^\beta \tag{C.135}$$

$$\frac{1}{p} + \frac{1}{q} = 1, \quad \frac{p+q}{pq} = 1, \quad p+q = pq \tag{C.136}$$

inequality:

$$|e^{ia} - e^{ib}| \tag{C.137}$$

Shwartz inequality

C.4 Curvilinear Coordinates

Techniques extremely close to those used in General Relativity, thus making the transition to the general theory easier.

The mathematics of vectors and tensors applies in GR much as it does in SR, with the restriction that vectors and tensors are defined independently at each spacetime point.

Definition In a coordinate change the dx^a transforms as a *contravariant* 4-vector

$$dx^{a\prime} = \frac{\partial dx^{a\prime}}{\partial x^b} dx^a \tag{C.138}$$

Similarly, if

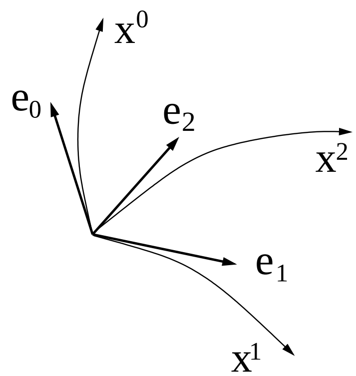


Figure C.1: Coordinate induced basis.

Under a coordinate transformation at the point **P** we pass from the basis $\{\mathbf{e}'_a\}$ to the basis $\{\mathbf{e}_a\}$ at **P**

$$\mathbf{e}'_a = a_a^{\ b}\mathbf{e}_b \tag{C.139}$$
$$\mathbf{e}_a = a'^{\ b}_{a}\mathbf{e}'_b, \tag{C.140}$$

$$\mathbf{e}'_a = a_a^{\ c}\mathbf{e}_c = a_a^{\ c}(a'^{\ b}_{c}\mathbf{e}'_b) = (a_a^{\ c}\, a'^{\ b}_{c})\mathbf{e}'_b \tag{C.141}$$

where in the second step we replaced $\mathbf{e}_c$ with $a'^{\ b}_{c}\mathbf{e}'_b$. Equation (C.141) which implies

$$a_a^{\ c}\, a'^{\ b}_{c} = \delta_a^{\ b} \tag{C.142}$$

i.e. $(a_a^{\ b})$ is the inverse matrix of $(a'^{\ b}_{a})$.

A natural basis for the tangent space T_p at point p is given by the partial derivatives with respect to the coordinates at that point. We denote this with the convenient notation

$$\hat{\mathbf{e}}_{(\mathbf{a})} = \partial_{\mathbf{a}}. \tag{C.143}$$

A basis for the cotangent space T_p^* is given by the gradients of the coordinate functions

$$\theta^{(\mathbf{a})} = \mathbf{dx^a} \tag{C.144}$$

C.5 Totally Antisymmetric symbol and Determinates

C.5.1 First Definition of Determinants

The determinant of 2 dimenional square matrices is defined as

$$\det \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = A_{11}A_{22} - A_{21}A_{12}. \tag{C.145}$$

The determinant of a 3-dimensional square matrix is defined by

$$\begin{aligned} \det \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} &= A_{11} \det \begin{pmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{pmatrix} - A_{12} \det \begin{pmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{pmatrix} \\ &\quad + A_{13} \det \begin{pmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix} \end{aligned} \tag{C.146}$$

the elements of thee top row being taken in order from left to right with alternate signs and multiplied by the determinants of 2-diemnsional matrices which remain when the row and column through the element are deleted.

The definition can be extended to determinants of square matrices of any order iteratively in a similar way; the determinant of an $n-$dimensional matrix is expressible in terms of determinants of $(n-1)-$dimension matrices formed from the original $n-$dimensional matrix.

C.5.2 The Levi-Civita Tensor

Determinants in $n-$dimensions can be written in a form making use of the Levi-Civita tensor defined by

$$\epsilon_{ij\ldots l} = \begin{cases} +1, & \text{if } ij\ldots l \text{ is an even permutation of } 12\ldots n \\ -1, & \text{if } ij\ldots l \text{ if an odd permutation of } 12\ldots n \\ 0, & \text{if any index is repeated.} \end{cases} \tag{C.147}$$

which simplifies the proof of basic properties of determinants.

It is easy to see for 2-dimensional matrices that the determinant can be written

$$\begin{aligned} \epsilon^{ij} A_{1i} A_{2j} &= \epsilon^{12} A_{11} A_{22} + \epsilon^{21} A_{12} A_{21} \\ &= \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \end{aligned} \tag{C.148}$$

Note that if we swap 1 and 2 around we get

$$\begin{aligned} \epsilon^{ij} A_{2i} A_{1j} &= \epsilon^{ij} A_{1j} A_{2i} \\ &= -\epsilon^{ji} A_{1j} A_{2i} \\ &= -\epsilon^{ij} A_{1i} A_{2j} \\ &= -\det(A). \end{aligned} \tag{C.149}$$

So we can write

$$\begin{aligned} 2\det(A) &= \epsilon^{ij} A_{1i} A_{2j} - \epsilon^{ij} A_{2i} A_{1j} \\ &= \epsilon^{ij} (A_{1i} A_{2j} - A_{2i} A_{1j}) \\ &= \epsilon^{ij} \epsilon^{kl} A_{ki} A_{lj} \end{aligned} \tag{C.150}$$

The determinant of a 3-dimensional matrix can be written

$$\det|\mathbf{A}| = \epsilon^{ijk} A_{1i} A_{2j} A_{3k} \tag{C.151}$$

We can easily see how this agrees with the iterative definition (C.146) by writing

$$\begin{aligned} \epsilon^{ijk} A_{1i} A_{2j} A_{3k} &= A_{11} \epsilon^{1jk} A_{2j} A_{3k} + A_{12} \epsilon^{2jk} A_{2j} A_{3k} + A_{13} \epsilon^{3jk} A_{2j} A_{3k} \\ &= A_{11} \epsilon^{1jk} A_{2j} A_{3k} - A_{12} \epsilon^{j2k} A_{2j} A_{3k} + A_{13} \epsilon^{jk3} A_{2j} A_{3k}. \end{aligned} \tag{C.152}$$

From the first term $\epsilon^{1jk}A_{2j}A_{3k}$; obviously elements from the first row of the matrix A are not included in this sum, and also because ϵ^{1jk} vanishes for $j = 1$ or $k = 1$ it doesn't include terms from the first column. Summation is taken over the remaining row and column indices subject to anti-symmetrisation giving the minor of A_{11}. Similarly argument holds for the other terms in (C.152).

This kind of argument generalises to $n-$dimensional determinants and we have

$$\det|\mathbf{A}| = \epsilon^{a_1 a_2 \ldots a_n} A_{1a_1} A_{2a_2} \ldots A_{na_n}. \tag{C.153}$$

We can also generalise (C.150) to $n-$dimensions. Consider 3-dimensions first

$$\det|\mathbf{A}| = +\epsilon^{ijk} A_{1i} A_{2j} A_{3k} = \epsilon^{123} \epsilon^{ijk} A_{1i} A_{2j} A_{3k} \tag{C.154}$$

$$\det|\mathbf{A}| = -\epsilon^{ijk} A_{1i} A_{2j} A_{3k} = \epsilon^{213} \epsilon^{ijk} A_{2i} A_{1j} A_{3k} \tag{C.155}$$

and so on for the other permutations of $1, 2, 3$. Since there are 3! ways of oredring $1, 2, 3$ we have

$$\det|\mathbf{A}| = \frac{1}{3!} \epsilon^{i'j'k'} \epsilon^{ijk} A_{i'i} A_{j'j} A_{k'k} \tag{C.156}$$

In $n-$dimensions

$$\det|\mathbf{A}| = \frac{1}{n!} \epsilon^{b_1 b_2 \ldots b_n} \epsilon^{a_1 a_2 \ldots a_n} A_{b_1 a_1} A_{b_2 a_2} \ldots A_{b_n a_n} \tag{C.157}$$

Proof of basic properties of determinants

This ϵ representation makes it easy to basic prove properties of determinates.

(a) Interchanging rows and columns results in the same determinant. This is the same as saying $\det(A^T) = \det(A)$ where A^T is the transpose of the matrix A. This follows easily follows from the equivalence of (C.157) with

$$\det|\mathbf{A}| = \epsilon^{b_1 b_2 \ldots b_n} A_{b_1 1} A_{b_2} \ldots A_{b_n n} \tag{C.158}$$

(similar to proving the equivalence of (C.153) and (C.157)) which is the same as

$$\epsilon^{b_1 b_2 \ldots b_n} (A^T)_{1b_1} (A^T)_{2b_2} \ldots (A^T)_{nb_n} = \det(A^T). \tag{C.159}$$

(b) Swapping columns introduces a minus sign. We need only prove that swapping any two adjacent columns introduces a minus sign, i.e.,

$$\det\begin{pmatrix} A_{11} & \dots & A_{1k} & A_{1k+1} & \dots & A_{1n} \\ A_{21} & \dots & A_{2k} & A_{2k+1} & \dots & A_{2n} \\ \vdots & \dots & \vdots & \dots & \ddots & \vdots \\ A_{n1} & \dots & A_{nk} & A_{nk+1} & \dots & A_{nn} \end{pmatrix} = -\det\begin{pmatrix} A_{11} & \dots & A_{1k+1} & A_{1k} & \dots & A_{1n} \\ A_{21} & \dots & A_{2k+1} & A_{2k} & \dots & A_{2n} \\ \vdots & \dots & \dots & \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nk+1} & A_{nk} & \dots & A_{nn} \end{pmatrix} \tag{C.160}$$

$$\begin{aligned} \epsilon_{ijk}A_{i1}A_{j2}A_{k3} &= \epsilon_{ijk}A_{j2}A_{i1}A_{k3} \\ &= \epsilon_{jik}A_{i2}A_{j1}A_{k3} \\ &= -\epsilon_{ijk}A_{i2}A_{j1}A_{k3} \end{aligned} \tag{C.161}$$

For $n-$dimensional determinants,

$$\begin{aligned} \epsilon_{a_1\dots a_k a_{k+1}\dots a_n}A_{a_1 1}\dots A_{a_k k}A_{a_{k+1}k+1}\dots A_{a_n n} &= \epsilon_{a_1\dots a_k a_{k+1}\dots a_n}A_{a_1 1}\dots A_{a_{k+1}k+1}A_{a_k k}\dots A_{a_n n} \\ &= \epsilon_{a_1\dots a_{k+1} a_k\dots a_n}A_{a_1 1}\dots A_{a_k k+1}A_{a_{k+1}k}\dots A_{a_n n} \\ &= -\epsilon_{a_1\dots a_k a_{k+1}\dots a_n}A_{a_1 1}\dots A_{a_k k+1}A_{a_{k+1}k}\dots A_{a_n n} \end{aligned} \tag{C.162}$$

Swapping rows introduces a minus sign. To see this first write $\det(A) = \det(A^T)$, swap adjacent columns of the matrix A^T to give another matrix, lets call it $(A')^T$, then $\det(A^T) = -\det((A')^T)$. Taking the transpose of $(A')^T$ gives a matrix A', this differs from A by the exchange of adjacent rows. We have altogether $\det(A) = -\det(A')$.

(c) Simply consequence of property (b) is that if two columns are the same the determinant vanishes. Similarly if two rows are the same then the vanishes.

(d) We have

$$\begin{aligned} \det\begin{pmatrix} A_{11} & \dots & A_{1k}+B_{1k} & \dots & A_{1n} \\ A_{21} & \dots & A_{2k}+B_{2k} & \dots & A_{2n} \\ \vdots & \ddots & \vdots & \ddots & \\ A_{n1} & \dots & A_{nk}+B_{nk} & \dots & A_n \end{pmatrix} &= \det\begin{pmatrix} A_{11} & \dots & A_{1k} & \dots & A_{1n} \\ A_{21} & \dots & A_{2k} & \dots & A_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nk} & \dots & A_{nn} \end{pmatrix} \\ &+ \det\begin{pmatrix} A_{11} & \dots & B_{1k} & \dots & A_{1n} \\ A_{21} & \dots & B_{2k} & \dots & A_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n1} & \dots & B_{nk} & \dots & A_{nn} \end{pmatrix} \end{aligned} \tag{C.163}$$

which are easily seen from ϵ

$$\begin{aligned}\epsilon_{a_1\ldots a_k\ldots a_n} A_{a_1 1}\cdots(A_{a_k k}+B_{a_k k})\cdots A_{a_n n} &= \epsilon_{a_1\ldots a_k\ldots a_n} A_{a_1 1}\cdots A_{a_k k}\cdots A_{a_n n} \\ &\quad + \epsilon_{a_1\ldots a_k\ldots a_n} A_{a_1 1}\cdots + B_{a_k k}\cdots A_{a_n n}. \end{aligned} \tag{C.164}$$

A similar result holds for rows and can be eablished by considering the determinant of the transpose of the matrix and using the analogous result for columns just proved, and then apply $\det(A^T)=\det(A)$ on the resulting two determinants.

(e) Multiplication of a column by a constant k

$$\det\begin{pmatrix} A_{11} & \ldots & kA_{1k} & \ldots & A_{13} \\ A_{21} & \ldots & kA_{2k} & \ldots & A_{23} \\ \vdots & \ldots & \vdots & \ddots & \vdots \\ A_{n1} & \ldots & kA_{nk} & \ldots & A_n \end{pmatrix} = k\det\begin{pmatrix} A_{11} & \ldots & A_{1k} & \ldots & A_{13} \\ A_{21} & \ldots & A_{2k} & \ldots & A_{23} \\ \vdots & \ldots & \vdots & \ddots & \vdots \\ A_{n1} & \ldots & A_{nk} & \ldots & A_n \end{pmatrix} \tag{C.165}$$

follows from

$$\epsilon_{a_1\ldots a_k\ldots a_n} A_{a_1 1}\cdots(kA_{a_k k})\cdots A_{a_n n} = k\epsilon_{a_1\ldots a_k\ldots a_n} A_{1a_1}A_{2a_2}A_{na_n} \tag{C.166}$$

Multiplication of a row by a constant k

$$\epsilon_{a_1 a_2\ldots a_n}(kA_{1a_1})A_{2a_2}A_{na_n} = k\epsilon_{a_1 a_2\ldots a_n}A_{1a_1}A_{2a_2}A_{na_n} \tag{C.167}$$

In particular $\det(kA)=k^n\det(A)$.

(f) The determinant of the product of two matrices A and B is equal to the product of the determinant of A and the determinant of B, that is, $\det(AB)=\det(A)\det(B)$.

$$\begin{aligned}\det|\mathbf{A}| &= \frac{1}{n!}\epsilon^{b_1 b_2\ldots b_n}\epsilon^{a_1 a_2\ldots a_n}(A_{b_1c_1}B_{c_1a_1})(A_{b_2c_2}B_{c_2a_2})\ldots(A_{b_nc_n}B_{c_3a_3}) \\ &= \frac{1}{n!}\sum_{perms\ of\ c}(\epsilon^{b_1b_2\ldots b_n}A_{b_1c_1}A_{b_2c_2}\ldots A_{b_nc_n})(\epsilon^{a_1a_2\ldots a_n}B_{c_1a_1}B_{c_2a_2}\ldots B_{c_3a_3}) \\ &= \frac{1}{n!}\sum_{perms\ of\ c}\det(A^T)\det(B) \\ &= \det(A)\det(B). \end{aligned} \tag{C.168}$$

C.5.3 Inverse Matrix

Demonstarte ideas with 3-dimensionals marices. Recal the determinant can be expanded in the frist terms of co-factors.

We show that the elements of middle row being taken in order from left to right with alternate signs and multiplied by the determinants of 2-diemnsional matrices which remain when the row and column through the element are deleted, except here the first term starts with a minus sign.

$$\begin{aligned} \det|\mathbf{A}| &= \epsilon^{ijk} A_{1i} A_{2j} A_{3k} \\ &= A_{21}\epsilon^{i1k} A_{1i} A_{3k} + A_{22}\epsilon^{i2k} A_{1i} A_{3k} + A_{23}\epsilon^{i3k} A_{1i} A_{3k} \\ &= -A_{21}\epsilon^{1ik} A_{1i} A_{3k} + A_{22}\epsilon^{i2k} A_{1i} A_{3k} + A_{23}\epsilon^{ik3} A_{1i} A_{3k} \end{aligned} \tag{C.169}$$

or

$$\det A = -A_{21} \begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix} + A_{22} \begin{vmatrix} A_{11} & A_{31} \\ A_{13} & A_{33} \end{vmatrix} - A_{23} \begin{vmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{vmatrix} \tag{C.170}$$

The determinant

$$-\begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix}$$

is called the cofactor of the matrix element A_{21}.

To find the inverse.

For a square matrix $\mathbf{A}$, form the matrix whose elements are the cofactors of $\mathbf{A}$, i.e.

$$\begin{pmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{pmatrix} \tag{C.171}$$

Consider the product of the matrix $\mathbf{A}$ with the matrix $\mathbf{C}^T$.

$$\mathbf{AC}^T = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{pmatrix} \tag{C.172}$$

The product of the first column of $\mathbf{C}$ and the first row of $\mathbf{A}$ is

$$A_{11}C_{11} + A_{12}C_{12} + A_{13}C_{13} = \det \mathbf{A}.$$

Generally, the product of any row of $\mathbf{A}$ with its cofactors is equal to $\mathbf{C}$,

$$\det A = \sum_k A_{ik} C_{ik} \quad \text{no summation over } i. \tag{C.173}$$

Other products give zero, for example,

$$\begin{aligned} A_{11}C_{21} + A_{12}C_{22} + A_{13}C_{23} &= A_{11} \begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix} - A_{12} \begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} + A_{13} \begin{vmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{vmatrix} \\ &= \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{11} & A_{12} & A_{13} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} \\ &= 0 \end{aligned} \tag{C.174}$$

We can see general why the other products are zero; if we replace A_{ik} in the sum in (C.173) by A_{jk}, with $j \neq i$, it gives a new deterninant which can be obtained from the determinant of A by replacing its $i-th$ row with the row $(A_{j1}, A_{j2}, A_{j3},)$. Then as this row will be repeated twice and the resulting derterminant vanishes. In general we have

$$\sum_k A_{ik} C_{jk} = A_{i1}C_{j1} + A_{i2}C_{j2} + A_{i3}C_{j3} = \delta_{ij} \det A \tag{C.175}$$

Or

$$\mathbf{A}\mathbf{C}^T = \mathbf{I} \det \mathbf{A}. \tag{C.176}$$

Same with columns.

$$\begin{aligned} \det |\mathbf{A}| &= \epsilon^{ijk} A_{i1} A_{j2} A_{k3} \\ &= A_{11} \epsilon^{1jk} A_{j2} A_{k3} + A_{21} \epsilon^{2jk} A_{j2} A_{k3} + A_{31} \epsilon^{3jk} A_{j2} A_{k3} \\ &= A_{11} \epsilon^{1jk} A_{j2} A_{k3} - A_{21} \epsilon^{j2k} A_{j2} A_{k3} + A_{31} \epsilon^{jk3} A_{j2} A_{k3} \end{aligned} \tag{C.177}$$

$$\det A = A_{11} \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} - A_{21} \begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix} + A_{31} \begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix} \tag{C.178}$$

$$\begin{aligned} \det |\mathbf{A}| &= \epsilon^{ijk} A_{i1} A_{j2} A_{k3} \\ &= \epsilon^{ijk} A_{j2} A_{i1} A_{k3} \\ &= A_{12} \epsilon^{i1k} A_{i1} A_{k3} + A_{22} \epsilon^{i2k} A_{i1} A_{k3} + A_{32} \epsilon^{i3k} A_{i1} A_{k3} \\ &= -A_{12} \epsilon^{1ik} A_{i1} A_{k3} + A_{22} \epsilon^{i2k} A_{i1} A_{k3} - A_{32} \epsilon^{ik3} A_{i1} A_{k3} \end{aligned} \tag{C.179}$$

Consider the product of the matrix $\mathbf{A}$ with the matrix $\mathbf{C}^T$.

$$\mathbf{C}^T\mathbf{A} = \begin{pmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \tag{C.180}$$

$$\mathbf{C}^T\mathbf{A} = \mathbf{I} \det \mathbf{A}. \tag{C.181}$$

Therefore from this and (C.176) we have

$$\mathbf{A}^{-1} = \frac{\mathbf{C}^T}{\det \mathbf{A}}. \tag{C.182}$$

This easily extends to higher dimensional square matrices. For any square matrix $\mathbf{A}$, the inverse matrix, $\mathbf{A}^{-1}$, exists if and only if $|\mathbf{A}| \neq 0$ and in this case

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\mathbf{C}^{\mathbf{T}}. \tag{C.183}$$

Or in component form

$$(\mathbf{A}^{-1})_{ij} = \frac{1}{\det \mathbf{A}}(\mathbf{C})_{ji}. \tag{C.184}$$

C.5.4 Cross Product, Area and Volume Vector Equations

Consider the vector formed by

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \tag{C.185}$$

which means

$$\vec{a} \times \vec{b} = (a_2 b_3 - b_2 a_3)\hat{\imath} - (a_1 b_3 - b_1 a_3)\hat{\jmath} + (a_1 b_2 - b_1 a_2)\hat{k}. \tag{C.186}$$

The dot product of $\vec{a}$ and $\vec{a} \times \vec{b}$ is

$$\vec{a} \cdot (\vec{a} \times \vec{b}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = 0. \tag{C.187}$$

Similarly $\vec{b}\cdot(\vec{a}\times\vec{b})=0$ therefore $\vec{a}\times\vec{b}$ is a vector perpendicular to the plane containing $\vec{a}$ and $\vec{b}$. Also note that

$$\vec{a}\times\vec{b}=-\vec{b}\times\vec{a}. \tag{C.188}$$

We consider the length of the cross product. We can simplify things by writing $\vec{a}=|\vec{a}|\hat{\imath}$ and $\vec{b}=|\vec{b}|\cos\theta\hat{\imath}+|\vec{b}|\sin\theta\hat{\jmath}$, then

$$|\vec{a}\times\vec{b}|=|\begin{vmatrix}\hat{\imath} & \hat{\jmath} & \hat{k}\\ |\vec{a}| & 0 & 0\\ |\vec{b}|\cos\theta & |\vec{b}|\sin\theta & 0\end{vmatrix}|=|\vec{a}||\vec{b}|\sin\theta. \tag{C.189}$$

The length is then

$$|\vec{a}\times\vec{b}|=|\vec{a}||\vec{b}|\sin\theta. \tag{C.190}$$

This can be written as general vector form

$$\begin{aligned}|\vec{a}\times\vec{b}| &= \sqrt{|\vec{a}|^2|\vec{b}|^2\sin^2\theta}\\ &= \sqrt{|\vec{a}|^2|\vec{b}|^2(1-\cos^2\theta)}\\ &= \sqrt{|\vec{a}|^2|\vec{b}|^2-(\vec{a}\cdot\vec{b})^2}\end{aligned} \tag{C.191}$$

and so holds in general.

The area of a paralleagram, see fig (C.5.4) (a), is equal to the base $\times$ height $=|\vec{a}|h=|\vec{a}||\vec{b}|\sin\theta=|\vec{a}\times\vec{b}|$. As we have an expression for the area in vector notation, we may orientate the parallelagram any way we want and the arae will still be given by the vector equation

$$Area=|\vec{a}\times\vec{b}| \tag{C.192}$$

The volume of a parallepiped is, see fig (C.5.4) (b), is equal to the area of base times height $=|\vec{a}\times\vec{b}|h=|\vec{a}\times\vec{b}||\vec{c}|\cos\theta=(\vec{a}\times\vec{b})\cdot\vec{c}$, where θ is the angle between $\vec{c}$ and $\vec{h}$ an so also the angle between $\vec{a}\times\vec{b}$ and $\vec{c}$. As we have an expression for the area in vector notation, we may orientate the parallelepiped any way we want and the arae will still be given by the vector equation

$$Volume=(\vec{a}\times\vec{b})\cdot\vec{c} \tag{C.193}$$

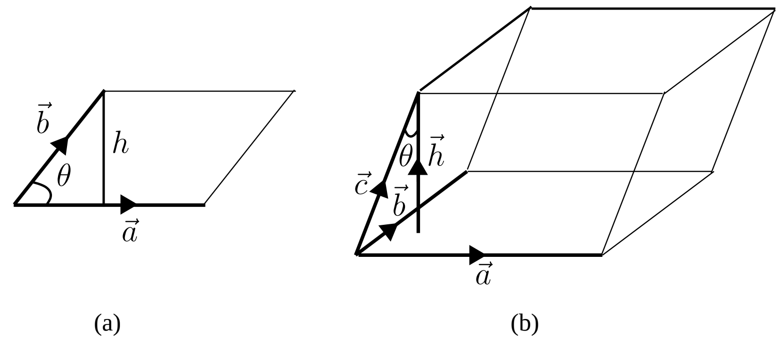


Figure C.2: (a) Area of a parallelagram written in vector form: $|\vec{a} \times \vec{b}|$ (b) Volume of a parallepiped written in vector form: $(\vec{a} \times \vec{b}) \cdot \vec{c}$.

C.5.5 Determinant as Volume Factor

The transformation defined by

$$M = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \tag{C.194}$$

maps $(0,0)$ to $(0,0)$, $(1,0,0)$ to (a_1, a_2), $(0,1)$ to (b_1, b_2) and maps $(1,1)$ to $(a_1 + b_1, a_2 + b_2)$. This gives a parallelagram in the image space - see fig (C.5.5)

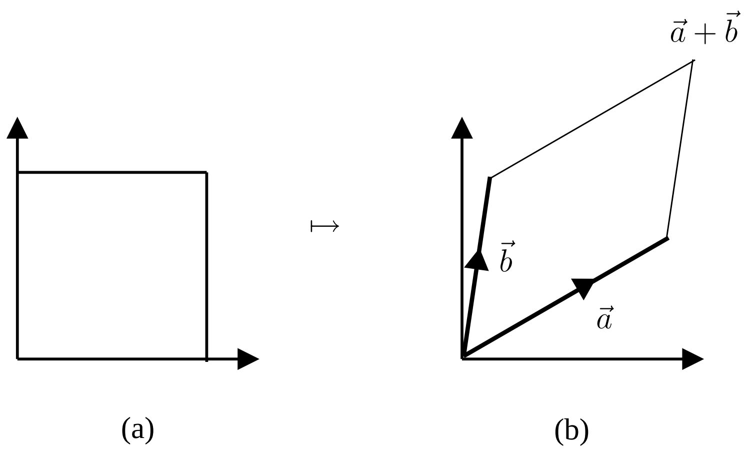


Figure C.3: Area of a parallelagram.

From (C.192) the area of the parallelagram can be written

$$
\begin{aligned}
|\vec{a} \times \vec{b}| &= | \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ a_1 & a_2 & 0 \\ b_1 & b_2 & 0 \end{vmatrix} | \\
&= | \begin{vmatrix} a_2 & 0 \\ b_2 & 0 \end{vmatrix} \hat{i} - \begin{vmatrix} a_1 & 0 \\ b_1 & 0 \end{vmatrix} \hat{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \hat{k} | \\
&= \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}
\end{aligned}
\tag{C.195}
$$

M alters the area by a factor of $\det M$.

The transformation defined by

$$
M = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}
\tag{C.196}
$$

maps a cube of unit volume in xyz space to a parallelepiped in the image space.

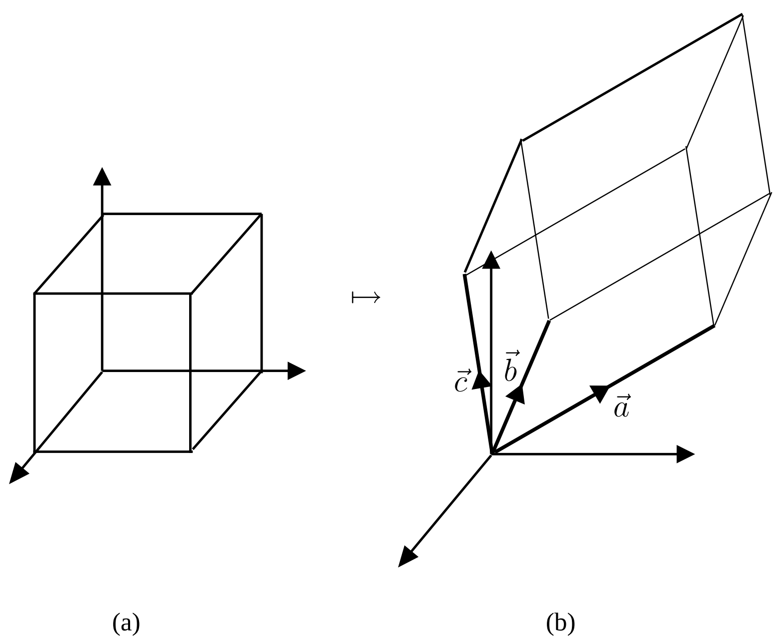


Figure C.4: Volume of a parallepiped.

The volume of the parallelepiped

$$
\begin{aligned}
\vec{a}\cdot\vec{b}\times\vec{c} &= (a_1\hat{i}+a_2\hat{j}+a_3\hat{k})\cdot\begin{vmatrix}\hat{i} & \hat{j} & \hat{k}\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix}\\
&= a_1\begin{vmatrix} b_2 & b_3\\ c_2 & c_3\end{vmatrix} - a_2\begin{vmatrix} b_1 & b_3\\ c_1 & c_3\end{vmatrix} + a_3\begin{vmatrix} b_1 & b_2\\ c_1 & c_2\end{vmatrix}\\
&= \begin{vmatrix} a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix}
\end{aligned}
\tag{C.197}
$$

M alters the volume by a factor of $\det M$.

We use these ideas in deriving the the volume element dV in other coordinate systems.

C.5.6 Generalisation to $n-$dimensions

How does this work in n dimensions. Given n vectors

$$
\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n.
$$

What volume do they span? If a transformation M maps a cube of unit volume in $d-$dimensional Euclidean space to the image space does the resulting object have volume given by $\det M$?

We assign "area" in $\mathbb{R}^2$ to the parallelogram defined by two vectors $\mathbf{x}, \mathbf{y}$. This is some function, $\omega(\mathbf{x}, \mathbf{y})$, of the two vectors $\mathbf{x}$ and $\mathbf{y}$. This should have the following properties:

i) If we double the length of one of the vectors, the area should double, i.e.

$$
\omega(\mu\mathbf{x}, \nu\mathbf{y}) = (\mu\nu)\omega(\mathbf{x}, \mathbf{y}).
$$

(note that if ν or μ have a negative sign, we have reversed the direction of a vector).

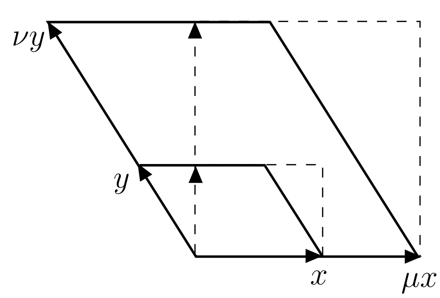


Figure C.5: areaofPar.

ii) satisfies

$$\omega(\mathbf{x_1} + \mathbf{x_2}, \mathbf{y}) = \omega(\mathbf{x_1}, \mathbf{y}) + \omega(\mathbf{x_2}, \mathbf{y}), \tag{C.198}$$

as is proved in fig (C.5.6). We may similarly prove that

$$\omega(\mathbf{x}, \mathbf{y_1} + \mathbf{y_2}) = \omega(\mathbf{x}, \mathbf{y_1}) + \omega(\mathbf{x}, \mathbf{y_2}), \tag{C.199}$$

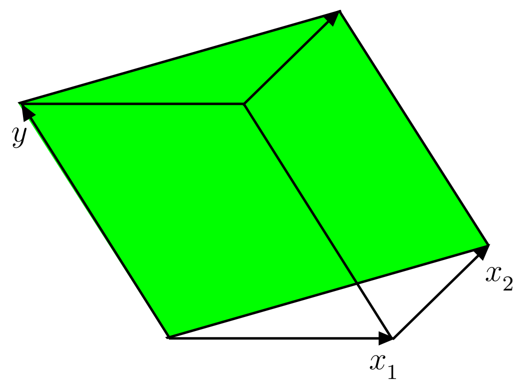


Figure C.6: areaofPar2. Illustrating proof of $\omega(\mathbf{x_1} + \mathbf{x_2}, \mathbf{y}) = \omega(\mathbf{x_1}, \mathbf{y}) + \omega(\mathbf{x_2}, \mathbf{y})$.

iii) If two sides coincide, the area should be zero. Thus

$$\omega(\mathbf{x}, \mathbf{x}) = 0.$$

These are bilinear form. The third shows that it must be skew-symmetric

$$\begin{aligned} 0 &= \omega(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) \\ &= \omega(\mathbf{x}, \mathbf{x}) + \omega(\mathbf{x}, \mathbf{y}) + \omega(\mathbf{y}, \mathbf{x}) + \omega(\mathbf{y}, \mathbf{y}) \\ &= \omega(\mathbf{x}, \mathbf{y}) + \omega(\mathbf{y}, \mathbf{x}) \end{aligned}$$

So

$$\omega(\mathbf{x}, \mathbf{y}) = -\omega(\mathbf{y}, \mathbf{x}). \tag{C.200}$$

These are exactly the properties of a 2-form. Similarly, a volume element is associated with 3-form. The orientaion of set of vectors is a choice of order in which we write them. If we interchange two vectors, the orientation changes. The algebra of forms is organised in a way that naturally reflects the orientation of the parallelagram. Notice that with respect to this regard that the area of the parallelagram given by the usual formula

$$area = \|\mathbf{x} \times \mathbf{y}\|$$

does not have the property of changing sign under interchange of the two vectors $\mathbf{x}$ and $\mathbf{y}$. This is the reason for calling $\omega(\mathbf{x}, \mathbf{y})$ the "area" (this together with the fact that we have not included a normaisation).

If we demand

1. If any two input vectors are identical then $\omega_n = 0$, i.e.,

$$\omega_n(\vec{v}_1, \ldots, \vec{w}, \ldots, \vec{w}, \ldots, \vec{v}_n) = 0. \tag{C.201}$$

2. If $n-1$ of the input vectors are fixed then ω_n is linear in the remaining variable.

3. $\omega_n(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) = 1$.

There is one and only one function ω_n which satisfies properties 1 to 3 above.

The function $\omega_n(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n)$ turns out to be the determinant of the $n \times n$ matrix formed from the vectors $(\vec{v}_1, \ldots, \vec{v}_n)$, namely

$$\begin{aligned}
\det \mathbf{A}_n &:= \det[\vec{v}_1, \ldots, \vec{v}_n] \\
&= \det \begin{pmatrix} (\vec{v}_1)^1 & \ldots & (\vec{v}_j)^1 & \cdots & (\vec{v}_n)^1 \\ & & & & \\ \vdots & \cdots & \vdots & & \\ & & & & \\ (\vec{v}_1)^{n-1} & \ldots & (\vec{v}_j)^{n-1} & \cdots & (\vec{v}_n)^{n-1} \\ (\vec{v}_1)^n & \ldots & (\vec{v}_j)^n & \cdots & (\vec{v}_n)^n \end{pmatrix}
\end{aligned}$$

It is easily verified that this determinant does indeed have the above properties. Property 1. follows from the fact that if two columns of the matrix $\mathbf{A}_n$ are the same the determinant vanishes. Property 2. follows from that fact that we have for determinants

$$
\begin{aligned}
& \det\left[\vec{v}_1,\ldots,\vec{v}_j+\vec{w}_j,\ldots,\vec{v}_n\right] \\
= & \det\begin{pmatrix} (\vec{v}_1)^1 & \ldots & (\vec{v}_j)^1+(\vec{w}_j)^1 & \cdots & (\vec{v}_n)^1 \\ \vdots & \ldots & \vdots & & \\ (\vec{v}_1)^{n-1} & \ldots & (\vec{v}_j)^{n-1}+(\vec{w}_j)^{n-1} & \cdots & (\vec{v}_n)^{n-1} \\ (\vec{v}_1)^n & \ldots & (\vec{v}_j)^n+(\vec{w}_j)^n & \cdots & (\vec{v}_n)^n \end{pmatrix} \\
= & \det\begin{pmatrix} (\vec{v}_1)^1 & \ldots & (\vec{v}_j)^1 & \cdots & (\vec{v}_n)^1 \\ \vdots & \ldots & \vdots & & \\ (\vec{v}_1)^{n-1} & \ldots & (\vec{v}_j)^{n-1} & \cdots & (\vec{v}_n)^{n-1} \\ (\vec{v}_1)^n & \ldots & (\vec{v}_j)^n & \cdots & (\vec{v}_n)^n \end{pmatrix} + \det\begin{pmatrix} (\vec{v}_1)^1 & \ldots & (\vec{w}_j)^1 & \cdots & (\vec{v}_n)^1 \\ \vdots & \ldots & \vdots & & \\ (\vec{v}_1)^{n-1} & \ldots & (\vec{w}_j)^{n-1} & \cdots & (\vec{v}_n)^{n-1} \\ (\vec{v}_1)^n & \ldots & (\vec{w}_j)^n & \cdots & (\vec{v}_n)^n \end{pmatrix} \\
= & \det\left[\vec{v}_1,\ldots,\vec{v}_j,\ldots,\vec{v}_n\right] + \det\left[\vec{v}_1,\ldots,\vec{w}_j,\ldots,\vec{v}_n\right].
\end{aligned}
$$

The vectors $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ can be choosen so that:

$$
\vec{e}_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \quad \ldots \quad, \quad \vec{e}_n = \begin{pmatrix}0\\0\\\vdots\\1.\end{pmatrix}
$$

Property 3. follows from

$$
\det\left[\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right] = \det\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix} = 1.
$$

We now prove that the only function which satisfies properties 1 to 3 above is this determinant.

Proof:

C.5.7 Recursive Definition of Volume

We define the n-dimensional volume

$$
V_{\vec{v}_1,\ldots,\vec{v}_n} := V_{\vec{v}_1,\ldots,\vec{v}_{n-1}} \|\vec{h}_n\| \tag{C.202}
$$

of the parallelepiped $P_{\vec{v}_1,\ldots,\vec{v}_n}$ recursively as the $(n-1)$-dimensional volume $V_{\vec{v}_1,\ldots,\vec{v}_{n-1}}$ multiplied by the length

$$\|\vec{h}_n\| = \left(\vec{h}_n^T \vec{h}_n\right)^{1/2} \tag{C.203}$$

of the vector

$$\vec{h}_n = \vec{v}_n - P_{\mathcal{W}_{n-1}}\ \vec{v}_n \tag{C.204}$$

that is orthogonal to the vector sunspace $\mathcal{W}_{n-1}$ spanned by the vectors $\vec{v}_1, \ldots, \vec{v}_{n-1}$

$$\mathcal{W}_{n-1} := \text{span}\{\vec{v}_1, \ldots, \vec{v}_{n-1}\}. \tag{C.205}$$

If $n = 1$, the volume $V_{\vec{v}_1}$ is defined as $\|\vec{v}_1\|$, which is the lebgth of $\vec{v}_1$.

For example, in fig () $V_{\vec{v}_1\vec{v}_2} = \|\vec{h}_2\| V_{\vec{v}_1} = \|\vec{h}_2\| \|\vec{a}_1\|$.

The recursive equation (C.202) defining the volume can be solved immediately for an n-dimensional parallelepiped $P_{\vec{h}_1,\ldots,\vec{h}_n}$ defined by n mutually orthogonal vectors $\vec{h}_1, \ldots, \vec{h}_n \in \mathbb{E}^n$:

$$V_{\vec{h}_1,\ldots,\vec{h}_n} := \|\vec{h}_1\| \cdots \|\vec{h}_n\|. \tag{C.206}$$

We prove by induction on n that the volume of an n-parallelepiped obeys the equation

$$\left(V_{\vec{h}_1,\ldots,\vec{h}_n}\right)^2 = \det\left[\mathbf{A}_n^T \mathbf{A}_n\right] \tag{C.207}$$

in which $\mathbf{A}_n$ is an $n \times n$ real matrix, the columns of which are the vectors that define the n-parallelepiped:

$$\mathbf{A}_n := [\vec{v}_1, \ldots, \vec{v}_n]. \tag{C.208}$$

If the list $[\vec{v}_1, \ldots, \vec{v}_n]$ is linearity dependent, that $\mathbf{A}_n$ is singular. It follows that $\mathbf{A}_n^T \mathbf{A}$ is singualr, and therefore that both the determinant and the volume must vanish. We assume from now on that $[\vec{v}_1, \ldots, \vec{v}_n]$ is linearly independent.

If $n = 1$, then $\mathbf{A}_1 = \vec{v}_1$ is an $n \times 1$ column vector and $\mathbf{A}_n^T \mathbf{A}_n = \vec{v}_l^T \vec{v}_l = \|\vec{v}_l\|^2$. Therefore the result holds for $n = 1$. We now assume that (C.207) holds for $n - 1$. The determinat in (C.207) is the determinant of the canoniocal inner product

$$
\det\left[\mathbf{A}_n^T\mathbf{A}_n\right] = \det\begin{pmatrix} (\vec{v}_1)^1 & \dots & (\vec{v}_1)^{n-1} & (\vec{v}_1)^n \\ \vdots & \dots & \vdots & \\ (\vec{v}_i)^1 & \dots & (\vec{v}_i)^{n-1} & (\vec{v}_i)^n \\ \vdots & \dots & \vdots & \\ (\vec{v}_n)^1 & \dots & (\vec{v}_n)^{n-1} & (\vec{v}_n)^n \end{pmatrix}\begin{pmatrix} (\vec{v}_1)^1 & \dots & (\vec{v}_j)^1 & \cdots & (\vec{v}_n)^1 \\ \vdots & \dots & \vdots & & \\ (\vec{v}_1)^{n-1} & \dots & (\vec{v}_j)^{n-1} & \cdots & (\vec{v}_n)^{n-1} \\ (\vec{v}_1)^n & \dots & (\vec{v}_j)^n & \cdots & (\vec{v}_n)^n \end{pmatrix}
$$

$$
\begin{aligned}
&= \det\left[\vec{v}_i^T\vec{v}_j\right] \\
&= \det\left[\vec{v}_i\cdot\vec{v}_j\right]. \qquad \text{(C.209)}
\end{aligned}
$$

The k-th column of $\mathbf{A}_n^T\mathbf{A}_n$ is equal to

$$
b_k = \begin{pmatrix} \vec{v}_1\cdot\vec{v}_k \\ \vdots \\ \vec{v}_{n-1}\cdot\vec{v}_k \\ \vec{v}_n\cdot\vec{v}_k. \end{pmatrix} \qquad \text{(C.210)}
$$

Define the vector $\vec{h}_n$ as in (C.204), and take $[\vec{v}_1, \cdots \vec{v}_{n-1}]$ as a basis of $\mathcal{W}_{n-1}$. Then

$$
\begin{aligned}
\vec{v}_n &= \vec{h}_n + P_{\mathcal{W}_{n-1}}\,\vec{h}_n \\
&= \vec{h}_n + \sum_{l=1}^{n} a_m^l\vec{v}_l. \qquad \text{(C.211)}
\end{aligned}
$$

The scalars a_m^l are the contravariant components of $P_{\mathcal{W}_{n-1}}\,\vec{h}_n$ relative to the basis choosen we have choosen in $\mathcal{W}_{n-1}$.

The nth-th column of $\mathbf{A}_n^T\mathbf{A}_n$ is equal to

$$
\begin{aligned}
\mathbf{b}_k &= \begin{pmatrix} \vec{v}_1\cdot\vec{v}_n \\ \vdots \\ \vec{v}_{n-1}\cdot\vec{v}_n \\ \vec{v}_n\cdot\vec{v}_n \end{pmatrix} \\
&= \|\vec{h}_n\|^2\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} + \sum_{l=1}^{n-1} a_m^l\begin{pmatrix} \vec{v}_1\cdot\vec{v}_l \\ \vdots \\ \vec{v}_{n-1}\cdot\vec{v}_l \\ \vec{v}_n\cdot\vec{v}_l \end{pmatrix} \\
&= \|\vec{h}_n\|^2\mathbf{e_n} + \sum_{l=1}^{n-1} a_m^l\mathbf{b}_l \qquad \text{(C.212)}
\end{aligned}
$$

where we have inserted (C.211) for $\vec{v}_n$. The term $\|\vec{h}_n\|^2\mathbf{e_n}$ occurs because

$$\vec{v}_i \cdot \vec{h}_n = \delta_n^i \vec{v}_n \cdot \vec{h}_n = \delta_n^i \|\vec{h}_n\|^2$$

By (C.212), the determinant of $\mathbf{A}_n^T\mathbf{A}_n$

$$\begin{aligned}
\det\left[\mathbf{A}_n^T\mathbf{A}_n\right] &= \det[\mathbf{b}_1, \ldots, \mathbf{b}_n] \\
&= \det\left[\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}, \|\vec{h}_n\|^2\mathbf{e}_n + \sum_{l=1}^{n-1} a_m^l \mathbf{b}_l\right] \\
&= \|\vec{h}_n\|^2 \det\left[\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}, \mathbf{e}_n\right] \qquad \text{(C.213)}
\end{aligned}$$

where we have used that none of the terms in the summation contribute to the determinant because each column $\mathbf{b}_l$ occurs twice. Expanding the determinant in minors of the last column, which consists of zeros except for a 1 in the last row, we obtain

$$\det\left[\mathbf{A}_n^T\mathbf{A}_n\right] = \|\vec{h}_n\|^2 \det\left[\mathbf{A}_{n-1}^T\mathbf{A}_{n-1}\right] \qquad \text{(C.214)}$$

in which $\mathbf{A}_{n-1}$ is the matrix of the vectors which span $\mathcal{W}_{n-1}$,

$$\mathbf{A}_{n-1} := [\vec{v}_1, \ldots, \vec{v}_{n-1}]. \qquad \text{(C.215)}$$

Then

$$\det\left[\mathbf{A}_{n-1}^T\mathbf{A}_{n-1}\right] = \left(V_{\vec{v}_1,\ldots,\vec{v}_{n-1}}\right)^2. \qquad \text{(C.216)}$$

Equation (C.207) follows from the recursive definition of the volume of an n-parallelepiped, (C.202).

It follows from this that the volume for an n-parallelepiped is

$$V_{\vec{v}_1,\ldots,\vec{v}_n} = |\det\left[\mathbf{A}_n\right]| \qquad \text{(C.217)}$$

in n-dimensional Euclidean space.

C.5.8 Linear Mappings and Volume Change

A non-singular linear mapping $A : \mathbb{E}^n \to \mathbb{E}^n$ transforms an n-parallelepiped into another n-parallelepiped $P_{\mathbf{A}\vec{v}_1,\ldots,\mathbf{A}\vec{v}_n}$, in which $\mathbf{A}$ is the $n \times n$ matrixthat realises A. There is a simply formula for the transformation of the volume $V_{\vec{v}_1,\ldots,\vec{v}_n}$ under $\mathbf{A}$:

$$V_{\mathbf{A}\vec{v}_1,\ldots,\mathbf{A}\vec{v}_n} = |\det[\mathbf{A}\mathbf{A}_n]| = |\det[\mathbf{A}]| \, |\det[\mathbf{A}_n]| \, . \tag{C.218}$$

It follows that the absolute value of the determinant of a matrix $\mathbf{A}$ is equal to the volume of the image of any n-parallelepiped to its original volume:

$$|\det[\mathbf{A}]| = \frac{V_{\mathbf{A}\vec{v}_1,\ldots,\mathbf{A}\vec{v}_n}}{V_{\vec{v}_1,\ldots,\vec{v}_n}} \tag{C.219}$$

C.5.9 Jacobian of Coordinate transformations

We have observed that the determinant is equal to the ratio of volumes before and after a linear mapping underlies the Jacobian used when changing variables in a multidimensional integral over a region R, such as

$$\int_R f(\mathbf{r}) d^n\mathbf{r}, \tag{C.220}$$

Here we give a heuristic derivation of the Jacobian formula. We begin by dividing the region of integration into parallelepiped bounded by infinitesimal displacements $d\mathbf{r}_1, \ldots, d\mathbf{r}_n$ along the directions of the coordinates $x^1, \ldots, x^n$. Because there are n coordinates, the location $\mathbf{r}_\mathbf{i}$ of each paralllelepiped can be indexed by a list of n integers,

$$\mathbf{i} := (i_1, \ldots, i_n). \tag{C.221}$$

The volume of the parallelepiped at $\mathbf{r}_\mathbf{i}$ is

$$V_{d\mathbf{s}_1,\ldots,d\mathbf{s}_n}(\mathbf{r}_\mathbf{i}) = |\det[d\mathbf{s}_1, \ldots, d\mathbf{s}_n]| \, . \tag{C.222}$$

A Riemann sumthat approximates the integral is

$$\int_R f(\mathbf{r}) d^n\mathbf{r} \approx \sum_{\mathbf{r}_\mathbf{i}} f(\mathbf{r}_\mathbf{i}) V_{d\mathbf{s}_1,\ldots,d\mathbf{s}_n}(\mathbf{r}_\mathbf{i}). \tag{C.223}$$

Let us change to a new set of coordinates $x'^1, \ldots, x'^n$ by expressing the old coordinates

$$x^k = \phi^k({x'}^1, \dots, {x'}^n) \tag{C.224}$$

and use the same points $\mathbf{r_i}$ (expressed in term of the new coordinates) to evaluate the function f. Although the mapping of coordinates may be nonlinear, the mapping of infintesimal displacements

$$d\mathbf{s} = \begin{pmatrix} dx^1 \\ \vdots \\ dx^n \end{pmatrix} \tag{C.225}$$

is linear. If one uses the chain rule of partial differentiation in the form

$$dx^k = \sum_{l=1}^{n} \frac{\partial \phi^k}{\partial {x'}^l} d{x'}^l, \tag{C.226}$$

then one recongnises a linear mapping

$$d\mathbf{s}' = \mathbf{J} d\mathbf{s} \tag{C.227}$$

the matrix of which is

$$\mathbf{J} = \begin{pmatrix} \frac{\partial \phi^1}{\partial {x'}^1} & \cdots & \frac{\partial \phi^1}{\partial {x'}^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi^n}{\partial {x'}^1} & \cdots & \frac{\partial \phi^n}{\partial {x'}^n} \end{pmatrix} \tag{C.228}$$

$\mathbf{J}$ is called the Jacobian matrix.

The volume of an infintesimal parallelepipedin the new coordinates,

$$\begin{aligned} V_{d\mathbf{s}'_1, \dots, d\mathbf{s}'_n} &= \left|\det\left[d\mathbf{s}'_1, \dots, d\mathbf{s}'_n\right]\right| \\ &= \left|\det\left[\mathbf{J} d\mathbf{s}_1, \dots, \mathbf{J} d\mathbf{s}_n\right]\right| \\ &= \left|\det\left[\mathbf{J}\right]\right| \left|\det\left[d\mathbf{s}_1, \dots, d\mathbf{s}_n\right]\right| \\ &= \left|\det\left[\mathbf{J}\right]\right| V_{d\mathbf{s}_1, \dots, d\mathbf{s}_n} \end{aligned} \tag{C.229}$$

In terms of the new coordinates, the Riemann sum is

$$\begin{aligned} \int_R f(\mathbf{r}) d^n r &\approx \sum_{\mathbf{i}} f(\phi(\mathbf{r}'_{\mathbf{i}})) \frac{V_{d\mathbf{s}_1, \dots, d\mathbf{s}_n}}{V_{d\mathbf{s}'_1, \dots, d\mathbf{s}'_n}} V_{d\mathbf{s}'_1, \dots, d\mathbf{s}'_n} \\ &\approx \sum_{\mathbf{i}} f(\phi(\mathbf{r}'_{\mathbf{i}})) \left|\det\left[\mathbf{J}\right]\right| V_{d\mathbf{s}'_1, \dots, d\mathbf{s}'_n} \end{aligned} \tag{C.230}$$

In the limit as one refines the Riemann sum, the expression in the second line converges to

$$\int_R f(\mathbf{r}) d^n r = \int_{R'} f(\mathbf{r}') \left|\det[\mathbf{J}]\right| d^n r'. \tag{C.231}$$

C.5.10 Summary of Transformations

1. The coordinate will change according to the relation

$$x \to x'(x). \tag{C.232}$$

2. The infinitesimal differentials dx^a transform as,

$$\begin{aligned} dx'^a &= \frac{\partial x'^a}{\partial x^b} dx^a \\ &= \Lambda^a{}_b \end{aligned} \tag{C.233}$$

3. Scalar fields transform as,

$$S(x) \to S(x'). \tag{C.234}$$

4. The volume element transforms as,

$$dx^1\, dx^2\, dx^3\, dx^4 = dx'^1\, dx'^2\, dx'^3\, dx'^4 |\det(\Lambda^a{}_b)| \tag{C.235}$$

where $|\det(\Lambda^a{}_b)|$ is the Jacobian factor calculated from this change of variables.

$$\sqrt{-g} = (-\det g_{a'b'})^{-1/2} \tag{C.236}$$

Then

$$\begin{aligned} \sqrt{-g} \to \sqrt{-g'} &= (-\det g_{a'b'})^{-1/2} \\ &= \sqrt{-\det(\Lambda^a{}_{a'}(\Lambda^b{}_{b'} g_{ab})} \\ &= (\det \Lambda^a{}_{a'} \det \Lambda^b{}_{b'})^{1/2} \sqrt{-\det g_{ab}} \\ &= |\det \Lambda^a{}_{a'}| \sqrt{-g} \end{aligned} \tag{C.237}$$

Tensor Densities

$$g' = \left| \frac{\partial x}{\partial x'} \right|^2 g \tag{C.238}$$

where $\left|\frac{\partial x}{\partial x'}\right|^2$ is the Jacobian associated with the coordinate transformation $x \to x'$.

$$\boxed{\sqrt{-g'} = \left| \frac{\partial x}{\partial x'} \right| \sqrt{-g}.} \tag{C.239}$$

Definition: A quantity which transforms as a scalar ... as a ***tensor density***.

The number of factors of the determinant $\left|\frac{\partial x}{\partial x'}\right|$, W, is called the ***weight*** of the density. The weight W takes the values of positive and negative integers.

Vector density

We take the covariant derivative of a vector density of weight W to be

$$\nabla_c L^a = \partial_c L^a + \Gamma^a_{bc} L^b - W\Gamma^b_{bc} \tag{C.240}$$

Suppose $W = +1$, and we took the trace of the covariant derivative, we get the result,

$$\nabla_a L^a = \partial_a L^a \tag{C.241}$$

so that the covariant divergence of a vector density of weight +1 is identical to the ordinary divergence. This result will be very important later on.

Generalized Kronecker delta

$$\delta^{ab}_{cd} = \begin{vmatrix} \delta^a_c & \delta^b_c \\ \delta^a_d & \delta^b_d \end{vmatrix} \tag{C.242}$$

and

$$\delta^{ab}_{cd} = \begin{vmatrix} \delta^a_d & \delta^b_d & \delta^c_d \\ \delta^a_e & 0\delta^b_e & \delta^c_e \\ \delta^a_f & \delta^b_f & \delta^c_f \end{vmatrix} \tag{C.243}$$

The product of alternating symbols give

$$\begin{aligned}
\epsilon^{abcd}\epsilon_{efgh} &= \delta^{abcd}_{efgh}, \\
\epsilon^{abcd}\epsilon_{efgd} &= \delta^{abc}_{efg} \\
\epsilon^{abcd}\epsilon_{efcd} &= 2\delta^{ab}_{ef} \\
\epsilon^{abcd}\epsilon_{ebcd} &= 3!\delta^{a}_{e} \\
\epsilon^{abcd}\epsilon_{abcd} &= 4!
\end{aligned} \tag{C.244}$$

Products of Delta functions and their symmetrization

$$D^{A'B'\ldots F'}_{A\ B\ \ldots F} = \delta^{A'}_{A}\delta^{B'}_{B}\ldots\delta^{F'}_{F} \tag{C.245}$$

$$\begin{aligned}
D^{A'B'\ldots F'}_{(A\ B\ \ldots F)} = \delta^{A'}_{(A}\delta^{B'}_{B}\ldots\delta^{F'}_{F)} = \frac{1}{n!}\left(\delta^{A'}_{A}\delta^{B'}_{B}\ldots\delta^{F'}_{F} + \text{all perms of A, B}, \ldots, \text{F}\right) \\
D^{(A'B'\ldots F')}_{A\ B\ \ldots F} = \delta^{(A'}_{A}\delta^{B'}_{B}\ldots\delta^{F')}_{F} = \frac{1}{n!}\left(\delta^{A'}_{A}\delta^{B'}_{B}\ldots\delta^{F'}_{F} + \text{all perms of A}', \text{B}', \ldots, \text{F}'\right)
\end{aligned} \tag{C.246}$$

For example

$$\begin{aligned}
D^{A'B'}_{(A\ B)} &:= \frac{1}{2!}\left(\delta^{A'}_{A}\delta^{B'}_{B} + \delta^{A'}_{B}\delta^{B'}_{A}\right) \\
D^{(A'B')}_{A\ B} &:= \frac{1}{2!}\left(\delta^{A'}_{A}\delta^{B'}_{B} + \delta^{B'}_{A}\delta^{A'}_{B}\right)
\end{aligned} \tag{C.247}$$

Obviously we have

$$D^{A'B'}_{(A\ B)} = D^{(A'B')}_{A\ B} = D^{(A'B')}_{(A\ B)} \tag{C.248}$$

This result is quite general. Consider

$$\begin{aligned}
\delta^{A'}_{(A}\ldots\delta^{C'}_{C}\ldots\delta^{E'}_{E}\ldots\delta^{G'}_{G)} &= \frac{1}{2}(\delta^{A'}_{(A}\ldots\delta^{\overset{\downarrow}{C'}}_{C}\ldots\delta^{\overset{\downarrow}{E'}}_{E}\ldots\delta^{G'}_{G)} + \delta^{A'}_{(A}\ldots\delta^{\overset{\downarrow}{E'}}_{C}\ldots\delta^{\overset{\downarrow}{C'}}_{E}\ldots\delta^{G'}_{G)}) \\
&= \delta^{A'}_{(A}\ldots\delta^{\overset{\downarrow}{E'}}_{C}\ldots\delta^{\overset{\downarrow}{C'}}_{E}\ldots\delta^{G'}_{G)}
\end{aligned} \tag{C.249}$$

that is, $D^{\ldots C'\ldots E'\ldots}_{(\ldots C\ldots E\ldots)} = D^{\ldots E'\ldots C'\ldots}_{(\ldots C\ldots E\ldots)}$ for arbitrary C' and E'. As $D^{A'B'\ldots F'}_{(A\ B\ \ldots F)}$ is symmetric under the interchange of any two of its upper indices we can symmetrize over some or all of them without changing the value

$$D^{A'B'\ldots F'}_{(A\ B\ \ldots F)} = D^{(A'B'\ldots F')}_{(A\ B\ \ldots F)} \tag{C.250}$$

Similar reasoning leads us to

$$D^{(A'B'\ldots F')}_{A\ B\ \ldots F} = D^{(A'B'\ldots F')}_{(A\ B\ \ldots F)} \tag{C.251}$$

So together we have

$$D^{A'B'\ldots F'}_{(A\ B\ \ldots F)} = D^{(A'B'\ldots F')}_{A\ B\ \ldots F} = D^{(A'B'\ldots F')}_{(A\ B\ \ldots F)}. \tag{C.252}$$

C.5.11 Variation of a Determinate

Consider finding the derivative of a determinate of a matrix whos elements are functions of coordinates. Take any square matrix $A = (a_{ij})$. Then its inverse , (b_{ij}) say, is given by (see C.184),

$$(b_{ij}) = \frac{1}{a}(A^T)^{ij} = \frac{1}{a}(A^{ji}), \tag{C.253}$$

where a is the determinate of a_{ij}, A^{ji} is the cofactor of a_{ij}, and the T denotes the transpose. Let us fix i, and expand the determinate a by the ith row. Then

$$a = \sum_{j=1}^{n} a_{ij} A^{ji} \tag{C.254}$$

where we have explicitly included the summation sign for clarity. If we differentiate both sides with respect to a_{ij}, then we get

$$\frac{\partial a}{\partial a_{ij}} = A^{ij}, \tag{C.255}$$

since a_{ij} does not occur in any of the cofactors A^{ij} (i fixed, j runs from 1 to n).

$$\frac{\partial a}{\partial a_{ij}} = a(\frac{1}{a}A^{ij}) = a(b^T)_{ij} \tag{C.256}$$

$$\frac{\partial a}{\partial x^k} = \frac{\partial a}{\partial a_{ij}}\frac{\partial a_{ij}}{\partial x^k} = a(b^T)_{ij}\frac{\partial a_{ij}}{\partial x^k} \tag{C.257}$$

$$\partial_c g = g g^{ab} \partial_c g_{ab}. \tag{C.258}$$

By definition of the inverse metric

$$g_{ac}(g^{-1})^{cb} = \delta_a^b \tag{C.259}$$

$$\delta g_{ac}(g^{-1})^{cb} + g_{ac}\delta(g^{-1})^{cb} = 0 \tag{C.260}$$

$$\frac{\partial g}{\partial g_{ab}} = g(\frac{1}{g}G^{ab}) = a g^{ab} \tag{C.261}$$

C.5.12 Tensor Densities

$$\delta\sqrt{g} = \frac{1}{2}\sqrt{g}g^{ab}\delta g_{ab} \tag{C.262}$$

Under a coordinate transformation a scalar density of weight n transforms as

$$S' = |J|^{n/2}S \tag{C.263}$$

We denote densities by placing n tilde's above the tensor if it is a density with a positive weight, or by placing n tilde's under the tensor. For example a tenor desity of weight 2 is denoted

$$\tilde{\tilde{S}} \tag{C.264}$$

$$\begin{aligned} \nabla_a \epsilon_{0123} &= \frac{\partial(-g)^{1/2}}{\partial x^a}[0123] - \Gamma^b_{0a}\epsilon_{b123} - \Gamma^b_{1a}\epsilon_{b123} - \Gamma^b_{2a}\epsilon_{b123} - \Gamma^b_{3a}\epsilon_{b123} \\ &= [\frac{\partial(-g)^{1/2}}{\partial x^a} - \Gamma^b_{ba}(-g)^{1/2}][0123]. \end{aligned} \tag{C.265}$$

We define the law of covariant derivitation of a tensor density as

$$\nabla_a(-g)^{1/2} = \partial_b(-g)^{1/2} - \Gamma^b_{ba}(-g)^{1/2}. \tag{C.266}$$

C.5.13 Divergence Theorem

$$\int (-g)^{1/2}\nabla_a \mathcal{L} = \int \nabla_a[(-g)^{1/2}\mathcal{L}] = \int \partial_a[(-g)^{1/2}\mathcal{L}] \tag{C.267}$$

$$\Gamma^a_{ab} = \partial_b \ln\sqrt{|g|} \tag{C.268}$$

$$\nabla_a A^a = \frac{1}{\sqrt{|g|}}\partial_a\left(A^a\sqrt{|g|}\right) \tag{C.269}$$

$$\int_V \sqrt{|g|}d^4x\nabla_a A^a = \int_{\partial V} d\Sigma_a A^a \tag{C.270}$$

$$d\Sigma_a = \frac{1}{3!}\sqrt{|g|}\epsilon_{abcd}dx^c \wedge dx^c \wedge dx^d \tag{C.271}$$

Details

(a)

Recall

$$\Gamma^c_{ab} = \frac{1}{2} g^{cd} \left\{ \frac{\partial g_{da}}{\partial x^b} + \frac{\partial g_{db}}{\partial x^a} - \frac{\partial g_{ab}}{\partial x^d} \right\} \tag{C.272}$$

Contracting over a and c,

$$\begin{aligned} \Gamma^a_{ab} &= \frac{1}{2} g^{ad} \left\{ \frac{\partial g_{da}}{\partial x^b} + \frac{\partial g_{db}}{\partial x^a} - \frac{\partial g_{ab}}{\partial x^d} \right\} \\ &= \frac{1}{2} g^{ad} \frac{\partial g_{da}}{\partial x^b} + \underbrace{\frac{1}{2} g^{ad} \left\{ \frac{\partial g_{db}}{\partial x^a} - \frac{\partial g_{ab}}{\partial x^d} \right\}}_{=0} \\ &= \frac{1}{2} g^{ad} \frac{\partial g_{da}}{\partial x^b} \end{aligned} \tag{C.273}$$

It is easily seen using (C.258) that

$$\Gamma^a_{ab} = \frac{1}{2} g^{ad} \frac{\partial g_{da}}{\partial x^b} = \frac{1}{\sqrt{|g|}} \frac{\partial \sqrt{g}}{\partial x^b} = \partial_b \ln \sqrt{|g|} \tag{C.274}$$

and we proven (C.268).

(b)

$$\nabla_a A^a = \partial_a A^a + \Gamma^a_{ab} A^b = \partial_a A^a + \frac{A^a}{\sqrt{|g|}} \frac{\partial \sqrt{g}}{\partial x^a} = \frac{1}{\sqrt{|g|}} \partial_a \left(A^a \sqrt{|g|} \right) \tag{C.275}$$

C.5.14 Summary of Tensor Calculus

- Linear spaces.
- Curvilinear coordinates
- Curved spaces.
- Vector and tensors: covariant and contravriant.
- Affine and metric connections.
- Differential manifolds.

C.5.15 Linear operators and Matrices

A linear operator T maps a vector space V onto itself which obeys linearity

$$T(a\hat{x} + b\hat{y}) = aT(\hat{x}) + bT(\hat{y}). \tag{C.276}$$

If we have a basis $\{\hat{e}_i, i = 1, \ldots, n\}$ for V, then

$$\hat{x} = \sum_{i=1}^{n} a_i \hat{e}_i \tag{C.277}$$

The action of a linear operator $\hat{T}$ on a basis vector $\hat{e}_i$ written $\hat{T}(\hat{e}_i)$ is a vector in the vector space, and so can be expressed with the basis:

$$T(\hat{e}_i) = \sum_{m=1}^{n} \hat{e}_m T_{mi} \tag{C.278}$$

Any vector $\vec{v}$ can be represented in the basis

$$\vec{v} = \sum_{i=1}^{n} a_i \hat{e}_i$$

The action of $\hat{T}$ on $\vec{v}$ is then

$$\begin{aligned} T(\hat{x}) &= T\left(\sum_{i=1}^{n} a_j \hat{e}_j\right) = \sum_{i=1}^{n} a_j T(\hat{e}_j) \\ &= \sum_{i=1}^{n} \hat{e}_i \sum_{j=1}^{n} T_{ij} a_j, \end{aligned} \tag{C.279}$$

where we have replaced each vector $T(\hat{e}_i)$ by its component form $\sum_{j=1}^{n} \hat{e}_i T_{ij}$.

Two successive linear transformations T and U acting on a space V produce the transformation UT:

$$
\begin{aligned}
UT(\hat{e}_i) &= U(T(\hat{e}_i)) = U\left(\sum_{j=1}^{n} \hat{e}_j T_{ji}\right) \\
&= \sum_{j=1}^{n} U(\hat{e}_j) T_{ji} \\
&= \sum_{j=1}^{n} \hat{e}_k U_{kj} T_{ji}
\end{aligned}
\tag{C.280}
$$

From which it follows that the components of UT are

$$
\sum_{j=1}^{n} U_{ij} T_{jk}. \tag{C.281}
$$

We see that this sum is just the matrix product of the respective matrices.

C.6 Group Theory

A group G is defined as a set of objects or operations (called elements) that may be combined or multiplied to form a well defined product and that satisfy the following four conditions. If we label the elements $a, b, c, \ldots$, then the conditions are:

(1) If a and b are any two elements, then the product ab is an element.

(2) The defined multiplication is associative, $(ab)c = a(bc)$

(3) There is a unit element I, with $Ia = aI = a$ for all elements a.

(4) Each element has an inverse $b = a^{-1}$, with $aa^{-1} = a^{-1}a = I$ for all elements a.

In physics, these abstract conditions will take on a physical meaning, for example in terms of transformations of vectors and tensors.

C.6.1 Examples of Groups

Translation in time: $t' = t + a^0$
Translation in space: $\vec{r}' = \vec{r} + \vec{a}$
Rotation in space : $\vec{r}' = R\vec{r}$

Lie groups

Roughly, a Lie group is a group with an infinite number of elements but which can be parametrized by one or several real numbers (the total of which is refered to as the *dimension* of the Lie group).

The simplest example of a Lie group is $SO(2)$, the group of rotations in the plane. Each element $R(\theta)$ is labelled by rotation angel θ, with multiplication acting as $R(\theta_1)R(\theta_2) = R(\theta_1 + \theta_2)$. Because the angle θ is defined only modulo 2π, the manifold of $SO(2)$ is a circle S^1.

interesting properties of Lie groups is that in neighbourhood of the identity element they can be expressed in terms of a set of generators T^a as

$$D(g) = \exp(-\alpha_a T^a) := \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \alpha_{a_1} \dots \alpha_{a_n} T^{a_1} \dots T^{a_n}, \tag{C.282}$$

where $\alpha_a \in \mathbb{C}$ are a set of coordinates of M in a neighbourhood of $\mathbb{1}$.

Internal symmetries

for e.g. $SU(2)$.

finite-dimensional Lie groups of imporance relativity translations, rotations, Lorentz, Poincare and special unitary group $SU(2)$. Apart from these finite-dimentional Lie groups, the infinie-dimensional ones play an important role in GR. These are the Lie group of *diffeomorphisms* of the spacetime manifold.

two groups are isomorphic if there is a one-to-one correspondence between their elements and if they have exactly the same structure.

The neighbourhood of a group element is chararterized by the neighbourhood of the corresponding parameter set.

Cosets

Let G be a group and H be a subgroup of G. The set of the form

$$gH = \{gh : h \in H\} \tag{C.283}$$

for $g \in G$ is called the (left) coset of H in G. There is one coset for each $g \in G$.

As we will see now, cosets define an equivalence on G, where $g \sim g'$ if the set gH is equal to the set $g'H$, in other words there is an $h \in H$ with $gh = g'$.

Cosets have the property that if cosets overlap they are the same coset, i.e., that for $g_1, g_2 \in G$, $g_1 \neq g_2$

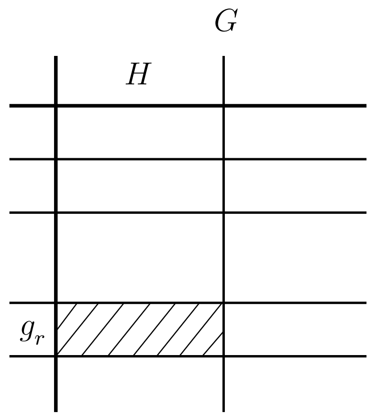


Figure C.7: coset. Suppose H is a subgroup of a finite group G, here the elements of H are listed first. The shaded box are the elements of the (left) coset of $g_r \in G$

$$g_1 H = g_2 H \tag{C.284}$$

or that different cosets do not overlap

$$g_1 H \cap g_2 H = \emptyset. \tag{C.285}$$

Proof: Suppose there is an overlap between the two cosets corresponding to the elements g_1 and g_2, that is, for some $h_1, h_2 \in H$ we have

$$g_1 h_1 = g_2 h_2.$$

Therefore

$$g_1 = g_2 h_2 h_1^{-1}.$$

If h is *any* element of H then

$$g_1 h = g_2 h_2 h_1^{-1} h = g_2 h'$$

where $h' = h_2 h_1^{-1} h$. Since H is a subgroup h' is also an element of H. Now since no element of G appears more than once in each row, if cosets do overlap at all they are in fact the same coset.

proving (C.284).

□

Right cosets are defined similarly, they are the sets

$$Hg = \{hg : h \in H\}. \tag{C.286}$$

These also define an equivalence realtion on the group G.

The equivalence class of a particular representative g is denoted

$$[g]. \tag{C.287}$$

Now the set of cosets gH, under the binary operator

$$[g] \cdot [g'] = [gg']. \tag{C.288}$$

forms a group if and only if H is what's known as a normal subgroup.

Definition Normal subgroups.

A normal N subgroup of a group G is

$$h \in H \text{ if } ghg^{-1} \in H \text{ for all } g \in G. \tag{C.289}$$

We write $N \triangleleft G$ to indicate that N is a normal subgroup of G.

We need to establish that induced multiplication between equivalence classes is independent of the representative chosen, i.e., that $[g_1'][g_2'] = [g_1][g_2]$, where g_1 is any element of $[g_1']$ and similarly for g_2.

$$[g_1'][g_2'] = [g_1h_1][g_2h_2] = [g_1h_1g_2h_2] = [g_1g_2hh_2] = [g_1g_2h] = [g_1g_2] = [g_1][g_2]. \tag{C.290}$$

$$[hg] = [h][g] = [e][g] = [eg] = [g] = [ge] = [g][e] \tag{C.291}$$

$[e] = [h]$

$$[e] = [gg^{-1}] = [g][g^{-1}] = [g][g]^{-1} \tag{C.292}$$

$[g]^{-1} = [g^{-1}]$

This makes G/N into a group, called the quotient group.

Definition The **Quotient group** is the subgroup is the collection of cosets, each being considered an element. The quotient group is denoted G/H.

We will be interested in other mathematical things such as algebras, rings, categories amongt others. There are analogous notions of the normal subgroups for each of these goining by the names of ideals, , quotient categories.

Simple and Semi-simple Lie Groups

A Lie group is called simple if it does not possess an invariant subgroup.

A Lie group is called semi-simple if it does not possess an abelian invariant subgroup. Note that a semi-simple subgroup can possess a non-abelian invariant sugroup.

An invariant subset is an ideal if

$$[A_i, G_j] = \sum a_{ijk} A_k \tag{C.293}$$

Cartan killing form k_{ij} built from the structure constants

$$k_{ij} = C_{iab} C_{jba} \tag{C.294}$$

Cartan killing form

The killing form k_{ab} is nondegenerate if the Lie algebra is semisimple

$$\det(k_{ij}) \neq 0. \tag{C.295}$$

Proof:

we have to demonstrate that $\det(k_{ij}) = 0$ if there is an abelian ideal. suppose that the Lie algebra has an abelian ideal. generators belonging to the ideal will be labelled marked indices. Then for the b' column of the killing form,

$$\det(k_{ij'}) = C_{iab} C_{j'ba} = C_{ia'b} C_{j'ba'} \tag{C.296}$$

as by (C.293) $C_{j'ba} = 0$ for those values of a which do not belong to the ideal. it follows from (C.296) that

$$\det(k_{ij'}) = -C_{a'ib} C_{j'ba'} = -C_{a'ib'} C_{j'b'a'} \tag{C.297}$$

since $C_{a'ib} = 0$ for all $b \neq b'$ the same reason. $[A_i, A_j] = 0$ for an abelian ideal so

$$C_{j'b'a'} = 0$$

holds. As the j' column of the killing form vanishes it follows that

$$\det(k_{ij}) = 0. \tag{C.298}$$

Cartan's condition can be stated that if a Lie algebra is semisimple its killing form is invertible, that is, there is an inverse $k^{ij} := (k_{ij})^{-1}$ for which

$$k^{ik} k_{kj} = \delta^i_j$$

holds.

C.6.2 Unitary Representations of Groups

R	Group elements.
$[G]$	The collection of group elements
R_a	An indexed group element.
$\alpha, \beta, \ldots$	Labels irreducible representations.
D	Matrix representation.
$D^{(\alpha)}$	Indexed irriduceable matrix representation.
$i, j, \ldots$	Labels components of matrix representations
h	Dimension of group.
χ	Character

Proof Let $\{A_a\}$ be a representation of the group G. Construct the Hermitian matrix $H = \sum_{a=1}^{h} A_a A_a^\dagger$

$$H^\dagger = [\sum_a^h (A_a A_a^\dagger)]^\dagger = \sum_a^h (A_a A_a^\dagger)^\dagger = \sum_a^h A_a A_a^\dagger = H \tag{C.299}$$

From the earlier discussion of matrix algebra, any Hermitian matrix can be diagonalized by a similarity transformation. Let

$$d = \sum_a^h U^\dagger A_a A_a^\dagger U = \sum_\alpha (U^\dagger A_a U)(U^\dagger A_a^\dagger U) = \sum_\alpha (U^\dagger A_a U)(U^\dagger A_a U)^\dagger \tag{C.300}$$

Hence, $d = \sum_a \tilde{A}_a \tilde{A}_a^\dagger$ where $\tilde{A}_a = U^\dagger A_a U = U^{-1} A_a U$. The elements of the diagonal d matrix are real and positive.

$$d_{jj} = \sum_{a}^{h} \sum_{k} (\tilde{A}_a)_{jk} (\tilde{A}_a^\dagger)_{kj} = \sum_{a}^{h} \sum_{k} (\tilde{A}_a)_{jk} (\tilde{A}_a)^*_{jk} = \sum_{a}^{h} \sum_{k} |(\tilde{A})_{jk}|^2 \tag{C.301}$$

for all $j = 1, \ldots, h$. Since d is diagonal, we can define its square-root $d^{1/2}$

$$d^{1/2} = \begin{pmatrix} d_{11}^{1/2} & 0 & \cdots & 0 \\ 0 & d_{11}^{1/2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{hh}^{1/2} \end{pmatrix} \tag{C.302}$$

and $d^{-1/2}$, which is given by an analogous expression. Evidently, $(d^{1/2})^2$ and $d^{1/2}d^{-1/2} = 1$, where I is the identity matrix. Diagonal matrices commute with each other, and we can write for the identity I

$$I = d^{-1/2} d d^{-1/2} = d^{-1/2} \sum_{a}^{h} \tilde{A}_a \tilde{A}_a^\dagger \, d^{-1/2}. \tag{C.303}$$

By the rearangement theorem, $\{\tilde{A}_b\tilde{A}_a\}$ all a, and any one b is equal to $\{\tilde{A}\}$ for all a. Hence,

$$\begin{aligned} I &= d^{-1/2} \sum_a (\tilde{A}_b \tilde{A}_a) \, (\tilde{A}_b \tilde{A}_a)^\dagger d^{-1/2} \\ &= d^{-1/2} \sum_a (\tilde{A}_b \tilde{A}_a \; \tilde{A}_a^\dagger \tilde{A}_b^\dagger) d^{-1/2} \\ &= d^{-1/2} \tilde{A}_b d^{1/2} \left(d^{-1/2} \sum_a \tilde{A}_a \tilde{A}_a^\dagger d^{-1/2} \right) d^{1/2} \tilde{A}_b^\dagger d^{-1/2} \end{aligned} \tag{C.304}$$

using (C.303) in this gives

$$\begin{aligned} I &= (d^{-1/2} \tilde{A}_b d^{1/2})(d^{-1/2} \tilde{A}_b d^{1/2})^\dagger \\ &= (d^{-1/2} U^{-1} A_b U d^{1/2})(d^{-1/2} U^{-1} A_b U d^{1/2})^\dagger. \end{aligned} \tag{C.305}$$

Let us define the matrix $B_a := d^{-1/2} U^{-1} A_a U d^{1/2}$. It is easy to see that B_a has the same multiplication table as A_α. If we have $A_a A_b = A_c$, then

$$\begin{aligned} B_a B_b &= (d^{-1/2} U^{-1} A_\alpha U d^{1/2})(d^{-1/2} U^{-1} A_b U d^{1/2}) \\ &= (d^{-1/2} U^{-1} A_a) \, (U d^{1/2} d^{-1/2} U^{-1}) \, (A_b d^{1/2} U) \\ &= (d^{-1/2} U^{-1} \; A_a A_b \; U d^{1/2}) \\ &= (d^{-1/2} U^{-1} A_c U d^{1/2}) = \; B_c \end{aligned} \tag{C.306}$$

Therefore $\{B_a\}$ is a unitary reprsentation of the group G.

Representation of finite groups can always be taken to be unitary. It is essential that the sum over $g \in G$ converge. This is guaranteed for a finite group, but may not work for infinite groups. In particular, non-compact Lie groups, such as the Lorentz group, have no finite dimensional unitary representations.

C.6.3 Schur's First Lemma

Theorem C.6.1 *A non-zero matrix which commutes with all of the matrices of an irreducible representation is a constant multiple of the unit matrix.*

Proof:

We can take these matrices to be unitary without loss of generality. Suppose there is a matrix M that commutes with all of the A_a but which is not a constant multiple of the unit matrix:

$$MA_a = A_a M \tag{C.307}$$

for all $a = 1, 2, \ldots, |G|$. By taking the edjoint of each of these equations, we obtain

$$A_a^\dagger M^\dagger = M^\dagger A_a^\dagger. \tag{C.308}$$

Since A_a are unitary

$$A_a^{-1} M^\dagger = M^\dagger A_a^{-1}. \tag{C.309}$$

Multiplying both this from left and right we obtain

$$M^\dagger A_a = A_a M^\dagger \tag{C.310}$$

So that, if M commutes with every matrix of the representation, then so does $M^\dagger$. As such the Hermitian matries

$$H_1 = M + M^\dagger, \quad H_2 = i(M - M^\dagger) \tag{C.311}$$

also commute with every matrix of the representation. We prove the statement of the theorem for Hermitician matrices. We start with

$$HA_a = A_a H. \tag{C.312}$$

Since a hermitian matrix H can be diagonalized by some U, $D = U^{-1}HU$.

$$U^{-1}HUU^{-1}A_aU = U^{-1}A_aUU^{-1}HU, \tag{C.313}$$

Let $\tilde{A}_a = U^{-1}A_aU$

$$D\tilde{A}_a = \tilde{A}_aD. \tag{C.314}$$

A non-zero matrix which commutes with all of the matrices of an irreducible representation is a constant multiple of the unit matrix.

$$(\tilde{A}_a)_{ij}(D_{ii} - D_{jj}) = 0. \tag{C.315}$$

(i) Suppose all the diagonal elements of D are distinct: $D_{ii} \neq D_{jj}$ if $i \neq j$. Then (C.315) implies that

$$(\tilde{A}_a)_{ij} = 0, \qquad i \neq j,$$

i.e., the off-diagonal elements of $\tilde{A}_\alpha$ must vanish. They form a reducible representatoin composed of d one-dimentional representations. Since the $\tilde{A}_a$ are obtained from the A_a by a similarity transformation, the A_a are a reducible representaion.

(ii) Suppose p of the diagonal elements of D are equal, but the remaining entries are distinct from these and from each other. Using similarity transfromations, (for example (C.316)), we can arrange the diagonal elements so that the first p are equal.

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} b & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix} \tag{C.316}$$

Hence, we can assume this is the case without loss of generality, i.e. $D_{11} = D_{22} = \cdots = D_{pp}$, $D_{mm} \neq D_{nn}$, otherwise. The $(\tilde{A}_i)_{mn}$ must vanish for any pair of unequal diagonal entries.

$$\tilde{A}_i = \begin{pmatrix} B_1 & 0 \\ 0 & B_2, \end{pmatrix} \tag{C.317}$$

where B_1 is a $p \times p$ matrix and B_2 is a $(p-d) \times (p-d)$ diagonal matrix.

We have shown that if a Hermtian matrix is not a multiple of the unit matrix and commutes with all the matrices of a representation, then that representation is necessarily reducible. Thus, if a non-zero Hermtian matrix which commutes with all the matrices of an irreducible representation that matrix must be a multiple of the unit matrix.

Given that $M + M^\dagger$ and $i(M - M^\dagger)$ are both Hermitician matrices that commute with all of the matrices of an irreducible representation they are a constant multiple of the unit matrix. we have $H_1 = M + M^\dagger = c_1 I$ and $H_2 = i(M - M^\dagger) = c_2 I$. We have that

$$M = H_1 - iH_2 = (c_1 - ic_2)I$$

and hence must be proportional to the unit marix.

□

C.6.4 Schur's Second Lemma

Theorem C.6.2 *Let $\{A_1, A_2, \ldots, A_{|G|}\}$ and $\{B_1, B_2, \ldots, B_{|G|}\}$ be two irreducible representations of a group G of dimensionalities d an d' respectively. If there is a matrix M such that*

$$MA_a = B_a M$$

for $a = 1, 2, \ldots, |G|$, then if $d = d'$, either $M = 0$ or the two representations are related by a similarity transformation. If $d \neq d'$, then $M = 0$.

As

$$MA_a = B_a M$$

then

$$A_a^\dagger M^\dagger = M^\dagger B_a^\dagger$$

using unitary

$$A_a^{-1} M^\dagger = M^\dagger B_a^{-1}. \tag{C.318}$$

Multiplying on the right by M, we get

$$A_a^{-1} M^\dagger M = M^\dagger B_a^{-1} M. \tag{C.319}$$

By the group properties B_a^{-1} is some B_b and A_a^{-1} is the corresponding A_b, and so by the posulate of the theorem, $MA_b = B_b M$, or

$$MA_b^{-1} = B_b^{-1}M$$

substituting this into (C.319), gives

$$A_a^{-1}M^\dagger M = M^\dagger M A_a^{-1}. \tag{C.320}$$

Thus, the $d' \times d'$ matrix $MM^\dagger$ commutes with all the matrices of an irreducible representation. By Schur's first lemma, $MM^\dagger$ must therefore be a constant multiple of the unit matrix,

$$MM^\dagger = cI. \tag{C.321}$$

(i) Say $d = d'$. If $c \neq 0$, equation (C.321) implies that

$$M^{-1} = \frac{1}{c}M^\dagger.$$

plus (C.318) can be rearranged

$$A_a = M^{-1}B_aM,$$

so the two representations are related by a similarity transfromation and are, therefore, equivalent and since $M^\dagger M = c^d$,it follows that $c^{-d'/2}M$ is a unitary matrix. If $c = 0$ then

$$(MM^\dagger)_{ij} = \sum_k M_{ik}(M^\dagger)_{kj} = \sum_k M_{ik}M_{jk}^* = 0. \tag{C.322}$$

By setting $i = j$, we obtain

$$\sum_k M_{ik}M^*{}_{ik} = \sum_k |M_{ik}|^2 = 0 \tag{C.323}$$

which implies that $M_{ik} = 0$ for all i and k, i.e., that M is the zero matrix.

(ii) Say $d \neq d'$. We take $d < d'$. Them M is a rectangular matrix with d columns and d' rows,

$$\begin{pmatrix} M_{11} & \cdots & M_{1d} \\ M_{21} & \cdots & M_{2d} \\ \vdots & \cdots & \vdots \\ \vdots & \cdots & \vdots \\ M_{d'1} & \cdots & M_{d'd}. \end{pmatrix} \tag{C.324}$$

We add $d' - d$ columns of zeros

$$N = \begin{pmatrix} M_{11} & \cdots & M_{1d} & 0 & \cdots & 0 \\ M_{21} & \cdots & M_{2d} & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & & \vdots \\ \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \cdots & & \vdots & & \vdots \\ M_{d'1} & \cdots & M_{d'd} & 0 & \cdots & 0 \end{pmatrix} \tag{C.325}$$

Taking the adjoint of this matrix yields

$$N^\dagger = \begin{pmatrix} M_{11}^* & M_{21}^* & \cdots & M_{d'1} \\ M_{12}^* & M_{22}^* & \cdots & M_{d'2} \\ \vdots & & \cdots & \vdots \\ M_{d1}^* & M_{2d}^* & \cdots & M_{d'd}^* \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \tag{C.326}$$

We have

$$NN^\dagger = MM^\dagger = cI.$$

Taking the determinant

$$\det(NN^\dagger) = \det(N)\det(N^\dagger) = c^{d'} = 0.$$

so $c = 0$, so that, as before we arrive at

$$\sum_k |N_{ik}|^2 = 0. \tag{C.327}$$

which implies that $N_{ik} = 0$ for all i and k, i.e. N is the zero matrix. Hence, $M = 0$.

□

C.6.5 Orthogonality relations

Let $D^{(\alpha)}(R)_{ij}$ be the (i,j) matrix elements of the α^{th} irreducible unitary representation of the element R of G. The range of R will be the h elements of G. The range of α will be the number of inequivalent irreducible representations. The range of i and j will be the dimension d_α of the α^{th} representation.

Theorem:

$$\sum_R D^{(\alpha)}(R)^*_{ij} D^{(\beta)}(R)_{i'j'} = \frac{h}{d_\alpha} \delta_{\alpha\beta} \delta_{ii'} \delta_{jj'} \tag{C.328}$$

Proof:

Let $D^{(1)}(R)$ not be equivalent to $D^{(2)}(R)$. We define the matrix M,

$$M = \sum_R D^{(2)}(R) X D^{(1)}(R^{-1}). \tag{C.329}$$

Then

$$\begin{aligned} D^{(2)}(S)M &= \sum_R D^{(2)}(S) \left[D^{(2)}(R) X D^{(1)}(R^{-1}) \right] D^{(1)}(S^{-1}) D^{(1)}(S) \\ &= \sum_R \left[D^{(2)}(S) D^{(2)}(R) \right] X \left[D^{(1)}(R^{-1}) D^{(1)}(S^{-1}) \right] D^{(1)}(S) \end{aligned} \tag{C.330}$$

If $SR = T$, then

$$D^{(\alpha)}(S) D^{(\alpha)}(R) = D^{(\alpha)}(SR) = D^{(\alpha)}(T)$$

and

$$D^{(\alpha)}(R^{-1}) D^{(\alpha)}(S^{-1}) = D^{(\alpha)}(R^{-1}S^{-1}) = D^{(\alpha)}(T^{-1})$$

So by the rearrangment theorem, (C.330) becomes

$$\begin{aligned} D^{(2)}(S)M &= \sum_T D^{(2)}(T) X D^{(1)}(T^{-1}) D^{(1)}(S) \\ &= M D^{(1)}(S) \end{aligned} \tag{C.331}$$

for all S.

Hence by Schur's second lemma, $M = 0$. Now choose X to be a matrix every element of which is zero except one, which we shall call $X_{j'i}$, and $X_{j'i} = 1$.

$$M = 0 = \sum_R D^{(2)}(R) X D^{(1)}(R^{-1}) \tag{C.332}$$

or

$$\begin{aligned} 0 &= \sum_{kl} \sum_R D^{(2)}(R)_{i'l} X_{kl} D^{(1)}(R^{-1})_{lj} \\ &= \sum_R D^{(2)}(R)_{i'j'} D^{(1)}(R^{-1})_{ij} \\ &= \sum_R D^{(1)}(R)^*_{ji} D^{(2)}(R^{-1})_{i'j'} \end{aligned} \tag{C.333}$$

The last step follows from the unitarity of the representation.

$$D^{(\alpha)}(R^{-1})_{ij} = [D^{(\alpha)}(R^{-1})]^{-1}_{ij} = D^{(\alpha)}(R)^*_{ji} \tag{C.334}$$

This proves the theorem for $\alpha \neq \beta$

$$\sum_i D^{(\alpha)}_{ii}(R) \sum_j D^{(\beta)}_{jj}(R)^* = g\delta_{\alpha\beta} \tag{C.335}$$

case(b): $\alpha = \beta = 1$

We define $M = \sum_R D^{(1)}(R) D^{(1)}(R^{-1})$. Then just as in case (a),

$$M D^{(1)}(S) = D^{(1)}(S) M \tag{C.336}$$

for all S. Hence $M = cI$. Let $X_{jj'} = 1$ for a particular j and j' and choose all other elements of X equal to zero.

$$M = cI \sum_R D^{(1)}(R) X D^{(1)}(R^{-1})$$

or in component form

$$c\delta_{ii'} = \sum_R D^{(1)}(R)_{ij} D^{(1)}(R^{-1})_{j'i'}. \tag{C.337}$$

Let $j = j'$ and sum over j:

$$c\sum_{i=1}^{d_1} 1 = \sum_R \sum_i D^{(1)}(R^{-1})_{j'i} D^{(1)}(R)_{ij}. \tag{C.338}$$

$$\begin{aligned}
cd_1 &= \sum_R \sum_i D^{(1)}(R^{-1}R)_{jj'} \\
&= \sum_R \sum_i D^{(1)}(I)_{jj'} \\
&= \sum_R \delta_{j'j} = h\delta_{j'j}
\end{aligned} \tag{C.339}$$

Therefore

$$c = \frac{h}{d_1}\delta_{jj'}. \tag{C.340}$$

We replace c in (C.337) by this value and rewrite $D^{(1)}(R^{-1})_{j'i'}$ as $D^{(1)}(R)^*_{i'j'}$ to get

$$\sum_R D^{(1)}(R)^*_{i'j'} D^{(1)}(R)_{ij} = \frac{h}{d_1}\delta_{ii'}\delta_{jj'}. \tag{C.341}$$

This establishes the theorem for general $\alpha = \beta$.

□

C.6.6 The Characters of a Representation

The characters are invariant under similarity transformations, all elements of one class have the same character.

First Orthogonal Relation

If α and β are irreducible and N_k is the number of elements in c_k, then

$$\sum_k \chi^{(\alpha)}(c_k)^* \chi^{(\beta)}(c_k) N_k = h\delta_{\alpha\beta} \tag{C.342}$$

Proof:

Starting with the orthogonality relatioins (C.328), theorem and setting $i = j$ and $i' = j'$ summing over i and i', we get

$$\sum_R \chi^{(\alpha)*}(R)\chi^{(\beta)}(R) = \frac{h}{d_\alpha}\delta_{\alpha\beta}\underbrace{\sum_{i,i'}\delta_{ii'}\delta_{ii'}}_{=d_\alpha}$$

or

$$\sum_R \chi^{(\alpha)}(c_k)^*\chi^{(\beta)}(c_k)N_k = h\delta_{\alpha\beta}. \tag{C.343}$$

□

The Second Orthogonality Relation If $D^{(\alpha)}$ is irreducible, then

$$\sum_\alpha \chi^{(\alpha)}(c_k)^*\chi^{(\alpha)}(c_l) = \frac{h}{N_l}\delta_{kl} \tag{C.344}$$

Proof:

$$\chi = \sum_\alpha a_\alpha \chi^{(\alpha)} \tag{C.345}$$

Defining the weighted inner product

$$\begin{aligned} \chi \cdot \lambda &= \sum_\alpha a_\alpha \chi(c_k)\cdot\chi(c_\beta) \\ &= \sum_\alpha a_\alpha h\delta_{ij} = ha_j \end{aligned} \tag{C.346}$$

Hence,

$$\begin{aligned} \chi &= \sum_\alpha \frac{\chi\cdot\chi^{(\alpha)}}{h}\chi^{(\alpha)} \\ &= \chi\cdot\sum_\alpha \frac{\chi^{(\alpha)}\chi^{(\alpha)}}{h} \end{aligned} \tag{C.347}$$

$$\chi(c_l) = \sum_{k,\alpha} N_k\chi(c_k)\chi^{(\alpha)}(c_k)\chi^{(\alpha)}(c_l)\frac{1}{h} \tag{C.348}$$

$$0 = \sum_k \chi(c_k)\left[\frac{N_k}{h}\sum_\alpha \chi^{(\alpha)}(c_k)\chi^{(\alpha)}(c_l) - \delta_{kl}\right] \tag{C.349}$$

□

C.6.7 Direct Products

Details Direct Product of matrices and vectors.

By definition the vector product of two 2-by-2 matrices is

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} \tag{C.350}$$

A more explicit form

$$A \otimes B = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{12}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix} \tag{C.351}$$

$$C = \begin{pmatrix} c_{11;11} & c_{11;12} & c_{12;11} & c_{12;12} \\ c_{11;21} & c_{11;22} & c_{12;21} & c_{12;22} \\ c_{21;11} & c_{21;12} & c_{12;11} & c_{22;12} \\ c_{21;21} & c_{21;22} & c_{22;21} & c_{22;22} \end{pmatrix} \tag{C.352}$$

In general the direct product C of two matrices A and B is defined in terms of matrix elements by

$$a_{ij}b_{kl} = c_{ik;jl}. \tag{C.353}$$

The row and column labels of C are ***composite*** labels:

$$M_{i1}, M_{i,2}, \ldots, M_{1,n} \tag{C.354}$$

whereas the as one goes along a row the read

$$c_{ik;1,1},\ c_{ik;1,2},\ \ldots,\ c_{ik;1,n},\ c_{ik;2,1},\ c_{ik;2,2},\ \ldots,\ c_{ik;2,n,},\ \ldots,\ c_{ik;n,1},\ c_{ik;n,2},\ \ldots,\ c_{ik;n,n,}. \tag{C.355}$$

the row label, is obtained from the

We prove these direct products have the same operations as matries.

Products

$$(A \otimes B)(C \otimes D) = AC \otimes BD \tag{C.356}$$

$$
\begin{aligned}
(A \otimes B)(C \otimes D) &= \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} \begin{pmatrix} c_{11}D & c_{12}D \\ c_{21}D & c_{22}D \end{pmatrix} \\
&= \begin{pmatrix} (a_{11}c_{11} + a_{12}c_{21})BD & (a_{11}c_{12} + a_{12}c_{22})BD \\ (a_{21}c_{11} + a_{22}c_{21})BD & (a_{21}c_{12} + a_{22}c_{22})BD \end{pmatrix} \\
&= AC \otimes BD \qquad \text{(C.357)}
\end{aligned}
$$

$$D \times D \qquad \text{(C.358)}$$

$$\chi^{(\alpha\times\beta)} = \sum_{ij} D^{(\alpha\times\beta)}_{(ii)(jj)}(G_a) = \sum_i D^{(\alpha)}_{ii}(G_a) \sum_j D^{(\beta)}_{jj}(G_a) \qquad \text{(C.359)}$$

The character of the representation of the direct product is equal to the product of the characters of the original representations α and β, which implies that

$$\chi^{(\alpha\times\beta)} = \chi^{(\alpha)} \cdot \chi^{(\beta)}. \qquad \text{(C.360)}$$

Thr representations resulting from (C.358) is in general reducible, that is

$$D^{(\alpha\times\beta)}(G_a) = \oplus_\gamma m_\gamma D^\gamma(G_a) \qquad \text{(C.361)}$$

$$m_\gamma = \frac{1}{g} \sum_p c_p \chi^{(\gamma)*} \chi^{(\alpha)}_p \chi^{(\beta)}_p. \qquad \text{(C.362)}$$

C.7 Continuous Groups, Lie Groups and Lie algebras

We are only interested in symmetery transformations are all based on *contiuous* quantities determing how summations over group elements are carried out.

The neighbourhood of a group element is characterized by the neighbourhood of the corresponding parameter set.

symetries e.g. rotational each point in three spacial manifold).

Instead of having to consider the group as a whole, for many purposes it is sufficient to consider the an infinitesimal transformation around the identity. Any finite transformation can then be constructed by the reapeted application of this infinitesimal transformation.

We will the properties of Lie groups and algebras in terms of specific examples, especially the two and three dimensional rotation groups and the Lorentz group.

C.7.1 Infinitesmal Generating Technique

A function

$$f(x+\delta x) = f(x) + \delta x \frac{df(x)}{dx} \tag{C.363}$$

$$f(x+2\delta x) = f(x+\delta x) + \delta x \frac{df(x+\delta x)}{dx} = f(x) + 2\delta x \frac{df(x)}{dx} + \delta x \frac{d^2 f(x)}{dx^2} \tag{C.364}$$

Binomial type expansion

$$f(x+N\delta x) = \sum_{r=0}^{N} \frac{N!}{r!(N-r)!} \delta x^r \frac{d^r f(x)}{dx^r} \tag{C.365}$$

In the limit $N \to \infty$ the factor $N!/(N-r)$ can be replaced by N^r. We put $N\delta x = a$ where, in the limit $N \to \infty$ a is a finite number.

$$f(x+a) = \sum_{r=0}^{\infty} \frac{a^r}{r!} \frac{d^r f(x)}{dx^r} = f(x) + a\frac{df(x)}{dx} + \frac{a^2}{2!}\frac{d^2 f(x)}{dx^2} + \dots \tag{C.366}$$

which we can formally write

$$f(x+a) = \exp\left(a\frac{d}{dx}\right) f(x) \tag{C.367}$$

Rotation operators

$$R(\delta\theta) = \begin{pmatrix} 1 & d\theta \\ -d\theta & 1 \end{pmatrix} = 1 + id\theta J \quad \text{where} \;\; J = i\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{C.368}$$

$$J^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{C.369}$$

$$\begin{aligned} R(\theta) &= \left(1 + \frac{i\theta}{N} J\right)^N = \exp(i\theta J) = I + iJ + \frac{(i\theta)^2}{2!}J^2 + \frac{(i\theta)^3}{3!}J^3 + \\ &= I\left[1 - \frac{\theta^3}{3!} + \frac{\theta^5}{5!}\dots\right] + iJ\left[1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!}\dots\right] = I\cos\theta + iJ\sin\theta \\ &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \end{aligned} \tag{C.370}$$

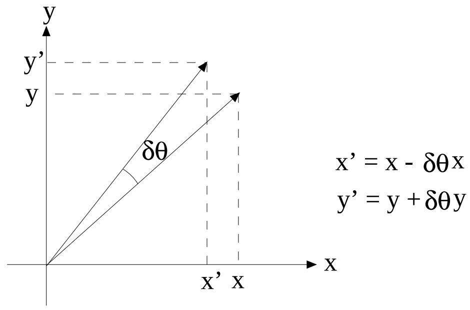


Figure C.8: infintesmal rotation.

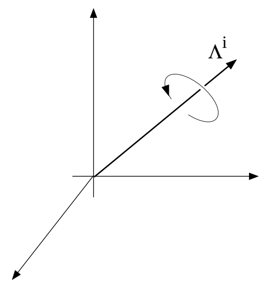


Figure C.9: infintesmal rotation.

$$\mathbf{r}' = \mathbf{r} + \delta\phi \times \mathbf{r} \tag{C.371}$$

expressed in component form

$$x_I' = x_i + \epsilon_{ijk}x_k = (\delta_{ik} + \epsilon_{ijk})x_k \tag{C.372}$$

$$\hat{R} = \begin{pmatrix} 1 & \delta\phi_z & \delta\phi_y \\ -\delta\phi_z & 1 & \delta\phi_x \\ \delta\phi_y & -\delta\phi_x & 1 \end{pmatrix} \tag{C.373}$$

$$\left(\mathbf{I} - \frac{i}{\hbar}\delta\phi \cdot \hat{\mathbf{L}}\right) \tag{C.374}$$

where

$$L_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad L_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad L_z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{C.375}$$

Thus, given two vector angles α and β, there must exist a third one, γ, such that

$$\exp(i\alpha\cdot\mathbf{J})\exp(i\beta\cdot\mathbf{J})=\exp(i\gamma\cdot\mathbf{J}). \tag{C.376}$$

For this to be so the matrices must have to satisfy some condition, which can be found by considering the angles α^i and β^i as small.

$$U(\delta\alpha_\mu)=I+i\alpha^iJ_i-\frac{1}{2}\sum_{i,j}\alpha^i\alpha^jJ_iJ_j+\dots\quad, \tag{C.377}$$

as J_i are *unitary*, i.e. $U^\dagger(\alpha^i)=U^{-1}(\alpha^i)$, then from() we obtain

$$U^\dagger(\alpha^i)I-i\alpha^iJ_i^\dagger=U^{-1}(\alpha^i)=I-i\alpha^iJ_i. \tag{C.378}$$

from this follows the *Hermiticity of the generators*

$$U_i^\dagger=U_i \tag{C.379}$$

Next we calculate the inverse operator to second order in α^i

$$U^{-1}(\alpha^i)=I-i\alpha^iJ_i-\frac{1}{2} \tag{C.380}$$

$$I+i(\alpha^iJ_i+\beta^iJ_i)-\frac{1}{2}(\alpha^iJ_i+\beta^iJ_i)^2=I+i(\alpha^iJ_i+\beta^iJ_i)-\frac{1}{2}(\alpha^i\alpha^j+\alpha^i\beta^j+\alpha^j\beta^i+\beta^i\beta^j)J_iJ_j\alpha^i\beta^i[J_i,J_j]+\dots \tag{C.381}$$

$$+\dots=i(\alpha^i+\beta^i)J_i-\frac{1}{2} \tag{C.382}$$

the commutator must be a linear combination of the J_i, that is

$$[J_i,J_j]=C_{ij}^kJ_k. \tag{C.383}$$

The generators of *any* Lie group must have such commutation relations. The coefficients C_{ij}^k are called the *structure constants*

The product of two rotations $\exp(-iT_y)\exp(-iT_z)$ can always be written as a single exponential, say $\exp(-i\alpha\cdot T)$ where $\alpha\cdot T:=\alpha_xT_x+\alpha_yT_y+\alpha_zT_z$. Suppose we set $\exp(-i\alpha\cdot T)\exp(-i\beta\cdot T)=\exp(-i\gamma\cdot T)$ and try to calculate γ in terms of α and β. If we expand the exponentials we find

$$
\begin{aligned}
[1 - i\alpha\cdot t - \frac{1}{2}(\alpha\cdot t)^2 &+ \dots][1 - i\beta\cdot t - \frac{1}{2}(\beta\cdot t)^2 + \dots] \\
&= \exp(-i(\alpha+\beta)\cdot t - \frac{1}{2}[\alpha\cdot t, \beta\cdot t] + \dots)
\end{aligned} \tag{C.384}
$$

(and this is known as the Campbell-Baker-Hausdorff theorem - see Appendix S). It is for this reason that we can learn all we need to know about Lie groups by studying the commutation algebras of the generators

C.7.2 General Structure of Lie Groups

An infinite group is a group that contains an infinite nuber of elements. The rotation group is an example of such a group.

$$
SO(3) = \exp(-i\hat{\phi}\cdot J). \tag{C.385}
$$

The fact that these matrices are functions of only three fundamental matrices $\{\hat{J}_k\} = \{\hat{J}_1\hat{J}_2\hat{J}_3\}$ allows us to represent them in a simple way.

The set of group elements are characterized by a set of continuous real parameters.

A continuous group G is said to be *compact* if the parameter space is finite and non-compact if it is infinite. The rotation group $SO(3)$ is an example of a compact group with the Lorentz group $SO(3,1)$ as an example of a non-compact group.

$$
\begin{aligned}
[A,[B,C]] &= A[B,C] - [B,C]A \\
&= ABC - ACB - BCA + CBA
\end{aligned} \tag{C.386}
$$

We now add and subtract the quantities BAC and CAB on the right-hand side of this equation and rearange the resulting expression

$$
\begin{aligned}
[A,[B,C]] &= ABC - ACB - BCA + CBA \\
&\quad -BAC + BAC + CAB - CAB \\
&= -C(AB - BA) + (AB - BA)C \\
&\quad +B(AC - CA) - (AC - CA)B \\
&= -[C,[A,B]] - [B,[C,A]]
\end{aligned} \tag{C.387}
$$

A simple rearrangement yields the *Jacobi identity*:

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \tag{C.388}$$

The important properties of the structure constants are the following

(i) They are antisymmetric in their lower indices

$$C_{ij}^{k} = -C_{ji}^{k} \tag{C.389}$$

(ii) The Jacobi identity defined by the infinitesimal generators (e.g. for the rotation group: $A = J_k,\ B = J_l,\ C = J_m$) leads to the condition on the scructure constants

$$C_{kl}^{n} C_{mn}^{p} + C_{lm}^{n} C_{kn}^{p} + C_{mk}^{n} C_{ln}^{p} = 0 \tag{C.390}$$

C.7.3 Rotations SO(3) and SU(2)

By a *representation* we mean a set of matrices T_x, T_y , and T_z with the same commutation relations as the t's. The T's of Eqs() and () are an examples in which the matrices are 3×3 and the representation is said to be of dimension three.

We recall the construction in standard quantum mechanics lectures.

$$[J_1, J_2] = iJ_3, \quad [J_2, J_3] = iJ_1, \quad [J_3, J_1] = iJ_2 \tag{C.391}$$

$$\hat{J}^2 \psi_{jm} = j(j+1)\psi_{jm}, \quad \hat{J}_3 \psi_{jm} = m\psi_{jm} \tag{C.392}$$

It is convenient to define

$$J_+ = iJ_1 - J_2, \ \text{ and its complex conjugate } \ J_- = -iJ_1 - J_2 \tag{C.393}$$

the commutation relations become

$$[J_z, J_+] = J_+, \ [J_z, J_-] = -J_-, \ [J_+, J_-] = 2J_z. \tag{C.394}$$

$$J_3\psi_1 = J_3 J_+ \psi = [J_3, J_+]\psi + J_+ J_3 \psi \tag{C.395}$$

$$J_+ J_- = J^2 - J_3(J_3 - 1) \tag{C.396}$$
$$J_- J_+ = J^2 - J_3(J_3 + 1) \tag{C.397}$$

$$J_+J_-\psi_{jm} = [j(j+1) - m(m-1)] \tag{C.398}$$
$$J_-J_+\psi_{jm} = [j(j+1) - m(m+1)] \tag{C.399}$$

$$U = e^{-i\theta_i\sigma_i} \tag{C.400}$$

$$\hat{\sigma}^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{C.401}$$

$$\tau^{\mathbf{i}}\tau^{\mathbf{j}} - \tau^{\mathbf{j}}\tau^{\mathbf{i}} = \epsilon_{ijk}\tau^{\mathbf{k}} \tag{C.402}$$

$SU(2)$ is the universal cover of $SO(3)$

By definition $SU(2)$ is the group of 2×2 special unitary matricies with determinant one. If

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

where $a, b, c, d \in \mathbb{C}$ then the requirements $P{-}1 = P^*$ and $\det P = 1$ translates to

$$\overline{a} = d \text{ and } \overline{b} = -c.$$

That is

$$P = \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix}$$

and the condition that $\det P = 1$ gives the condition

$$a\overline{a} + b\overline{b} = 1$$

This says that P is fully determined by a vector $(a, b) \in \mathbb{C}^2$ of length one. If we write a, b in terms of their real and imaginary parts, then the above condition becomes

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$$

that is the unit sphere S^3 in $\mathbb{R}^4$, which is simple connected as the unit sphere S^2 in $\mathbb{R}^3$ is.

Details The Pauli matrices form a vector space

$$\tau_i = \frac{\hat{\sigma}_i}{2} \tag{C.403}$$

They form linear independent complete vector space for 2×2 matrices

$$\begin{aligned} I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \hat{\sigma}^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \hat{\sigma}^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \hat{\sigma}^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{C.404}$$

$$\begin{aligned} & a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 + bI = \mathbf{a}\cdot\sigma \\ & = \begin{pmatrix} b + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & b - a_3 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \mathbf{A} \end{aligned} \tag{C.405}$$

$$TrI = 2, \qquad Tr\sigma_i = 0 \tag{C.406}$$

Hence,

$$A = \sum_i \hat{\sigma}_i \frac{1}{2} Tr(A\hat{\sigma}_i) + (\frac{1}{2} TrA)I \tag{C.407}$$

C.7.4 Spin Direct Products

A subspace V is said to be **invariant** if it is mapped into itself by application of each matrix element of the group.(refine definition) A less trival example of a reduciable representation is the "addition of angular momentum" in quantum mechanics

We combine two irreducible representations by forming the product space $V = V_1 \otimes V_2$.

$$
\begin{aligned}
\hat{\sigma}_1^x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} & & 1 & \\ & & & 1 \\ 1 & & & \\ & 1 & & \end{pmatrix}, \\
\hat{\sigma}_1^y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} & & -i & \\ & & & -i \\ i & & & \\ & i & & \end{pmatrix}, \\
\hat{\sigma}_1^z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix},
\end{aligned}
\tag{C.408}
$$

and

$$
\begin{aligned}
\hat{\sigma}_2^x &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} & 1 & & \\ 1 & & & \\ & & & 1 \\ & & 1 & \end{pmatrix}, \\
\hat{\sigma}_2^y &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} & -i & & \\ i & & & \\ & & & -i \\ & & i & \end{pmatrix}, \\
\hat{\sigma}_1^z &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}.
\end{aligned}
\tag{C.409}
$$

Note that

$$\left[\hat{\sigma}_1^i, \hat{\sigma}_2^j\right] = 0 \tag{C.410}$$

for any $i, j \in \{x, y, z\}$ as can be shown directly or better still using the rule for multiplying direct products of matrices :$(A \otimes B)(C \otimes D) = (AC \otimes BD)$,

$$\left[\hat{\sigma}_1^i, \hat{\sigma}_2^j\right] = (\hat{\sigma}^i \otimes I)(I \otimes \hat{\sigma}^j) - (I \otimes \hat{\sigma}^j)(\hat{\sigma}^i \otimes I) = (\hat{\sigma}^i \otimes \hat{\sigma}^j) - (\hat{\sigma}^i \otimes \hat{\sigma}^j) = 0 \tag{C.411}$$

These operators act on the direct product space $V \otimes V$, with, elements

$$\eta \otimes \omega = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \otimes \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = \begin{pmatrix} \eta_1\omega_1 \\ \eta_1\omega_2 \\ \eta_2\omega_1 \\ \eta_2\omega_2 \end{pmatrix} \tag{C.412}$$

$$\begin{aligned}
|\psi_{(1,1)} > &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\
|\psi_{(1,0)} > &= \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\
|\psi_{(1,-1)} > &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\
|\psi_{(0,0)} >= &= \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}.
\end{aligned} \tag{C.413}$$

This result

$$S := \frac{1}{2}(\hat{\sigma} \otimes I + I \otimes \hat{\sigma}) \tag{C.414}$$

are generators of a *reducible* 4-dimensional reperentation of $su(2)$ as is easily seen,

$$\begin{aligned}
[S_i, S_j] &= \frac{1}{4}\big[(\hat{\sigma}_i \otimes I + I \otimes \hat{\sigma}_i)(\hat{\sigma}_j \otimes I + I \otimes \hat{\sigma}_j) - i \leftrightarrow j\big] \\
&= \frac{1}{4}\big[(\hat{\sigma}_i\hat{\sigma}_j \otimes I + \hat{\sigma}_i \otimes \hat{\sigma}_j + \hat{\sigma}_j \otimes \hat{\sigma}_i + I \otimes \hat{\sigma}_i\hat{\sigma}_j - i \leftrightarrow j\big] \\
&= \frac{1}{2}\big[\frac{1}{2}(\hat{\sigma}_i\hat{\sigma}_j - \hat{\sigma}_j\hat{\sigma}_i) \otimes I + \frac{1}{2}I \otimes (\hat{\sigma}_i\hat{\sigma}_j - \hat{\sigma}_j\hat{\sigma}_i)\big] \\
&= \epsilon_{ijk}\frac{1}{2}(\hat{\sigma}_k \otimes I + I \otimes \hat{\sigma}_k) \\
&= \epsilon_{ijk} S_k.
\end{aligned} \tag{C.415}$$

There exists a unitary matrix U such that

$$USU^\dagger = \begin{pmatrix} S_{(S=1)} & \\ & S_{(S=0)} \end{pmatrix}, \tag{C.416}$$

where $S_{(S=1)}$ and $S_{(S=0)}$ are irreducible representaions of dimension 3 and 1, respectively.

$$\begin{aligned} S^2 &= \frac{1}{4}(\hat{\sigma} \otimes I + I \otimes \hat{\sigma})^2 \\ &= \frac{1}{4}(\hat{\sigma}^2 \otimes I + 2\hat{\sigma} \otimes \hat{\sigma} + I \otimes \hat{\sigma}^2) \\ &= \frac{1}{2}\hat{\sigma} \otimes \hat{\sigma} + \frac{3}{2} I \otimes I, \end{aligned} \tag{C.417}$$

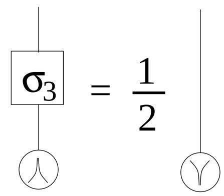
since $\hat{\sigma}^2 = 3I$.

$$\vec{S}^z = \hat{\sigma}^z \otimes I + I \otimes \hat{\sigma}^z \tag{C.418}$$

$$\vec{S}^z|S, S^z >= S^z|S, S^z > \tag{C.419}$$

$$\vec{S}^2|S, S^z >= S(S+1)|S, S^z >, \quad S^z = -S, \ldots, S; \;\; S = 0, 1. \tag{C.420}$$

I will first go through the explicit calculation of the action of the z-component of the angular momentum, $\hat{J}_3$, and the total angular momentum, $\hat{J}^2$, so the reader may have a better feel for what is going on in the abstract diagrammatic calculation (and to appreciate the simplicity of the diagrammatic version). Moreover, I wish to go through the calculation for the total angular momentum operator both ways because of its importance in finding the (main sequence of) eigenvalues of the area operator (see appendix on geometric operators).



Details

$$
\begin{aligned}
\omega_{AB\ldots F}(i=0) &= \delta_{A0}\delta_{B0}\ldots\delta_{E0}\delta_{F0} && \text{(C.421)}\\
\omega_{AB\ldots F}(i=1) &= \delta_{A1}\delta_{B0}\delta_{C0}\ldots\delta_{E0}\delta_{F0}\\
&\quad +\delta_{A0}\delta_{B1}\delta_{C0}\ldots\delta_{E0}\delta_{F0}\\
&\quad +\delta_{A0}\delta_{B0}\delta_{C1}\ldots\delta_{E0}\delta_{F0}+\ldots\\
&\vdots\\
\omega_{AB\ldots F}(i=2s) &= \delta_{A1}\delta_{B1}\ldots\delta_{E1}\delta_{F1} && \text{(C.422)}
\end{aligned}
$$

$$\hat{\sigma}_z = \sum_{k=1}^{2s+1} 1\otimes\cdots\otimes 1\left(\frac{\hat{\sigma}^3}{2}\right)\otimes\cdots\otimes 1 \tag{C.423}$$

which written in component form reads

$$
\begin{aligned}
\frac{1}{2}\hbar\hat{\sigma}^3_{AA'BB'\ldots EE'FF'} &:= \frac{1}{2}\hbar\hat{\sigma}^3_{AA'}\delta_{BB'}\ldots\delta_{FF'}\\
&\vdots\\
&\quad \frac{1}{2}\hbar\hat{\sigma}^3_{FF'}\delta_{AA'}\delta_{BB'}\ldots\delta_{EE'}\\
&\quad (AA';BB';\ldots;EE';FF'=0,1). && \text{(C.424)}
\end{aligned}
$$

We verify that the "states" $\omega(i)$ fulfill the eigenvalue equation

$$
\begin{aligned}
\hbar\frac{1}{2}\hat{\sigma}^3\omega(i) &= \hbar(s-i)\omega(i)\\
i &= 0,1,\ldots,2s. && \text{(C.425)}
\end{aligned}
$$

Using

$$\hat{\sigma}^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{C.426}$$

and (C.424) we get

$$\begin{aligned}
\left(\frac{1}{2}\hbar\,\omega(i=0)\right)_{AB\ldots F} &= \hbar\frac{1}{2}\hat{\sigma}^3_{AA'BB'\ldots EE'FF'}\,\omega_{A'B'\ldots E'F'}(i=0) \\
&= \hbar\frac{1}{2}\hat{\sigma}^3_{AA'}\delta_{A'0}\delta_{B0}\ldots\delta_{E0}\delta_{F0} + \\
&\quad + \delta_{A0}\hbar\frac{1}{2}\hat{\sigma}^3_{BB'}\delta_{B'0}\ldots\delta_{E0}\delta_{F0} + \ldots \\
&\quad + \delta_{A0}\delta_{B0}\ldots\delta_{E0}\hbar\frac{1}{2}\hat{\sigma}^3_{FF'}\delta_{F'0} \\
&= \frac{\hbar}{2}2s(\delta_{A0}\delta_{B0}\ldots\delta_{F0}) = \hbar s\omega_{AB\ldots F}(i=0)
\end{aligned} \tag{C.427}$$

$$\begin{aligned}
\left(\frac{1}{2}\hbar\,\omega(i=1)\right)_{AB\ldots F} &= \hbar\frac{1}{2}\hat{\sigma}^3_{AA'}\delta_{A'1}\delta_{B0}\ldots\delta_{E0}\delta_{F0} + \delta_{A1}\frac{\hbar}{2}\hat{\sigma}^3_{BB'}\delta_{B'0}\ldots\delta_{F0} + \\
&+ \cdots + \delta_{A1}\delta_{B0}\ldots\frac{\hbar}{2}\hat{\sigma}^3_{FF'}\delta_{F'0} + \frac{\hbar}{2}\hat{\sigma}^3_{AA'}\delta_{A'0}\delta_{B1}\ldots\delta_{F0} \\
&+ \delta_{A0}\frac{\hbar}{2}\hat{\sigma}^3_{BB'}\delta_{B'1}\ldots\delta_{F0} + \ldots \\
&+ \delta_{A0}\delta_{B1}\ldots\frac{\hbar}{2}\hat{\sigma}^3_{FF'}\delta_{F'1} + \ldots \\
&= \frac{\hbar}{2}\left[-\delta_{A1}\delta_{B0}\ldots\delta_{F0} + (2s-1)\times\delta_{A1}\delta_{B0}\ldots\delta_{F0} + \right. \\
&+ \delta_{A0}\delta_{B1}\ldots\delta_{F0} - \delta_{A0}\delta_{B1}\ldots\delta_{F0} + \\
&+ \left.(2s-3)\times\delta_{A0}\delta_{B1}\ldots\delta_{F0} + \ldots\right] \\
&= \hbar(s-1)\omega_{AB\ldots F}(i=1),
\end{aligned} \tag{C.428}$$

$$\begin{aligned}
&\vdots \\
\left(\frac{1}{2}\hbar\,\omega(i=2s)\right)_{AB\ldots F} &= \\
&= \hbar\frac{1}{2}\hat{\sigma}^3_{AA'}\delta_{A'1}\delta_{B1}\ldots\delta_{F1} + \delta_{A1}\frac{\hbar}{2}\hat{\sigma}^3_{BB'}\delta_{B'1}\ldots\delta_{E1}\delta_{F1} + \ldots \\
&+ \cdots + \delta_{A1}\delta_{B1}\ldots\frac{\hbar}{2}\hat{\sigma}^3_{FF'}\delta_{F'1} \\
&= -\hbar s\omega_{AB\ldots F}(i=2s).
\end{aligned} \tag{C.429}$$

$\square$

$$\hat{\sigma}^2 = \sum_{k=1}^{2s+1} \sum_{k'=1}^{2s+1} 1 \otimes \cdots \otimes \left(\frac{\hat{\sigma}^k}{2}\right) \otimes \cdots \otimes \left(\frac{\hat{\sigma}^{k'}}{2}\right) \otimes \cdots \otimes 1 \tag{C.430}$$

$$(\hbar^2/4)\hat{\sigma}^2_{AA'BB'\dots FF'} \tag{C.431}$$

has eigenvalue $\hbar^2 s(s+1)$

$$\left(\frac{1}{4}\hat{\sigma}^2\right)_{AA'BB'\dots FF'} = \frac{1}{4} \tag{C.432}$$

$$\hat{\sigma}^2 = \sum_{k=1}^{2s+1} 1 \otimes \cdots \otimes \left(\frac{\hat{\sigma}^k \hat{\sigma}^k}{4}\right) \otimes \cdots \otimes 1 \tag{C.433}$$

$$\hat{\sigma}^2_{AA'} = (\hat{\sigma}^1)^2_{AA'} + (\hat{\sigma}^2)^2_{AA'} + (\hat{\sigma}^3)^2_{AA'} = 3\delta_{AA'} \tag{C.434}$$

$$\begin{aligned} \left(\frac{1}{4}\hat{\sigma}^2\right)\omega(i=0) &= \left[\frac{3}{2}s + \frac{1}{4}(2s-1)2s\right]\omega(i=0) \\ &= s(s+1)\omega(i=0). \end{aligned} \tag{C.435}$$

$$\begin{aligned} \left(\frac{1}{4}\hat{\sigma}_{AA'}\hat{\sigma}_{BB'}\right)\delta_{A'1}\delta_{B'0} &= \\ &= \left(\hat{\sigma}^1_{AA'}\hat{\sigma}^1_{BB'} + \hat{\sigma}^2_{AA'}\hat{\sigma}^2_{BB'} + \hat{\sigma}^3_{AA'}\hat{\sigma}^3_{BB'}\right)\delta_{A'1}\delta_{B'0} \\ &= \frac{1}{4}\left((+1)\delta_{A0}\delta_{B1} + (+1)\delta_{A0}\delta_{B1} + (-1)\delta_{A1}\delta_{B0}\right) \end{aligned} \tag{C.436}$$

$$\begin{aligned} \left(\frac{1}{4}\hat{\sigma}_{AA'}\hat{\sigma}_{BB'}\right)(\delta_{A'1}\delta_{B'0} + \delta_{A'0}\delta_{B'1}) &= \\ \frac{1}{4}(\delta_{A0}\delta_{B1} + \delta_{A1}\delta_{B0}) \end{aligned} \tag{C.437}$$

C.7.5 Direct Products and Clebsch-Gordan Coefficients

$$\tau^{\mathbf{i}}_{(\mathbf{j})} = \sum_{k=1}^{2s+1} 1 \otimes \cdots \otimes \left(\frac{\hat{\sigma}^k}{2}\right) \otimes \cdots \otimes 1 \tag{C.438}$$

We wish to calculate $\tau^{\mathbf{i}}_{(\mathbf{j})}\tau^{\mathbf{j}}_{(\mathbf{j})} - \tau^{\mathbf{j}}_{(\mathbf{j})}\tau^{\mathbf{i}}_{(\mathbf{j})}$. The terms of (5.24) for $k \neq k$ wont contribute to the commutator as the order of multiplication irrelevant,

$$\sum_{k \neq k'} \left(1 \otimes \cdots \otimes \left(\frac{\hat{\sigma}^k}{2}\right) \otimes \cdots \otimes 1\right) \left(1 \otimes \cdots \otimes \left(\frac{\hat{\sigma}^{k'}}{2}\right) \otimes \cdots \otimes 1\right) - (k \leftrightarrow k') = 0 \tag{C.439}$$

$$\begin{aligned}
\tau^{\mathbf{i}}_{(\mathbf{j})}\tau^{\mathbf{j}}_{(\mathbf{j})} - \tau^{\mathbf{j}}_{(\mathbf{j})}\tau^{\mathbf{i}}_{(\mathbf{j})} &= \\
&= \sum_{k=1}^{2s+1} \left[1 \otimes \cdots \otimes \left(\frac{\hat{\sigma}^i}{2}\right)\left(\frac{\hat{\sigma}^j}{2}\right) \otimes \cdots \otimes 1 - 1 \otimes \cdots \otimes \left(\frac{\hat{\sigma}^j}{2}\right)\left(\frac{\hat{\sigma}^i}{2}\right) \otimes \cdots \otimes 1 \right] \\
&= \sum_{k=1}^{2s+1} 1 \otimes \cdots \otimes \left(\frac{\hat{\sigma}^i\hat{\sigma}^j}{4} - \frac{\hat{\sigma}^j\hat{\sigma}^i}{4}\right) \otimes \cdots \otimes 1 \\
&= \epsilon_{ijk} \sum_{k=1}^{2s+1} 1 \otimes \cdots \otimes \frac{\hat{\sigma}^k}{2} \otimes \cdots \otimes 1 \\
&= \epsilon_{ijk}\tau^{\mathbf{k}}_{(\mathbf{j})}
\end{aligned} \tag{C.440}$$

$A \otimes B$

$$a_{ij}b_{kl} = c_{ik;jl}. \tag{C.441}$$

the row and column labels of the matrix elements of C are *composite* labels: the row label ik, is obtained from the row labels of the matrix elements of A and B and the column label, jl is obtained from the corresponding column labels.

$$\begin{aligned}
\Delta_j(\theta) &= \sum_{m=-j}^{j} e^{-im\theta} \\
&= e^{ij\theta}\frac{e^{-i(2j+1)\theta} -}{e^{-i\theta} - 1} \\
&= \frac{e^{i(j+\frac{1}{2})\theta} - e^{-i(j+\frac{1}{2})\theta}}{e^{i\frac{1}{2}\theta} - e^{-i\frac{1}{2}\theta}} \\
&= \frac{\sin\left[(j+\frac{1}{2})\theta\right]}{\sin(\frac{1}{2}\theta)}
\end{aligned} \tag{C.442}$$

$$\int_0^{2\pi} \overline{\Delta_k(\theta)} \Delta_j(\theta) \frac{1-\cos\theta}{2\pi}\, d\theta = \delta_{kj}. \tag{C.443}$$

Since the spinor indices take only one of two values an antisymmetrizable over three indices of a multivalent spinor $\tau_{\dots ABC \dots}$ is zero. In particular we have the *Jacobi identity*

$$\epsilon_{A[B}\epsilon_{CD]} = 0 = \epsilon_{AB}\epsilon_{CD} + \epsilon_{AC}\epsilon_{DB} + \epsilon_{AD}\epsilon_{BC} \tag{C.444}$$

$$\tau_{\dots AB \dots} = \tau_{\dots (AB) \dots} + \frac{1}{2}\epsilon_{AB}\tau_{\dots C}{}^{C}{}_{\dots}. \tag{C.445}$$

The proof

$$\epsilon_{AB}\tau_{C}{}^{C} - \tau_{AB} + \tau_{BA} = 0 \tag{C.446}$$

$$\tau_{[AB]} = \frac{1}{2}\epsilon_{AB}\tau_{C}{}^{C}. \tag{C.447}$$

Since $\tau_{AB} = \tau_{(AB)} + \tau_{[AB]}$ the result follows

$$\tau_{AB} = \tau_{(AB)} + \frac{1}{2}\epsilon_{AB}\tau_{C}{}^{C} \tag{C.448}$$

It should be evident that this result will also apply to the more general case where we consider only two particular indices of a multivalent spinor $\tau_{\dots ABC \dots}$ of valence > 2, hence we have the result (C.445).

The same is true for the followings cases.

(a) Case τ_{AB} we already have

(b) Case τ_{ABC}

$$\begin{aligned} 3\tau_{(ABC)} &= \tau_{A(BC)} + \tau_{B(AC)} + \tau_{C(AB)} \\ &= 3\tau_{A(BC)} - (\tau_{A(BC)} - \tau_{B(AC)}) - (\tau_{A(BC)} - \tau_{C(AB)}) \\ &= 3\tau_{A(BC)} - \epsilon_{AB}\sigma_C - \epsilon_{AC}\sigma_B \end{aligned} \tag{C.449}$$

where $\sigma_C = \epsilon^{AB}(\tau_{A(BC)} - \tau_{B(AC)})$.

This rearanged gives

$$\tau_{A(BC)} = \tau_{(ABC)} + \frac{1}{3}\epsilon_{AB}\sigma_C + \frac{1}{3}\epsilon_{AC}\sigma_B. \tag{C.450}$$

$$\tau_{A(BC)} = \tau_{ABC} - \frac{1}{2}\epsilon_{BC}\tau_{AD}{}^{D} \tag{C.451}$$

Using (C.451) in (C.450) we have the desired expansion for τ_{ABC}

$$\tau_{ABC} = \tau_{(ABC)} + \frac{1}{2}\epsilon_{BC}\tau_{AD}{}^{D} + \frac{1}{3}\epsilon_{AB}\ \sigma_{C} + \frac{1}{3}\epsilon_{AC}\ \sigma_{B}. \tag{C.452}$$

or more generally

$$\tau_{\ldots ABC\ldots} = \tau_{\ldots(ABC)\ldots} + \frac{1}{2}\epsilon_{BC}\tau_{\ldots AD}{}^{D}{}_{\ldots} + \frac{1}{3}\epsilon_{AB}\ \sigma_{\ldots C\ldots} + \frac{1}{3}\epsilon_{AC}\ \sigma_{\ldots B\ldots}. \tag{C.453}$$

where $\sigma_{\ldots C\ldots} = \epsilon^{AB}(\tau_{\ldots A(BC)\ldots} - \tau_{\ldots B(AC)\ldots})$

(c) Case $\tau_{A\ldots F}$:

Proof is by induction

Any spinor $\tau_{(A\ldots F)}$

$$n\tau_{(ABC\ldots F)} = \tau_{A(BC\ldots F)} + \tau_{B(AC\ldots F)} + \cdots + \tau_{F(BC\ldots A)}. \tag{C.454}$$

$$\tau_{A(BCD\ldots F)} - \tau_{B(ACD\ldots F)} = \frac{1}{2}\epsilon_{AB}\sigma_{(CD\ldots F)} \tag{C.455}$$

where $\sigma_{(CD\ldots F)} = \epsilon^{AB}(\tau_{A(BCD\ldots F)} - \tau_{B(ACD\ldots F)})$

$$n\tau_{(ABC\ldots F)} = n\tau_{A(BC\ldots F)} - (\tau_{A(BC\ldots F)} - \tau_{B(AC\ldots F)}) - \cdots - (\tau_{A(BC\ldots F)} - \tau_{F(BC\ldots A)}) \tag{C.456}$$

$$\tau_{A(BC\ldots F)} = \tau_{(ABC\ldots F)} + \frac{1}{n}\epsilon_{AB}\rho_{(C\ldots F)} + \cdots + \frac{1}{n}\epsilon_{AF}\rho_{(B\ldots F)} \tag{C.457}$$

where

$$\rho_{(C\ldots F)} = \epsilon^{AB}(\tau_{A(BC\ldots F)} - \tau_{B(AC\ldots F)}) \tag{C.458}$$

C.7.6 Recoupling Theory

Recoupling Theory of Three Angular Momenta - 6-j-Symbols

$$|j_{12}(j_1, j_2), j(j_{12}, j_3) > \; = \; | \tag{C.459}$$

C.7.7 SO(3,1) and SL(2,C)

$$x' = \frac{x + vt}{\sqrt{1 - c^2 t^2}}, \quad y' = y, \quad z' = z, \quad t' = \frac{t + vx/c^2}{\sqrt{1 - c^2 t^2}} \tag{C.460}$$

$$\gamma = \cosh\phi, \quad \gamma\beta = \sinh\phi, \tag{C.461}$$

$$\begin{pmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix} = \begin{pmatrix} \cosh\phi & \sinh\phi & 0 & 0 \\ \sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}. \tag{C.462}$$

$$K_x = -i \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{C.463}$$

$$K_y = -i \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_z = -i \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{C.464}$$

$$J_x = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad J_y = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad J_z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{C.465}$$

$$\begin{aligned} [K_x, K_y] &= -iJ_z \\ [J_x, K_x] &= 0 \\ [J_x, K_y] &= iK_z \end{aligned} \tag{C.466}$$

$$\begin{aligned} A &= \frac{1}{2}(J + iK), \\ B &= \frac{1}{2}(J - iK). \end{aligned} \tag{C.467}$$

$$\begin{aligned} [A_x, A_y] &= iA_z \text{and cyclic perms,} \\ [B_x, B_x] &= iB_z \text{and cyclic perms,} \\ [A_x, B_y] &= 0 \ \ (i, j = x, y, z). \end{aligned} \tag{C.468}$$

This shows that A and B each generate a group $SU(2)$ $SU(2) \otimes SU(2)$

$$X(x) := \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix} \tag{C.469}$$

unique vector x^μ

$$x^\mu = \frac{1}{2}\mathrm{tr}(\sigma_\mu X) \tag{C.470}$$

$$\begin{aligned} \det X(x) &> 0 \quad \text{if } x \text{ is a timelike vector} \\ &= 0 \quad \text{if } x \text{ is on the light cone} \\ &< 0 \quad \text{if } x \text{ is a spacelikevector} \end{aligned} \tag{C.471}$$

$$A = \exp\left(-i\frac{\theta}{2}(\vec{n}\cdot\sigma)\right) = \cos\frac{\theta}{2}I - i(\vec{n}\cdot\sigma)\sin\frac{\theta}{2} \tag{C.472}$$

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, C) \tag{C.473}$$

$$A = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} \begin{pmatrix} \cos\beta & \sin\beta e^{i\gamma} \\ -\sin\beta e^{-i\gamma} & \cos\beta \end{pmatrix} B \tag{C.474}$$

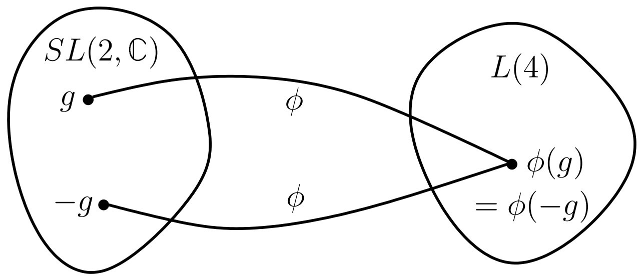


Figure C.10:

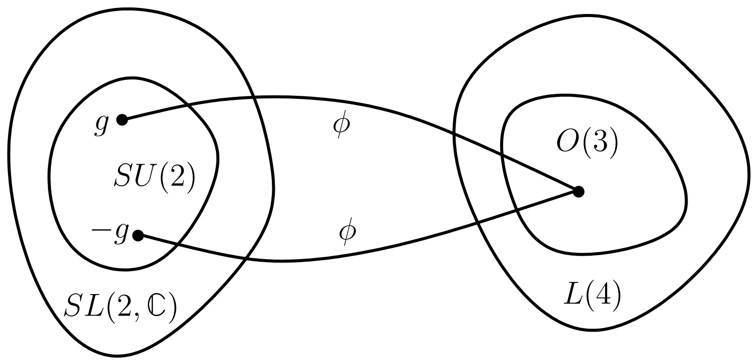


Figure C.11:

C.7.8 SO(4)

$$U(g) = 1 - \frac{1}{2} J_i^j \epsilon_j^i + \mathcal{O}, \tag{C.475}$$

where the J_i^j are $N \times N$ matrices

$$[J_i^j, J_k^l] = \delta_k^j J_i^l - \delta_l^i J_k^j + \delta_k^i J_j^l - \delta_l^j J_i^k \ , \quad i, j, kl = 1, \dots, n. \tag{C.476}$$

$$SO(4) = SO(3) \oplus SO(3). \tag{C.477}$$

$$L^{\pm} := \frac{1}{2}(L \pm M') \tag{C.478}$$

The structure constants with respect to this new basis are given by

$$[(L^{\pm})^i, (L^{\pm})^j] = i\hbar \epsilon^{ij}{}_k (L^{\pm})^k \ , \quad [(L^+)^i, (L^-)^j] = 0 \tag{C.479}$$

C.7.9 Conformal Group

In Minkowskian spacetime a *conformal transformation* are coordinate transformations $x \to x'(x)$ which are such that the induced change in the metric is a positive rescaling by a positive function:

$$ds'^2 = \Omega(x)^2\, ds^2 \tag{C.480}$$

where $\Omega(x)$ is a real-valued function. Geometrically conformal transformations leaves thae *angles* but changes distances; it involves dilations (rescalings). Tranformations with this property are used in the design of geographic maps. Penrose diagrams bring infinity onto a page.

$$ds^2 = dt^2 - dr^2 - r^2(d\theta^2 + \sin\theta d\phi^2) \tag{C.481}$$

We introduce double null coordinates

$$u = t + r \tag{C.482}$$
$$v = t - r \tag{C.483}$$

$dudv = (dt + dr)(dt - dr) = dt^2 - dr^2$ and $(u - v)^2 = (t + r - t + r)^2 = 4r^2$ so the line element becomes

$$ds^2 = dudv - \frac{1}{4}(u - v)^2(d\theta^2 + \sin\theta d\phi^2). \tag{C.484}$$

$$p = \tan^{-1} u, \quad q = \tan^{-1} v. \tag{C.485}$$

$$ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta = \frac{1}{4}\sec^2 p \sec^2 q[4dpdq - \sin^2(p - q)(d\theta^2 + \sin\theta d\phi^2)] \tag{C.486}$$

$$d\bar{s}^2 = \bar{g}_{\alpha\beta}dx^\alpha dx^\beta = dpdq - \sin^2(p - q)(d\theta^2 + \sin\theta d\phi^2) \tag{C.487}$$

$$\Omega = \frac{1}{4}\sec^2 p \sec^2 q \tag{C.488}$$

Now finally we introduce the coordinates

$$t' = p + q \tag{C.489}$$
$$r' = p - q \tag{C.490}$$

with the coordinate range

$$-\pi < t' + r' < \pi, \tag{C.491}$$
$$-\pi < t' - r' < \pi, \tag{C.492}$$
$$r' > 0 \tag{C.493}$$

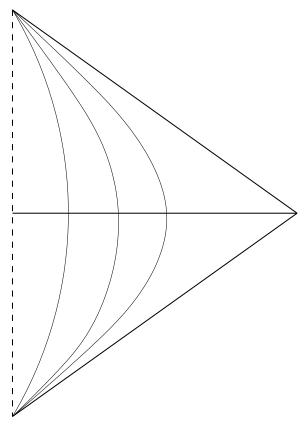


Figure C.12: Penrose diagram for Minkowskian spacetime.

C.7.10 Group Integration: The Haar Measure

For a compacct group, we can replace the sums over group elements that occur in the representation theory of finite groups, by convergent integrals over the group elements using the invariant Haar measure, which is usually denoted by $d[g]$. The invariance property is expressed by $d[g_1 g] = d[g]$. This allows us to make a change of variables transfromation, $g \to g_1 g$, identaical to that which played such an important role in deriving the finite group theorems. Consequently, all the results from finite groups, such as the existence of an invariant inner product and the orthoganality theorems, acn be taken over by the simple replacement of a sum by an integral.

As a vector space, $R(G)$ has a distinguished basis given by the characters $\chi_i(g)$ of the irreducible representations. Recall that the orthog-onality relations for characters are essentially the same as in the finite group case, with the sum over group elements replace by an integral

$$\int \overline{\chi_i(g)} \chi_j(g) dg = \delta_{ij} \tag{C.494}$$

where i and j are labels for irreducible representations and dg is the standard Haar measure, normalized so that the volume of g is 1.

For the case of finite group one important property of these groups is the rearrangement theoremis given for fixed S,

$$\sum_R f(R) = \sum_R f(SR) \tag{C.495}$$

$$\int f(R)dR = \int f(SR)dR \tag{C.496}$$

To clarify the relation to the finite group rearangement theorem we write the integral on the left-hand side of (C.495) as an integral over the parameters. the density of group elements is arranged so that the density of points at SR is the same as that at R.

$$\int f(R)\, dR = \int f(R)\mu(R)\, da, \tag{C.497}$$

where $\mu(R)$ is the density of group elements in the parameter space in the neighborhood of R.

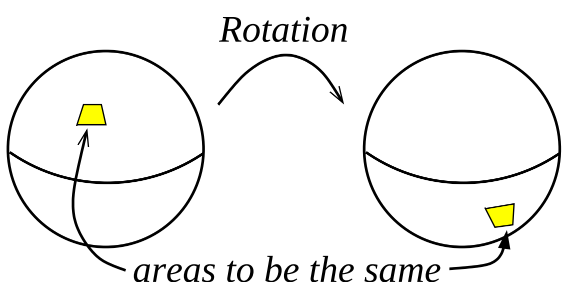


Figure C.13: Haarmeas1.

uniform measure over a 3-sphere S^3

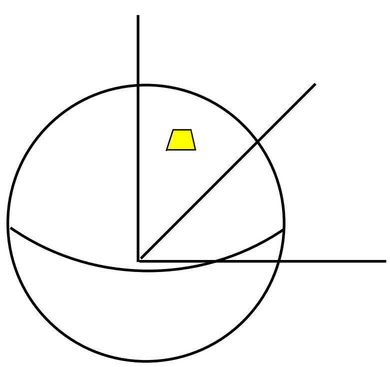


Figure C.14: HaarmeasSO3.

Of course the measure is

$$dU = d\theta \cos\theta d\phi \tag{C.498}$$

as under a rotation

$$d\theta \cos\theta d\phi \mapsto d\theta' \cos\theta' d\phi' \tag{C.499}$$

Approach that is easily extended to other Lie groups.

$$dU = \pi^{-2} da_1 da_2 da_3 \delta\left(\sum_{i=1}^{2} a_i^2 - 1\right) \tag{C.500}$$

$SU(2)$ a 3-sphere in 4-dimensional Euclidean space.

$$dU = \pi^{-2} da_0 da_1 da_2 da_3 \delta\left(\sum_{i=0}^{3} a_i^2 - 1\right) \tag{C.501}$$

$$\mu(C) = \mu(0)/\left(\frac{\partial c}{\partial a}\right)_{a=0} \tag{C.502}$$

$$\begin{aligned}
&\int dU U_{ij} = 0 \\
&\int dU U_{i_1 j_1} U_{i_2 j_2} = \frac{1}{2}\epsilon_{i_1 j_1}\epsilon_{i_2 j_2} \\
&\int dU U_{ij}(U^{-1})_{kl} = \frac{1}{2}\delta_{jk}\delta_{il}
\end{aligned} \tag{C.503}$$

$$\int dU |\mathrm{Tr} U|^2 = 1 \tag{C.504}$$

Details

$$\int dU U_{ij}(U^{-1})_{kl} = A\delta_{jk}\delta_{il} \tag{C.505}$$

The constant A is easily found by setting $j = k$ summing over the index

$$\int dU \delta_{il} = A\delta_{jj}\delta_{il} = 2A\delta_{il}, \tag{C.506}$$

noting $\int dU = 1$ implies $A = 1/2$.

$$\int dU |\mathrm{Tr} U|^2 = \int dU U_{ij}(U^{\dagger})_{kl} = \int dU U_{ij}(U^{-1})_{kl} = \frac{1}{2}\delta_{ji}\delta_{ij} = \frac{1}{2}\delta_{jj} = 1 \tag{C.507}$$

For any compact group G the Haar measure is the unique measure dU on G which obeys invarianc (C.496) and normalization: $\int_G dU = 1$.

$$dc = J\left[\frac{\partial c}{\partial a}\right] da \tag{C.508}$$

where $J[\partial c/\partial a]$ is the *Jacobian* defined by

$$J\left[\frac{\partial c}{\partial a}\right] = \frac{\partial(c_1, c_2, \ldots, c_h)}{\partial(a_1, a_2, \ldots, a_h)} \equiv det\left|\frac{\partial c_i}{\partial a_i}\right| \tag{C.509}$$

Example 1: $SO(2)$ $R(\theta_1)R(\theta_2) = R(\theta_1 + \theta_2)$. Thus, from $R(\theta') = R(\theta)R(\epsilon) = R(\theta + \epsilon)$, where $R(\epsilon)$ is the infintesimal transformation, we have

$$\theta' = \theta + \epsilon \tag{C.510}$$

Accordingly $d\theta'/d\epsilon = 1$, so we have that $\mu(\theta) = \mu_0$. Through normalization

$$\int_{-\pi}^{\pi} \mu(\theta) d\theta = 2\pi\mu_0 = 1 \tag{C.511}$$

we obtain $\mu(\theta) = 1/(2\pi)$.

Unitary Group $U(2)$

The unitary group $U(N)$ has N^2 generators

$$[\mathcal{C}_{im}, \mathcal{C}_{jn}] = \delta_{jm}\mathcal{C}_{in} - \delta_{in}\mathcal{C}_{jm}. \tag{C.512}$$

a representation of $U(2)$ is

$$(\mathcal{C}_{im}) = (\delta_{im}) \tag{C.513}$$

where (δ_{im} denotes the matrix which has a "1" at the intersection of the ith row and the mth column and zeros everywhere else i.e.

$$\begin{aligned} \mathcal{C}_{11} &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathcal{C}_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ \mathcal{C}_{21} &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \mathcal{C}_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \tag{C.514}$$

An arbitrary group element of $U(2)$ is given by

$$\exp\left(-i\sum_{k,l=1}^{2}\theta_{kl}\mathcal{C}_{kl}\right) \tag{C.515}$$

where the "angles" θ parametrize the group tranformation.

The transition to $SU(2)$ can be made by constructing traceless matrices from the $\mathcal{C}_{kl}$

All in all

$$\begin{aligned}\tilde{\mathcal{C}}_{11} &= \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2}\end{pmatrix}, \quad \tilde{\mathcal{C}}_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0\end{pmatrix}, \\ \tilde{\mathcal{C}}_{21} &= \begin{pmatrix} 0 & 0 \\ 1 & 0\end{pmatrix}, \quad \tilde{\mathcal{C}}_{22} = \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{1}{2}\end{pmatrix}\end{aligned} \tag{C.516}$$

Finally for $\mu(c)$ we obtain

$$\mu(c) 1/J[\partial c/\partial a]_{a=0} \tag{C.517}$$

U(2), SU(2) The measure is given by

$$\begin{aligned}\mu(U(2)) &= |\epsilon_1-\epsilon_2|^2 = |\exp(-i\theta_{11})-\exp(-i\theta_{22})|^2 \\ &= |\exp\left(-\frac{1}{2}i(\theta_{11}+\theta_{22})\right)|^2 \times |\exp\left(-\frac{1}{2}i(\theta_{11}-\theta_{22})\right) \\ &\quad |\exp\left(\frac{1}{2}i(\theta_{11}-\theta_{22})\right)|^2 \\ &= 4\sin^2\left(\frac{1}{2}(\theta_{11}-\theta_{22})\right).\end{aligned} \tag{C.518}$$

restriction to $SU(2)$ yields $\theta_{11}+\theta_{22}=0$. With $\theta_{11}-\theta_{22}=\phi$ then

$$\mu(SU(2)) = 4\sin^2\frac{1}{2}\phi. \tag{C.519}$$

Normalization $C\int_0^{4\pi}\mu(SU(2))\,d\phi = 1$ determines constant C;

$$C\int_0^{4\pi} d\phi\sin^2\frac{1}{2}\phi = \frac{C}{2}\int_0^{4\pi} d\phi\,[1-\cos\phi] = \frac{C}{2}\int_0^{4\pi} d\phi = 2\pi \tag{C.520}$$

C.7.11 Peter-Weyl theorem

The irreducible representation functions $D^l(R)_m^{m'}$ satisfy orthogonality and completeness relations. In fact, they form a complete basis in the space of square integrable functions defined on the group parameter space. This is the Peter-Weyl theorem. That the irreducible representation functions form a complete basis for square-integrable functions $f(R) \in L^2$ can be expressed as

$$f(R) = \sum_{\alpha ab} f_\alpha^{ab} D_{ab}^{(\alpha)}(R) \tag{C.521}$$

Moreover the Peter-Weyl theorem states that the χ_r form a basis of the space of square-integrable class functions on the group G. That means that every square-integrable function $f(U)$, obeying

$$f(U) = f(VUV^{-1}), \tag{C.522}$$

can be expanded into characters:

$$f(U) = \sum_{r \in G} f_r \chi_r(U) \tag{C.523}$$

$$f_r = \int dU \overline{\chi_r}(U) f(U). \tag{C.524}$$

(compare with discrete case: $\lambda(c_l) = \sum_\alpha a_\alpha \chi^{(\alpha)}(c_l)$ where $a_\alpha = \sum_k N_k \chi^{(\alpha)}(c_k)^* \lambda(c_k)$) the completeness relation can be written as

$$\sum_{r \in G} f_r \chi_r(U) \overline{\chi_r}(V) = \delta(UV^{-1}). \tag{C.525}$$

$$\int dU \overline{\chi_r}(U) \chi_s(U) = \delta_{rs}. \tag{C.526}$$

The invariant $\delta-$function on G is defined by means of

$$\int dU f(U) \delta(UV^{-1}) = f(V), \tag{C.527}$$

and obeys

$$\delta(U) = \delta(U^{-1}). \tag{C.528}$$

The charactres are

$$\chi_j(U) = \frac{\sin\left(j+\frac{1}{2}\right)\phi}{\sin\frac{1}{2}\phi}. \tag{C.529}$$

$$f_j = \frac{1}{2\pi}\int_0^{2\pi} d\phi\ \sin\frac{1}{2}\phi\ \sin\left(j+\frac{1}{2}\right)\phi\ f(\phi). \tag{C.530}$$

As an example consider the group $U(1)$

$$D_{(\mu)}(\theta) = e^{i\mu\theta}, \tag{C.531}$$

where $\mu = 0, \pm1, \pm2, \ldots$.

$$D_{(\mu)}(\theta + 2\pi) = D_{(\mu)}(\theta). \tag{C.532}$$

The Haar measure is

$$\Omega = \frac{1}{2\pi}\, d\theta. \tag{C.533}$$

Theorem applied to $U(1)$ gives

$$\frac{1}{2\pi}\int_0^{2\pi} d\theta e^{-i\mu\theta} e^{i\nu\theta} = \delta_{\mu\nu} \tag{C.534}$$

where $0, \pm1, \pm2, \ldots$.

The *Peter-Weyl Theorem* applied to $U(1)$ gives the Fourier series theory:

$$f(\theta) = \sum_{-\infty}^{\infty} f_n \frac{e^{in\theta}}{\sqrt{2\pi}}, \tag{C.535}$$

where $f(\theta) \in L^2(U(1))$.

C.7.12 Analogies

$$\int dx\delta(x)f(x)=f(0) \quad\Longleftrightarrow\quad \int_G dU\delta(U)f(U)=f(I) \tag{C.536}$$

$$\int dy\delta(y-a)f(y)=f(a) \quad\Longleftrightarrow\quad \int_G dU\delta(U\gamma^{-1})f(U)=f(\gamma) \tag{C.537}$$

$$f(x)=\sum_n a_n e_n(x) \quad\Longleftrightarrow\quad f(g)=\sum_j a_j\chi_j(g) \quad (f(g)=f(\gamma g\gamma^{-1}))$$

$$a_n=\int\overline{e_n(x)}f(x)dx \quad\Longleftrightarrow\quad a_j=\int_G\overline{\chi_j(g)}f(g)dg \tag{C.538}$$

$$\delta(x-y)=\sum_n\overline{e_n(x)}e_n(y) \quad\Longleftrightarrow\quad \delta(gx^{-1})=\sum_j\overline{\chi_j(g)}\chi_j(x) \tag{C.539}$$

$$\delta(x+a-a) \quad\Longleftrightarrow\quad \delta(xgx^{-1})=\delta(g) \tag{C.540}$$

$$\int dx\overline{e_m(x)}e_n(x)=\delta_{mn} \quad\Longleftrightarrow\quad \int D^j_{mn}(g)\overline{D^{j'}_{m'n'}(g)}dg=\delta_{jj'}\delta_{mm'}\delta_{nn'} \tag{C.541}$$

$$\int dy\overline{e_m(y)}e_n(x-y)=\delta_{mn}e_n(x) \quad\Longleftrightarrow\quad \int_G\overline{\chi_{j'}(g)}\chi_j(gx)dg=V_G\delta_{jj'}\frac{\chi_j(x)}{d_j} \tag{C.542}$$

Exercise

(a):

Recall how we prove $\int dx\delta(x)f(x)=f(0)$ implies $\int dy\delta(y-a)h(y)=h(a)$ where $h(y)=f(y-a)$

make the substitution $x=y-a$

$$h(a)=f(0)=\int dy\delta(y-a)f(y-a)=\int dy\delta(y-a)h(y) \tag{C.543}$$

prove

$$\int_G dU\delta(U)f(U)=f(I) \quad\Rightarrow\int_G dU\delta(U\gamma^{-1})f(U)=f(\gamma) \tag{C.544}$$

define $h(U\gamma^{-1})=g(U)$

$$h(\gamma)=f(1) \tag{C.545}$$

(b):

$$f(g)=\sum_\Lambda\phi^{\alpha\beta}_\Lambda D^{(\Lambda)}_{\alpha\beta}(g) \tag{C.546}$$

$$f(g_1,g_2)=\sum_{\Lambda_1\Lambda_2}\phi^{\alpha_1\alpha_2\beta_1\beta_2}_{\Lambda_1\Lambda_2}D^{(\Lambda_1)}_{\alpha_1\beta_1}(g_1)D^{(\Lambda_2)}_{\alpha_2\beta_2}(g_2) \tag{C.547}$$

$$f(g_1, g_2 = C) = \sum_{\Lambda_1} a_{\Lambda_1}^{\alpha_1\beta_1}(g_2 = C) D_{\alpha_1\beta_1}(g_1) \tag{C.548}$$

$$a_{\Lambda_1}^{\alpha_1\beta_1}(g_2) = \sum_{\Lambda_2}^{\alpha_2\beta_2} D_{\alpha_2\beta_2}(g_2) \tag{C.549}$$

$$a_{\Lambda_1}^{\alpha_1\beta_1} =: \phi_{\Lambda_1\Lambda_2}^{\alpha_1\alpha_2\beta_1\beta_2} \tag{C.550}$$

Using $D^j_{ml}(g)D^j_{ln}(x) = D^j_{mn}(gx)$

$$\int \overline{D^{j'}_{m'n'}(g)} D^j_{mn}(gx) dg = \delta_{jj'}\delta_{mm'} D^j_{nn'}(x) \tag{C.551}$$

contracting m, m' and nn' we egt (C.542).

Riesz-Fishcher theorem

Parseval's equation :

$$\sum_{-\infty}^{\infty} |c_n|^2 = \int |f(x)|^2 dx. \tag{C.552}$$

$$f_n(x) = \sum_{k=-n}^{n} c_k e_k(x) \tag{C.553}$$

converge to the vector f in the sense of L_2:

$$||f - f_n|| \to 0. \tag{C.554}$$

limit in the mean of the f_n's.

$$||f_m - f_n||^2 = \sum_{|k|=n+1}^{m} |c_k|^2. \tag{C.555}$$

if c_n are given complex numbers for which $\sum_{-\infty}^{\infty} |c_n|^2$ converges, then there exists a function f in L_2. If we grant the completeness of L_2 as a metric space this is easy to prove.

This tells us that the f's form a Cauchy sequence in L_2; and since L_2 is complete, there exists a function f in L_2 such that $f_n \to f$.

It is apparent that

C.7.13 Clebsch-Gordan

$$D_{(\mu)} \times D_{(\nu)} = \sum_{\otimes\sigma} n_{\mu\nu}^{(\Lambda)} D_{(\Lambda)} \tag{C.556}$$

where $n_{\mu\nu}^{(\Lambda)}$ is the number of times that the Λ-th irreducible representation occurs in the product representation (**??**).

$$\chi_{(\mu\times\nu)}(g) = \chi_{(\mu)}(g)\chi_{(\nu)}(g). \tag{C.557}$$

Thus by

$$n_{\mu\nu}^{(\Lambda)} = \int_G dg\overline{\chi_{(\Lambda)}(g)}\chi_{(\mu)}(g)\chi_{(\nu)}(g) \tag{C.558}$$

C.7.14 Semi-direct Products

A full Lorentz transformation can be decomposed into an ordinary spacial rotaion, followed by a boost, followed by a further ordinary rotation.

Definition (First definition)

Suppose N is a normal subgroup of G and H is another subgroup of G such that $N \cap H = E$ (the identity of G) and every element of G can be written in a unique way as

$$g = ba, \quad b \in H,\ a \in N,$$

then G is said to be a **semi-direct product** of H and N, written $G = H \otimes_S N$.

Definition (Second definition)

We form a new group whose elements are the elements of $H \times N$ and multiplication given by

$$(h_1, n_1) \cdot (h_2, n_2) = (h_1 h_2\ ,\ n_1 \rho_{h_1}(n_2)) \quad \text{with} \quad \rho_{h_1}(n_2) = h_1 n_2 h_1^{-1}. \tag{C.559}$$

Note $h_1 n_2 h_1^{-1} \in N$ as N is a normal subgroup of G.

Check it forms a group

The identity element of this group is (E, E):

$$(h,n)\cdot(E,E) = (h, h\; hEh^{-1}) = (h,n),$$
$$(E,E)\cdot(h,n) = (h, E\; EnE^{-1}) = (h,n).$$

The inverse of (h,n) is (h^{-1},n') where $n' = h^{-1}n^{-1}h$:

$$(h,n)\cdot(h^{-1},n') = (E, n\; hn'h^{-1}) = (E,E),$$
$$(h^{-1},n')\cdot(h,n) = (E, n'\; h^{-1}nh) = (E,E).$$

Associativity $[(h_1,n_1)\cdot(h_2,n_2)]\cdot(h_3,n_3) = (h_1,n_1)\cdot[(h_2,n_2)\cdot(h_3,n_3)]$:

$$\begin{aligned}
[(h_1,n_1)\cdot(h_2,n_2)]\cdot(h_3,n_3) &= (h_1h_2, n_1\; h_1n_2h_1^{-1})\cdot(h_3,n_3) \\
&= (h_1h_2h_3, n_1\; h_1n_2h_1^{-1}\; h_1h_2n_3(h_1h_2)^{-1})
\end{aligned}$$

$$\begin{aligned}
(h_1,n_1)\cdot[(h_2,n_2)\cdot(h_3,n_3)] &= (h_1,n_1)\cdot(h_2h_3, n_2\; h_2n_3h_2^{-1}) \\
&= (h_1h_2h_3, n_1\; h_1(n_2\; h_2n_3h_2^{-1})h_1^{-1})
\end{aligned}$$

□

Equivalence of the two defintions

for example $H \cap N$ might be empty.

for each $h \in H$ the inner automorphism $x \to hxh^{-1}$ takes N to N and defines an automorphism

$$\rho_h(n) = hnh^{-1}$$

Moreover,

$$\begin{aligned}
\rho_{h_1}(\rho_{h_2}(n)) &= \rho_{h_1}(h_2nh_2^{-1}) = h_1h_2nh_2^{-1}h_1^{-1} \\
&= \rho_{h_1\cdot h_2}(n)
\end{aligned}$$

Thus $h \to \rho_h$ is a homomorphism of H into the group of automorphisms of N, we write $\rho \in Hom(H, Aut(N))$.

Sub-groups

The group G is the semigroup product of N by H with homomorphism ρ.

Recall $\rho(h)(n) = hnh^{-1}$. Note that if $\rho \in Hom(H, Aut(N))$

Obviously there is the subgroup of G composed of elements of the form:

$$(h_i, E) \quad \text{for all } h_i \in H.$$

$$H \otimes_S I \simeq H$$

there is the subgroup of G composed of elements of the form:

$$(E, n_i) \quad \text{for all } n_i \in N.$$

as

$$(E, n_1) \cdot (E, n_2) = (E, E, n_1 \; Eh_2E) = (E, n_1 n_2)$$

$$I \otimes_S N \simeq N$$

Conversely, suppose that we are given two groups N and H and a homomorphism $h \to \rho_h$ of H into the group of all automorphisms of N. We may then define a group $H \otimes_S N$ with respect to ρ as follows.

□

Example The group RT is a semi-direct product of the rotation group $R(2)$ and the group of all trnaslations $T(2)$

Proof:

Example The group ?? is a semi-direct product of the rotation group $R(3)$ and the group of all trnaslations $T(3)$

Proof:

The group of all translations and rotations has six generators $\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{J}_1, \hat{J}_2, \hat{J}_3$. More concisely, it is called the transaltion-rotaion group and has the translations $(\hat{p}_\nu)$ as an abelian subgroup. It is obvious that

$$\hat{R}\hat{T}\hat{R}^{-1} = \hat{T}',$$

where $\hat{R}$ is a rotation and $\hat{T}$ a translation, is again a pure translation $\hat{T}'$ (see fig (5.24)). Consequently the translation group is an invariant abelian subgroup of the translation-rotation group.

The group consists of pairs of the pairs (r, t) with $t \in T$, $r \in O(3, \mathbb{R})$ and multiplication rule

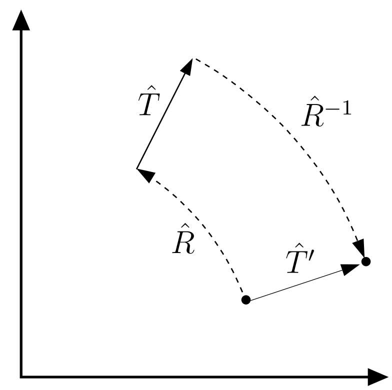


Figure C.15: TranRotGrF.

$$(r_1, t_1) \cdot (r_2, t_2) = (r_1 r_2, t_1 r_1 t_2 r_1^{-1}). \tag{C.560}$$

Example The Poincare group P includes the abelian subgroup $T(4)$ of all translations in Minkowski spacetime in addition to the Lorentz transformations. The Poincare group P is a semi-direct product of the Lorentz group $L(4)$ and the group of all translations $T(4)$ on Minkowski spaectime, that is,

$$P = L(4) \otimes_S T(4). \tag{C.561}$$

Proof:

$T(4)$ is a normal abelian subgroup of P as for $t \in T(4)$ and any $p \in P$, $ptp^{-1} \in T(4)$.

An element $p \in P$ is then denoted $p = (\Lambda^\mu_\nu, a^\mu)$.

Examples from Quantum Gravity

(i) Three diemnsional gravity:

where $ISU(2)$ is the (universal cover of the) group of Euclidean transformations. It is the semigroup:

$$ISU(2) = SU(2) \otimes_S \mathbb{R}^3 \tag{C.562}$$

Its elements are written as $(u, \vec{a})$ where u is an element of $SU(2)$ and $\vec{a}$ is a vector in $\mathbb{R}^3$. We have

$$(u_1, \vec{a}_1) \cdot (u_2, \vec{a}_2) = (u_1 u_2, u_1 \vec{a}_2 + \vec{a}_1), \tag{C.563}$$

the notation $u\vec{a}$ means $U(u)\vec{a}$ where U is the vectorial representation of $SU(2)$.

(ii) black hole gauge group is the semi-direct product of ....

(iii) Gauge group of LQG

The group of these symmetries is the semi-direct product the group of smooth local gauge transformations with the group of smooth diffeomorphisms on Σ.

The Irreducible Representations of Semi-direct Products

C.8 Infinite-Dimensional Group Representations

C.8.1 Group Actions

If G is a permutation group, then every element of the group permutates elements of the set $\{1, 2, \cdots n\}$. We say that the group G acts on the set $\{1, 2, \cdots n\}$.

Group action on a set Let G be a group and let A be a some set. Suppose that we have a map $T : G \times A \to A$ such that for every fixed $g \in G$ the map $a \to T(g, a)$ is a permutation of the set A.

We require the compatibility conditions:

(i) $T(g_1, T(g_2, a)) = T(g_1 g_2, a)$ for all $g_1, g_2 \in G$ and $a \in A$.

(ii) $T(e, a) = a$ for all $a \in A$

where $e = e_G$ the identity of G.

define what is called T is a **group action** and that G acts on A by the action of T.

We can shorten the notion by writing $g \cdot a$ instead of $T(g, a)$.

These conditions are designed so that the map $G \to S_A$ is a group homomorphism.

C.8.2 Countable and Locally Compact Topological Groups

Let S be a set on which some group acts as a group of transformations. Let us write sx for the transfrom by x of $s \in S$. Then

$$(sx)y = s(xy) \qquad \text{for all } s \in S, x, y \in G.$$

and

$$se = s$$

where e is the identity element of G.

We assume that G is a topological group which is separable and locally compact and S comes with a 'volume element'. We assume that S is a Borel space with a measure μ; that is, that we are given a family $\mathcal{B}$ of subsets of S, closed with respect to completions and countable uniions, and measure μ assigning a nonnegative real number or ∞ to each subset $E \in \mathcal{B}$ so that

$$\mu(E_1 + E_2 + \cdots) = \mu(E_1) + \mu(E_2) + \cdots$$

whenever $E_i \cap E_j = 0$ for all $i \neq j$. The members of $\mathcal{B}$ are called *Borel sets*. We assume S is a union of countably many sets of finite measure and that

If S_1 and S_2 are Borel spaces, then a function from S_1 to S_2 is a Borel function if $g^{-1}(E)$ is a Borel set in S_1 whenever E is a Borel set in S_2. Any topological space becomes a Borel space if we define the Borel sets to be the sets one can obtain from closed sets by the taking of complements, countable unions, and countaale intersections. We extend theorems already proven to much more general context as the notion of Borel space and Borel function are quite general. A function can be widly discontinuous and still be a Borel function.

We shall say that the measure μ is *invariant* if

$$\mu(Ex) = \mu(E) \qquad \text{for all Borel sets } E \text{ and all } x \in G. \tag{C.564}$$

C.8.3 Haar Measure

Given an $n-$dimensional manifold and a nowhere-vanishing orieneted $n-$form η, we can make a measure on $\mathcal{M}$ by defining the integral of f against μ to be the integral of the $n-$form $f\eta$.

It is not hard to show that on an $n-$dimensional Lie group G, there exists a nowhere-vanishing $n-$form that is invariant under left translations and that this form is unique up to a constant. Integrating fucntions against this form gives a letf-invariant measure (i.e., a left Haar measure).

C.8.4 Summary of Group theory

- Unitrary representations.
- Irreducible representations.
- Schur's lemmas.
- Orthoganality theorems.

C.9 Manifolds and Elementary Topology

Roughly, a manifolds are sets on which, at least around each point, everything looks Euclidean.

Organized set of points with a structure - a division into convenient family of subset families.

It is this need for care, to ensure we can rely on calculations we do, that motivates much of this course, illustrates why we empathize accurate argument as well as getting the "correct" answers, and explains why in the rest of this section we need to revise elementary notions.

C.9.1 Sets and Mappings Between Sets

We need to be able to talk easily about certain subsets of R. We say that I is an **open interval** if

$$I = (a, b) = x \in R : a < x < b. \tag{C.565}$$

Thus an open interval excludes its end points, but contains all the points in between. x is always separated from F by

In contrast a **closed interval** contains both its end points, and is of the form

$$I = [a, b] = \{x \in R : a \leq x \leq b\}. \tag{C.566}$$

It is also sometimes useful to have **half-open intervals** like $(a, b]$ and $[a, b)$. It is trivial that $[a, b] = (a, b) \cup_{\{a\}} \cup_{\{b\}}$.

The two end points a and b are points in R. It is sometimes convenient to allow also the possibility $a = -\infty$ and $b = +\infty$; it should be clear from the context whether this is being allowed. If these extensions are being excluded, the interval is sometimes called a *finite interval*, just for emphasis.

Of course we can easily get to more general subsets of R. So $(1, 2) \cup [2, 3] = (1, 3]$ shows that the union of two intervals may be an interval, while the example $(1, 2) \cup (3, 4)$ shows that the union of two intervals *need not* be an interval.

An open subset V of X is a subset where if $x \in V$, then there is a $\delta > 0$ such that an open ball $B_\delta(x)$ is entirely contained in V.

Notation	Meaning
$a \in A$	a belongs to A
$A \subset B$	A is included in B
$A = B$	A is identical to B
$A \cup B$	The union of A and B ($\cup$ is for **u**inion)
$A \cap B$	The intersection A and B ($\cap$ is for **in**tersection)
$A - B$	The set of elements of A not included in B

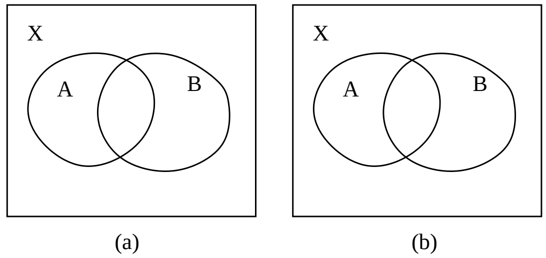


Figure C.16: .

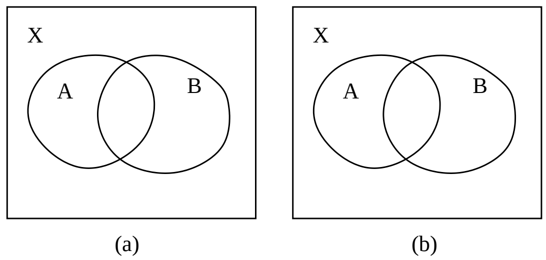


Figure C.17: .

C.9.2 Continuity

A function f is continuous at a point p if whenever we can force the distance between $f(x)$ and $f(p)$ to be as small as desired by taking the distance between x and p to be small enough.

The definition of continuity of a function f on the real line: $f : R \mapsto R$ is continuous at x_0 if for any positive number ϵ, there exists a positive number δ such that if $|y - x_0| < \delta$, then $|f(y) - f(x_0)| < \epsilon$. The setting for this definition is the real line, a Euclidean space.

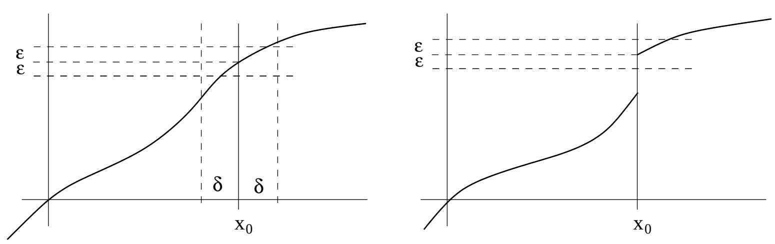


Figure C.18: (a) (b) f is discontinuous at p_1.

We can generalize this definition a little by considering mapping between spaces with a metric. Let X, Y be two spaces with metrics d_1 and d_2, respectively. Then, $f : X \to Y$ is continuous at $x_0 \in X$ if for any $\epsilon > 0$, there exists $\delta > 0$, such that if $d_1(x, x_0) < \delta$, then $d_2(f(x), f(x_0)) < \epsilon$.

However, the notion of continuity does not depend on a metric.

We need to be able to talk about a function near a point a. If we want to look at the points a distance less than d for a. we are looking at an interval $(a-d, a+d)$. We call such an interval a neighbourhood of a.

Definition A subset $\mathcal{U}$ is **open** if given $a \in \mathcal{U}$, there is some $\delta > 0$ such that $(a-\delta, a+\delta) \subseteq \mathcal{U}$.

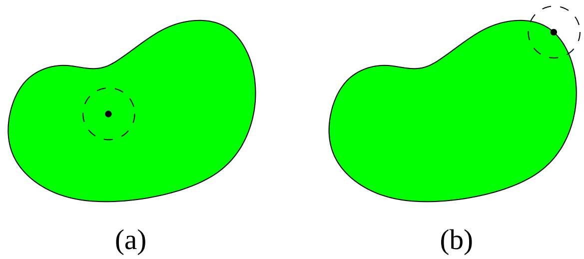


Figure C.19: Open sets interior points (b) .

In fact this is the same as saying that given $a \in \mathcal{U}$, there is some open interval containing a which lies in $\mathcal{U}$ - a set is open if it contains a neighbourhood of each of its points. This definition has the effect that if a function is defined on an open set that its behaviour near point a of interest from both sides.

We can transfer such things as limits and calculus from Euclidean space. A ***topology*** is a structure added to an arbitrary point set which enables one to define a convergent sequence and to define a continuous function in a general setting.

A ***topology***, $\mathcal{T}$, on a set X is defined to be a specified family of open subsets on X satisfying the following 3 properties:
(i) The empty set, $\emptyset$, and the space X belong to $\mathcal{T}$.
(ii) The union of any number (possibly infinite) of open subsets belonging to $\mathcal{T}$ is also in $\mathcal{T}$.
(iii) The intersection of any finite number (nor infinite) open subsets in $\mathcal{T}$ also belongs to $\mathcal{T}$.

The set X together with a topology $\mathcal{T}$ is called a topological space.

Let $f : x \to Y$ be a function between two topological spaces $(X, \mathcal{T}_X)$, and $(Y, \mathcal{T}_Y)$. If $x \in X$, then $f(x) \in Y$ is the image of x under f. Let U be an open set in X (i.e. $U \in \mathcal{T}_X$). The image of U under f is the subset $V = f(U) \subset Y$, the range of f with domain U. If V is a subset of Y, then the inverse image of V by f. The mapping f is defined to be *continuous* when the inverse image of any open set is open. That is, f is continuous if $U = f^{-1}(V) \in \mathcal{T}_X$ when $V \in \mathcal{T}_Y$.

The transition functions transform the coordinates of one overlapping patch into another.

The usual Jacobian of ψ_{ij} is

$$\begin{aligned}(\Lambda^a{}_b) &= \frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)} \\ &= \begin{pmatrix} \frac{\partial x'^1}{\partial x^1} & \ldots & \frac{\partial x'^1}{\partial x^n} \\ \ldots & \ldots & \\ \frac{\partial x'^n}{\partial x^1} & \ldots & \frac{\partial x'^n}{\partial x^n} \end{pmatrix}\end{aligned} \tag{C.567}$$

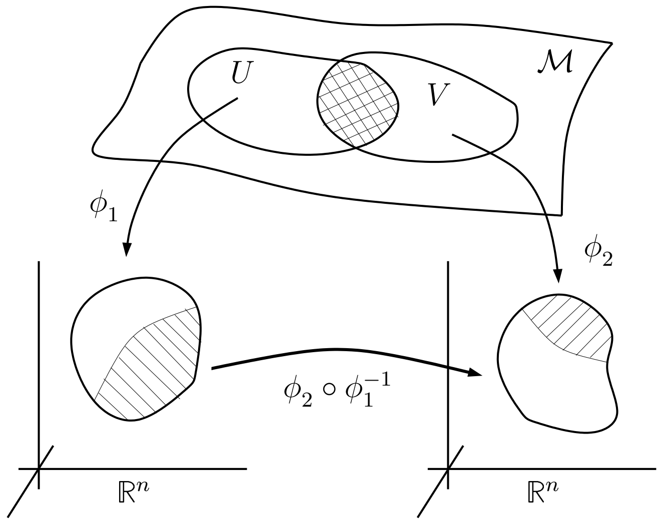


Figure C.20: (U, ϕ_1) and (V, ϕ_2) are two coordinate patches on X. Transition functions, $\phi_2 \circ \phi_2^{-1}$, are ordinary functions that go from points of one R^n space onto another, i.e. $\phi_2 \circ \phi_2^{-1} : R^n \mapsto R^n$. The domain and range of the transition function are the shaded regions in R^n.

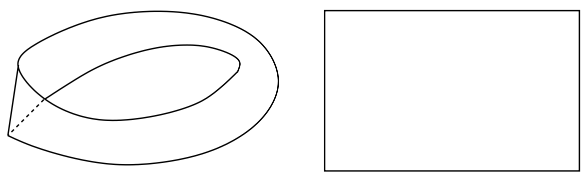


Figure C.21: Möbius.

C.10 Elementary Tensor Analysis

One often needs to do a change of coordinates, either because we prefer to use a different choice of coordinates valid in some patch, or because we need to transform to new patch which covers a different portion of the manifold.

Einstein summation convention

$$\sum_a (..)^a \ (..)_a \to (..)^a \ (..)_a \tag{C.568}$$

Contravariant Vectors

$$dz = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \tag{C.569}$$

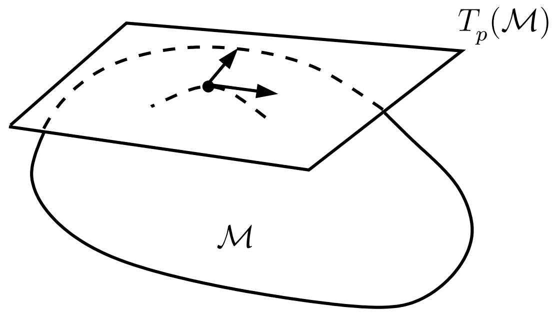


Figure C.22: tangtoM.

the infinitesimal displacement in two different coordinate systems x^a and x'^a is related by

$$dx'^a = \sum_b \frac{\partial x^{a\prime}}{\partial x^b} dx^b \tag{C.570}$$

where $x^{a\prime}/x^b$ is evaluated at the point $p = x^a = x'^a(x^a)$. **Einstein summation convention**

$$dx^{a\prime} = \frac{\partial x^{a\prime}}{\partial x^b} dx^b \tag{C.571}$$

We set

$$\Lambda^a_b = \frac{\partial x^{a\prime}}{\partial x^b} \tag{C.572}$$

The infinitesimal displacement is the prototype of a geometric object which is called a contravariant vector. A set of quantities X_a are said to be the components of a contravariant vector if they transform, under a change of coordinates, as

$$X'^a = \frac{\partial x^{a\prime}}{\partial x^b} X^b \tag{C.573}$$

Covariant Vectors

$$df = \frac{\partial f}{\partial x^a} dx^a = 0 \tag{C.574}$$

$$\frac{\partial f'}{\partial x^{a\prime}} = \frac{\partial f}{\partial x^b} \frac{\partial x^b}{\partial x^{a\prime}} \tag{C.575}$$

$f = f(x^a) = f(x^{a\prime}) = f'$.

A set of quantities X_a are said to be the components of a covariant vector if they transform as

$$X'_a = \frac{\partial x^b}{\partial x'^a} X_b \tag{C.576}$$

Covariant vectors can be interpreted as linear functionals mapping vectors to $\mathbb{R}$.

$$\mathbf{V} = V^a \mathbf{e}_a = V'^a \mathbf{e}'_a \tag{C.577}$$

It is evident that if the vector $\mathbf{V}$ is to be invariant the basis vectors $\{\mathbf{e}_a\}$ must transform under coordinate transformations as the components of a covector. Similarly, basis covectors $\{\mathbf{e}^a\}$ must change under a coordinate transform as the components of a contravariant vector.

Tensors

A set of quantities T_{ab} are said to be the components of a covariant tensor of second?? if they transform as

$$T'_{ab} = \frac{\partial x^c}{\partial x'^b} \frac{\partial x^d}{\partial x'^a} T_{cd} == \tag{C.578}$$

Contractions of Tensors

$$\frac{\partial x^{a\prime}}{\partial x^c} \frac{\partial x^c}{\partial x^{b\prime}} = \delta^b_a \tag{C.579}$$

C.10.1 Affine Connection

$$\begin{aligned} dX^a &= X^a(Q) - X^a(P) \\ &= X^a(x^b + dx^b) - X^a(x^b) \\ &= \frac{\partial X^a}{\partial x^b} dx^b \end{aligned}$$

is not tensorial. Under a coordinate transformation a partial derivative of a vector X^a transforms as

$$\frac{\partial X'^a}{\partial x'^b} = \frac{\partial x^c}{\partial x'^b} \frac{\partial}{\partial x^c} \left(\frac{\partial x'^a}{\partial x^d} X^d \right) = \underbrace{\frac{\partial x^c}{\partial x'^b} \frac{\partial x'^a}{\partial x^d} \frac{\partial X^d}{\partial x^c}}_{tensorial} + \underbrace{\frac{\partial^2 x'^a}{\partial x^c \partial x^d} \frac{\partial x^c}{\partial x'^b} X^d}_{inhomogeneous} \tag{C.580}$$

This is not a tensorial transformation because of the inhomogeneous term.

have

$$dx^b \nabla_b X^a = \frac{\partial X}{\partial x^b} dx^b + \Gamma^a_{cb} X^c dx^b \tag{C.581}$$

For $\nabla_b X^a$ to transform as a tensor of type $(1,1)$ then the connection must transform as

$$\Gamma'^{\alpha}_{\beta\gamma} = \frac{\partial x'^{\alpha}}{\partial x^{\delta}} \frac{\partial x^{\sigma}}{\partial x'^{\beta'}} \frac{\partial x^{\tau}}{\partial x'^{\gamma}} \Gamma^{\delta}_{\sigma\tau} + \frac{\partial x'^{\alpha}}{\partial x^{\delta}} \frac{\partial^2 x^{\delta}}{\partial x'^{\beta} \partial x'^{\gamma}} \tag{C.582}$$

Any quantity that transforms as (C.582) is called an affine connection. If the second term on the right-hand side were absent, then this would be the transformation law of a tensor of type (1,2).

$$T^{\alpha}_{\beta\gamma} = \Gamma'^{\alpha}_{\beta\gamma} - \Gamma'^{\alpha}_{\gamma\beta} \tag{C.583}$$

is a tensor called the *torsion tensor*. If the torsion tensor vanishes then the connection is said to be *symmetric* or (*torsion-free*), i.e.

$$\Gamma'^{\alpha}_{\beta\gamma} = \Gamma'^{\alpha}_{\gamma\beta} \tag{C.584}$$

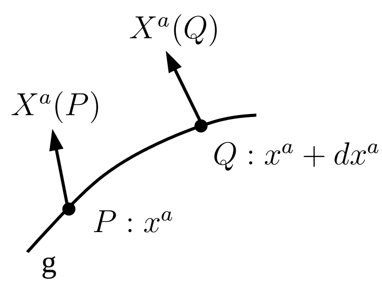


Figure C.23: connection.

Covariant differentiation can be extended to other types of tensors by demanding that the covariant derivative obeys the product rule of differential calculus. The covariant derivative of a scalar field is the same as its partial derivative,

$$\nabla \phi = \partial_a \phi. \tag{C.585}$$

Demanding that the covariant derivative obeys the Leibniz rule, then we find

$$\nabla_b X_a = \partial_b X_a - \Gamma^c_{ab} X_c \tag{C.586}$$

$$\nabla_c T^{a\dots}_{b\dots} = \partial_c T^{a\dots}_{b\dots} + \Gamma^a_{dc} T^{d\dots}_{b\dots} + \dots - \Gamma^d_{bc} T^{a\dots}_{d\dots} - \dots. \tag{C.587}$$

The covariant derivative contracted with X.

$$\nabla_X T^{a\dots}_{b\dots} := X^c \nabla_c T^{a\dots}_{b\dots} \tag{C.588}$$

$$\frac{D}{Du} T^{a\dots}_{b\dots} = \nabla_X T^{a\dots}_{b\dots} \tag{C.589}$$

C.10.2 Affine Geodesic

A geodesic is the closest thing there is to a straight line curved space time. In we are given a metric we can define a geodesic as the shortest distance between two points. However, there is a more general definition of a geodesic is its velocity vector is parallel transported along the curve it traces out in spacetime (it follows its own nose, so to speak). In other words, the parallely propagated vector at any point of the curve is parallel, that is, proportional to the vector at this point:

$$\frac{dx^b}{du} \nabla_b \left(\frac{dx^a}{du} \right) = \lambda(u) \frac{dx^a}{du} \tag{C.590}$$

The L.H.S. becomes

$$\begin{aligned} \frac{dx^b}{du} \nabla_b \left(\frac{dx^a}{du} \right) &= \frac{dx^b}{du} \frac{\partial}{\partial x^b} \left(\frac{dx^a}{du} \right) + \Gamma^a_{bc} \frac{dx^b}{du} \frac{dx^a}{du} \\ &= \frac{d^2 x^a}{du^2} + \Gamma^a_{bc} \frac{dx^b}{du} \frac{dx^c}{du}. \end{aligned}$$

We have the geodesic equation

$$\frac{d^2 x^a}{du^2} + \Gamma^a_{bc} \frac{dx^b}{du} \frac{dx^c}{du} = \lambda(u) \frac{dx^a}{du}. \tag{C.591}$$

Note that this definition did not involve a metric, it only refers to a connection. Other geometric notions can also be defined using a connection only. Such description are desirable when we come to formulate the quantum theory where there is background metric. In fact LQG is formulated with a connection of a frame field (which we will be coming to presently). In the next section we shall verify that the differential equation derived from the principle of shortest distance is as in (C.591) but with the connection Γ^a_{bc} as a particular function of the metric. This connection is called the metric connection.

C.10.3 The Metric Connection

The vanishing of the covariant derivative of the metric is equivalent to requiring that the length of a vector is unchanged under parallel propagation. To see this first consider the covariant derivative of the length squared l^2 of a arbitrary vector l^μ,

$$\begin{aligned} 0 = \partial_\sigma(l^2) = \nabla_\sigma(l^2) &= \nabla_\sigma(g_{\mu\nu}l^\mu l^\nu) \\ &= l^\mu l^\nu \nabla_\sigma g_{\mu\nu} + g_{\mu\nu} l^\nu \nabla_\sigma l^\mu + g_{\mu\nu} l^\mu \nabla_\sigma l^\nu \\ &= l^\mu l^\nu \nabla_\sigma g_{\mu\nu} + 2 l_\mu \nabla_\sigma l^\mu \end{aligned} \tag{C.592}$$

By assumption the vector vector l^μ is parallel propagated i.e. $\nabla_\sigma l^\mu = 0$

$$l^\mu l^\nu \nabla_\sigma g_{\mu\nu} = 0 \tag{C.593}$$

since l^μ was arbitrary this implies,

$$\nabla_\sigma g_{\mu\nu} = 0 \tag{C.594}$$

(It should be noted that the vanishing of the covariant derivative was motivated by a physical requirement and not a mathematical one). This condition is sufficient to determine the connection as a function of the metric. We show this by using

$$\nabla_\sigma g_{\mu\nu} = \partial_\sigma g_{\mu\nu} + \Gamma^\rho_{\sigma\mu} g_{\nu\rho} + \Gamma^\rho_{\sigma\nu} g_{\mu\rho} = 0 \tag{C.595}$$

$\Gamma_{\mu\nu,\sigma} = g_{\sigma\lambda}\Gamma^\lambda_{\mu\nu}$. Written out explicitly, with cyclic rotating of indices,

$$\begin{aligned} \partial_a g_{bc} + \Gamma_{ab,c} + \Gamma_{ac,b} = 0 &\quad \leftrightarrow \quad (\nabla_\sigma g_{ab} = 0) \\ \partial_\nu g_{\lambda\mu} + \Gamma_{\nu\lambda,\mu} + \Gamma_{\lambda\nu,\mu} = 0 &\quad \leftrightarrow \quad (\nabla_\nu g_{\lambda\mu} = 0) \\ \partial_\lambda g_{\mu\nu} + \Gamma_{\lambda\mu,\nu} + \Gamma_{\nu\mu,\lambda} = 0 &\quad \leftrightarrow \quad (\nabla_\lambda g_{\mu\nu} = 0) \end{aligned} \tag{C.596}$$

These three equations are identical. But by adding the first two equations and subtracting the last (and remembering that the Christoffel symbol is symmetric in the lower case indices), we find:

$$\Gamma^c_{ab} = \frac{1}{2} g^{cd}\{\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab}\} \tag{C.597}$$

Covariant Derivative

$$\nabla_a V_b = \partial_a V_b + \Gamma^c_{ab} V_c \tag{C.598}$$

Metric Geodesic

$$\left(\frac{ds}{du}\right)^2 = g_{ab}\frac{dx^a}{du}\frac{dx^b}{du}. \tag{C.599}$$

$$s = \int_{p_1}^{p_2} ds = \int_{p_1}^{p_2} \frac{ds}{du}du = \int_{p_1}^{p_2} \left(g_{ab}\frac{dx^a}{du}\frac{dx^b}{du}\right)^{1/2} du. \tag{C.600}$$

$$\frac{d^2x^a}{du^2} + \Gamma^a_{bc}\frac{dx^b}{du}\frac{dx^c}{du} = \left(\frac{d^2s}{du^2}\Big/\frac{ds}{du}\right)\frac{dx^a}{du}, \tag{C.601}$$

$$u = \alpha s + \beta, \tag{C.602}$$

where α and β are constants, then the righthand side vanishes.

we assume $ds \neq 0$

equations for a metric geodesic

$$\frac{d^2x^a}{ds^2} + \Gamma^a_{bc}\frac{dx^b}{ds}\frac{dx^c}{ds} = 0 \tag{C.603}$$

and

$$g_{ab}\frac{dx^a}{ds}\frac{dx^b}{ds} = -1, \tag{C.604}$$

where Γ^a_{bc} is given by (C.597).

Proof.

$$\frac{d}{du}\frac{\partial \mathcal{L}}{\partial \dot{x}^a} - \frac{\partial \mathcal{L}}{\partial x^a} = 0 \tag{C.605}$$

We instead minimize this instead. So we use $\mathcal{L} = g_{ab}\dot{x}^a\dot{x}^b$ in

$$2\mathcal{L}\left[\frac{d}{du}\frac{\partial \mathcal{L}}{\partial \dot{x}^a} - \frac{\partial \mathcal{L}}{\partial x^a}\right] = 0 \tag{C.606}$$

which can be rewritten as

$$\frac{d}{du}\left(\frac{\partial \mathcal{L}^2}{\partial \dot{x}^a}\right) - \frac{\partial \mathcal{L}^2}{\partial x^a} = 2\frac{\partial \mathcal{L}}{\partial \dot{x}^a}\frac{d\mathcal{L}}{du} \tag{C.607}$$

Substituting for $\mathcal{L}^2$, the left hand side gives

$$\begin{aligned}
\frac{d}{du}\left(\frac{\partial\mathcal{L}^2}{\partial\dot{x}^a}\right) - \frac{\partial\mathcal{L}^2}{\partial x^a} &= \frac{d}{du}\left[\frac{\partial}{\partial\dot{x}^a}(g_{bc}\dot{x}^b\dot{x}^c)\right] - \frac{\partial}{\partial x^a}(g_{bc}\dot{x}^b\dot{x}^c) \\
&= \frac{d}{du}(2g_{ab}\dot{x}^b) - (\partial_a g_{bc})\dot{x}^b\dot{x}^c \\
&= 2g_{ab}\ddot{x}^b + 2\partial_c g_{ab} - \partial_a g_{bc}\dot{x}^b\dot{x}^c \\
&= 2g_{ab}\ddot{x}^b + 2\dot{x}^b\dot{x}^c[\frac{1}{2}(\partial_c g_{ba} + \partial_b g_{ca} - \partial_a g_{bc})] \qquad \text{(C.608)}
\end{aligned}$$

$$\begin{aligned}
2\frac{\partial\mathcal{L}}{\partial\dot{x}^a}\frac{d\mathcal{L}}{du} &= 2\frac{\partial}{\partial\dot{x}^a}(g_{bc}\dot{x}^b\dot{x}^c)\frac{d}{du}\frac{ds}{du} \\
&= 2(g_{bc}\dot{x}^b\dot{x}^c)^{-1/2}g_{ad}\dot{x}^d\frac{d^2s}{du^2} \\
&= 2\left(\frac{d^2s}{du^2}\Big/\frac{ds}{du}\right)g_{ab}\dot{x}^b. \qquad \text{(C.609)}
\end{aligned}$$

multiplying through by g^{ad}

Lighlike inertial motion cannot be characterized with reference to proper time parameterization since the proper time along a null curve vanishes. However, this does not prevent us from characterizing such motion in the same manner as in the timelike case. To this end we go back to the more general definition of a geodesic as a curve $x^a(\lambda)$ with the property that the coordinate acceleration $d^2x^a/d\lambda^2$ at every point p is parallel to the tangent vector.

It is independent of the parameterization of the curve

C.10.4 Curvature

Let ξ^a be a contravariant vector field with

$$\dot{\xi}^c(x(\tau)) = 0. \qquad \text{(C.610)}$$

We take the curve to be small so that we can write

$$\xi^c(x) = \xi^c + \xi^c_{,b}x^b + \mathcal{O}(x^2). \qquad \text{(C.611)}$$

In an affine space without metric the term 'small' and 'large' appear to be meaningless. However, since differentiability is required, the small size limit *is* well defined. Thus, it is more precise to state that the curve is *infinitesimally* small.

If there is a strong gravitational field the contravariant vector may not return to its original value going around the loop once and have deviation $\delta\xi^a$. We find:

$$\begin{aligned}
\oint d\tau\dot{\xi} &= 0 \\
\delta\xi^c &= \oint d\tau \frac{d}{d\tau}\xi^c(x(\tau)) = -\oint \Gamma^c_{ab}\frac{dx^b}{d\tau}\xi^a(x(\tau))d\tau \\
&= -\oint d\tau\big(\Gamma^c_{ab} + \Gamma^c_{ab,d}x^d\big)\frac{dx^b}{d\tau}\big(\xi^a + \xi^a_{,d}x^d\big). \qquad \text{(C.612)}
\end{aligned}$$

where we chose the function $x(\tau)$ to be v. small, so that terms $\mathcal{O}(x^2)$ can be neglected. We have for a closed curve,

$$\oint d\tau\frac{dx^b}{d\tau} = 0 \quad \text{and} \quad \nabla_a\xi^c \approx 0 \to \xi^c_{,a} \approx -\Gamma^c_{ab}\xi^b, \qquad \text{(C.613)}$$

so that (5.24) becomes

$$\delta\xi^c = \frac{1}{2}\left(\oint x^d\frac{dx^b}{d\tau}d\tau\right)R^c_{abd}\xi^a + \mathcal{O}(x^2). \qquad \text{(C.614)}$$

covariant derivative of $T_{\nu\gamma}$ is given by

$$\nabla_\mu T_{\nu\gamma} = \partial_\mu T_{\nu\gamma} + \Gamma^\delta_{\mu\nu}T_{\delta\Gamma} + \Gamma^\delta_{\mu\gamma}T_{\nu\delta} \qquad \text{(C.615)}$$

$$\begin{aligned}
\nabla_c(\nabla_d V^a) &= \partial_c(\nabla_d V^a) + \Gamma^e_{cd}(\nabla_e V^a) - \Gamma^a_{ce}(\nabla_b V^e) \qquad \text{(C.616)} \\
&= \partial_c(\partial_d V^a + \Gamma^a_{df}V^f) + \Gamma^e_{cd}(\partial_e V^a + \Gamma^c_{df}V^f) - \Gamma^a_{ce}(\partial_b V^d + \Gamma^d_{bf}V^f)
\end{aligned}$$

$$\begin{aligned}
\nabla_c\nabla_d V^a - \nabla_d\nabla_c V^a &= \partial_c\partial_d V^a - \partial_d\partial_c V^a \ \text{(we take} = 0) \\
&+ \partial_c\Gamma^a_{df}V^f - \partial_d\Gamma^a_{cf}V^f \\
&+ \Gamma^e_{cd}\partial_e V^a - \Gamma^e_{dc}\partial_e V^a \quad (= 0 \ \text{ as } \Gamma^e_{cd} = \Gamma^e_{dc})
\end{aligned} \qquad \text{(C.617)}$$

$$\nabla_c\nabla_d V_a - \nabla_d\nabla_c V^a = \qquad \text{(C.618)}$$

where $R^\alpha_{\beta\gamma\delta}$

$$R^{\alpha}_{\beta\gamma\delta} \tag{C.619}$$

$$R^{\rho}_{\mu\nu\sigma} = \partial_{[\alpha}\Gamma_{\alpha} \tag{C.620}$$

C.10.5 Gaussian Normal Coordinates

$$x'^{a} = x^{a} + Q^{a}_{bc}\frac{x^{b}x^{c}}{2} \text{ where } Q^{a}_{bc} = Q^{a}_{cb} \tag{C.621}$$

$$\frac{\partial x'^{a}}{\partial x^{d}} = \delta^{a}_{d} + Q^{a}_{bd}x^{b} \tag{C.622}$$

$$\frac{\partial^{2} x'^{a}}{\partial x^{d}\partial x^{e}} = Q^{a}_{de} \tag{C.623}$$

$$x^{a} \stackrel{*}{=} 0 \tag{C.624}$$

$$\left[\frac{\partial x'^{a}}{\partial x^{b}}\right]_{P} = \delta^{a}_{b}, \quad \left[\frac{\partial x^{b}}{\partial x'^{a}}\right]_{P} = \delta^{a}_{b} \tag{C.625}$$

$$\Gamma'^{a}_{bc} = \frac{\partial x'^{a}}{\partial x^{d}}\frac{\partial x^{e}}{\partial x'^{b}}\frac{\partial x^{f}}{\partial x'^{c}}\Gamma^{d}_{ef} - \frac{\partial x^{d}}{\partial x'^{b}}\frac{\partial x^{e}}{\partial x'^{c}}\frac{\partial^{2} x'^{a}}{\partial x^{d}\partial x^{e}} \tag{C.626}$$

$$\left[\Gamma'^{a}_{bc}\right]_{P} = [\Gamma^{a}_{bc}]_{P} - Q^{a}_{bc} \tag{C.627}$$

choose $Q^{a}_{bc} = [\Gamma^{a}_{bc}]_{P}$

Mathematically this says that there exists a coordinate system the space time manifold is locally Minkowskian - one can always find a local coordinate system in which the metric tensor takes the pseudo-Euclidean form and the connection Γ^{a}_{bc} vanish at a point. In geometric terms, such a freely-falling frame of reference represents a local coordinate system.

Defined only in the immediate vicinity of a physically defined location (for example the place where two particles intersect), whose coordinate axes are very close to straight line and mutually perpendicular.

The physical statement is the gravitational field can be made to vanish at the place where two particles intersect, by going into free-fall.

C.10.6 Bianchi Identities

$$R^{\alpha}{}_{\beta\gamma\delta} = \Gamma^{\alpha}{}_{\beta\delta,\gamma} - \Gamma^{\alpha}{}_{\beta\gamma,\delta} \tag{C.628}$$

$$R_{\alpha\beta\gamma\delta} = -R_{\alpha\beta\delta\gamma} = R_{\beta\alpha\gamma\delta} = R_{\gamma\delta\alpha\beta} \tag{C.629}$$

Cyclic Bianchi identity

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\delta\beta\gamma} + R_{\alpha\gamma\delta\beta} \equiv 0 \tag{C.630}$$

$$\nabla_\alpha R_{\delta\sigma\beta\gamma} + \nabla_\gamma R_{\delta\sigma\alpha\beta} + \nabla_\beta R_{\delta\sigma\gamma\alpha} \equiv 0. \tag{C.631}$$

Define the tensor

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R, \tag{C.632}$$

the so-called Einsten tensor (its name comming from that it appears on the left-hand side of Einstein's field equations of general relativity). We have the contracted Bianchi identity

$$\nabla_\alpha G^{\alpha\beta} = 0. \tag{C.633}$$

which follows from the Bianchi identity,

$$\nabla_\sigma R^{\alpha}{}_{\beta\mu\nu} + \nabla_\mu R^{\alpha}{}_{\beta\nu\sigma} + \nabla_\nu R^{\alpha}{}_{\beta\sigma\mu} = 0$$

Contracting α and μ.

$$\nabla_\sigma R^{\mu}{}_{\beta\mu\nu} + \nabla_\mu R^{\mu}{}_{\beta\nu\sigma} + \nabla_\nu R^{\mu}{}_{\beta\sigma\mu} = 0$$

then contracting with $g^{\sigma\beta}$ gives

$$\nabla_\sigma R^{\mu\sigma}{}_{\mu\nu} + \nabla_\mu R^{\mu\sigma}{}_{\nu\sigma} - g^{\sigma\beta}\nabla_\nu R^{\mu}{}_{\beta\mu\sigma} = 0$$

or

$$\nabla_\mu (R^{\mu\nu} - \frac{1}{2} R g^{\mu\nu}) = 0.$$

C.10.7 Conformal Tensor, Ricci tensor and Ricci Scalar

C.10.8 The Weyl Tensor

The conformal tensor describes the components of the Riemann tensor, that are not contained in the Ricci tensor. The Ricci tensor being the contraction of the Riemann tensor, the rest of the information of the curvature is contained in the trace free part of the Riemann tensor, called the Weyl tensor,

$$C_{abcd} = R_{abcd} + \frac{1}{2}(g_{ad}R_{cb} + g_{bc}R_{da} - g_{ac}R_{bd} - g_{bd}R_{ca}) + \frac{1}{6}(g_{ac}g_{db} - g_{ad}g_{cb})R. \tag{C.634}$$

Constructed to have the same symmetries of the curvature tensor:

$$\begin{aligned} &C_{abcd} = -C_{abdc} = -C_{bacd} = C_{cdab}, \\ &C_{abcd} + C_{adbc} + C_{acdb} \equiv 0. \end{aligned} \tag{C.635}$$

$$C^a{}_{bad} \equiv 0 \tag{C.636}$$

C.10.9 Index Free Formulism

necessary to exploit coordinate systems. In this approach, tensor quantities are defined in terms of their components and the transformation rules for the latter under coordinate changes.

Physical quantities are (coordinate free) geometric objects scalar fields, vectors (not there components) and so on. The laws of physics are expressible as geometric relationships between these geometric objects.

$$\begin{aligned} X[f] &= X^a \frac{\partial}{\partial x^a} f \\ &= \\ &= X'^a \frac{\partial}{\partial x'^a} f \\ &= X'[f] \end{aligned} \tag{C.637}$$

the torsion $T(X,Y) = T^a{}_{bc}X^bY^c$

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] \tag{C.638}$$

$R(X,Y)Z = R^a{}_{bcd}X^cY^dZ^b$

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z + \nabla_{[X,Y]} Z \tag{C.639}$$

An element $\xi \in T_p\mathbb{R}^n$ is identified with a mapping which takes every smooth function f, defined on any neighbourhood of p, to its directional derivative at p along ξ, denoted $\xi[f]$; that is

$$\xi[f] = \frac{d}{dt} f(p + t\xi)\Big|_{t=0}. \tag{C.640}$$

Notation for vector and covector coordinates basis

Suppose we have a vector ξ and a coordinate system x^a with basis vectors $\{\mathbf{e_a}\}_x$. Recall that for the vector ξ to remain unchanged under a coordinate transformation the basis vectors $\{\mathbf{e_a}\}_x$ should transform as the components of a covector, this suggests the following alternative notation

$$\{\mathbf{e_a}\}_x \quad \equiv \quad \{\frac{\partial}{\partial \mathbf{x^a}}\} \tag{C.641}$$

so that we would write

$$\xi = \xi^a \frac{\partial}{\partial \mathbf{x^a}},$$

and under a transformation to the new coordinates x' with new basis vectors we have

$$\xi = \xi^a \frac{\partial}{\partial \mathbf{x^a}} = \xi^b \Big(\frac{\partial x'^a}{\partial x^b} \frac{\partial}{\partial \mathbf{x'^a}}\Big) = \xi'^a \frac{\partial}{\partial \mathbf{x'^a}}$$

and hence

$$\xi'^a = \xi^b \frac{\partial x'^a}{\partial x^b}$$

which corresponds to the definition of the components of a contravariant vector. For covector ω and a coordinate system with basis covectors $\{\mathbf{e^a}\}_x$. For the covector ω to remain unchanged under a coordinate transformation the co-basis vectors $\{\mathbf{e^a}\}_x$ should transform as the components of a vector, this suggests the following alternative notation

$$\{\mathbf{e^a}\}_x \quad \equiv \quad \{\mathbf{dx^a}\} \tag{C.642}$$

so that we would write

$$\omega = \omega_a \mathbf{dx^a}.$$

$$\left(\mathbf{e^a}, \mathbf{e_b}\right) = \delta^a_b \quad \equiv \quad \left(\mathbf{dx^a}, \frac{\partial}{\partial \mathbf{x^b}}\right) = \delta^a_b \tag{C.643}$$

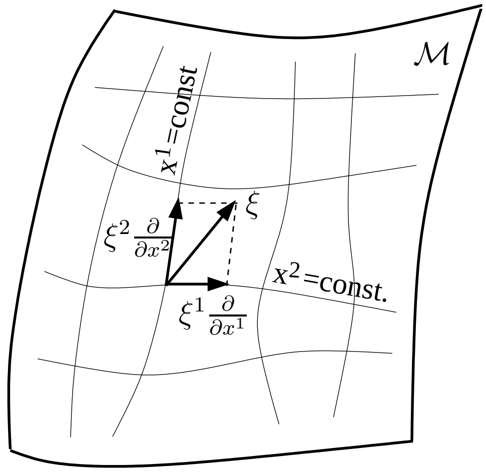


Figure C.24: coordbasevec. The vector ξ may be thought of as being composed of $\xi = \xi^1 \partial/\partial \mathbf{x^1} + \xi^2 \partial/\partial \mathbf{x^2}$. ξ^1 and ξ^2 are the components of ξ in the $(x^1, x^2)-$coordinate system.

The contraction of the vector and covector would go as

$$\begin{aligned} \left(\omega, \xi\right) &= \omega_a \xi^b \left(\mathbf{dx^a}, \frac{\partial}{\partial \mathbf{x^b}}\right) \\ &= \omega_a \xi^b \delta^a_b \\ &= \omega_a \xi^a \end{aligned} \tag{C.644}$$

A tensor with raised indices has contravariant components and is therefore expanded in terms of basis vectors, i.e.

$$\mathbf{T} = T^{ab} e_a \otimes e_b$$

A tensor with lowered indices has covariant components and is therefore expanded in terms of basis one-forms, i.e.

$$\mathbf{T} = T^{ab} \mathbf{e_a} \otimes \mathbf{e_b}$$

We have already seen an example of a tensor with lowered indicies, the metric tensor

$$\mathbf{g} = g_{ab} \omega^{\mathbf{a}} \otimes \omega^{\mathbf{b}}.$$

We have notation

$$\mathbf{T}(\omega_a, \omega_b) = T^{ab}$$

In the $x-$coordinates the metric tensor $\mathbf{g}$ would be written as

$$\mathbf{g} = g_{ab}\mathbf{dx^a} \otimes \mathbf{dx^b}.$$

In this notation an arbitrary $(p, q)-$tensor T is expanded in the above notation as

$$\mathbf{T} = T^{a_1 \ldots a_p}_{b_1 \ldots b_q} \frac{\partial}{\partial \mathbf{x^{a_1}}} \otimes \cdots \otimes \frac{\partial}{\partial \mathbf{x^{a_p}}} \otimes \mathbf{dx^{b_1}} \otimes \cdots \otimes \mathbf{dx^{b_q}} \tag{C.645}$$

C.11 Differential Geometry

C^r-function f if all its representatives $\psi \circ f \circ \phi^{-1} : R^m \to R^n$: (**??**)

In many undergraduate texts, one fixes a covering and coordinate patches and writes any tensor in terms of its value in some coordinate system. This approach is convenient an teaching of elementary GR, but it can obscure the coordinate independent meaning of important concepts. A more preferable formulation of the principle is based on modern differentiable geometry: such a formulation is coordinate free. Physical quantities are (coordinate free) geometric objects scalar fields, vectors (not there components) and so on. The laws of physics are expressible as geometric relationships between these geometric objects.

More advanced texts tend to use a coordinate free formulation, which is what we will present in the next few sections. We will see how the coordinate free approach replaces the 'tensor component' description.

Derive coordinate-free form of equations. These equations will involve objects such as vector fields, one-forms and scalar functions and geometric operations such as . The geometric objects belong to the manifold itself, be it a space-time or phase space or others, rather than any coordinate system on it. We will often use formulation in derivations employing local coordinate systems, but the definition used will hold on every chart, and hence they hold globally, making them chart independent. So we will have shown that they can be written in a consistent coordinate-independent manner.

C.11.1 Tangent Vectors

There are different ways of defining tangent vectors. Smooth manifolds embedded in Euclidean space

We don't want a definition dependent on embedding out space into a larger space. We can define tangent vectors in a way that is defined intrinsically to the manifold. a tangent vector being tangent to a curve in the manifold.

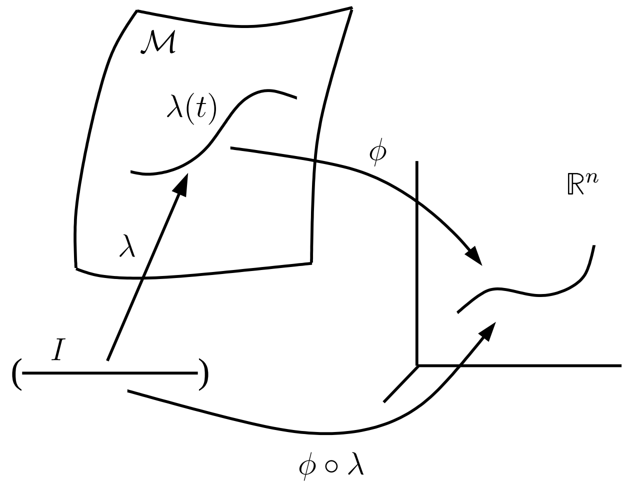


Figure C.25: tangvector. maps the tangent spaces of $\mathcal{M}$ *linearly* into the .

We will generalize the notion of a tangent vector to manifolds in a coordinate free way.

consider a curve in $\mathbb{R}^n$ $\lambda : (0,1) \to \mathbb{R}^n$. The tangent vector to a point $p = \lambda(t_1)$ is

$$\left(\frac{d\lambda^1}{dt}\Big|_{t=t_1}, \dots, \frac{d\lambda^n}{dt}\Big|_{t=t_1} \right) \tag{C.646}$$

where $\lambda(t) = (\lambda^1(t), \dots, \lambda^n(t)) \in \mathbb{R}^n$.

Basis Vectors

The rate of change of $f(\lambda(t))$ at $t = 0$ along the curve is

$$\frac{d}{dt} f(\lambda(t))\Big|_{t=0} \tag{C.647}$$

In terms of a coordinate system, this becomes

$$\left[\frac{\partial}{\partial x^a} f(x^a) \right] \frac{dx^a}{dt} (\lambda(t))\Big|_{t=0} \tag{C.648}$$

In other words, $f(\lambda(t))$ at $t = 0$ is given by applying the differential operator X to f, where

$$\frac{d\lambda(t)}{dt} = X^a \frac{\partial f}{\partial x^a} =: X[f]. \tag{C.649}$$

Thus are the components of X_p in the basis

$$\frac{\partial}{\partial x^a}\bigg|_p. \tag{C.650}$$

The coordinate transformation

$$y^b = y^b(x^1, \ldots, x^n), \quad b = 1, \ldots, n.$$

If we have a coordinate system in a neighbourhood U of P, then the coordinate basis $\{\frac{\partial}{\partial x^a}\}$.

$$\hat{V} = V^a \frac{\partial}{\partial x^a} = V'^b \hat{e}_b. \tag{C.651}$$

The numbers $\{V^a\}$ are the components of $\hat{V}$ on $\{\frac{\partial}{\partial x^a}\}$. The numbers $\{V^a\}$ are the components of $\hat{V}$ on $\{\frac{\partial}{\partial y^a}\}$.

$$\frac{\partial}{\partial x^1} = \frac{\partial y^1}{\partial x^1}\frac{\partial}{\partial y^1} + \frac{\partial y^2}{\partial x^1}\frac{\partial}{\partial y^2} + \cdots + \frac{\partial y^n}{\partial x^1}\frac{\partial}{\partial y^n}$$

and similarly for other x^as.

$$X = X^a \frac{\partial}{\partial x^a} = X'^a \frac{\partial}{\partial y^a} \tag{C.652}$$

This shows that X^a and X'^a are related by

$$X'^a = X^b \frac{\partial y^a}{\partial x^b}. \tag{C.653}$$

the components of the vector transform in such a way that the vector itself is left invariant.

The basis need not be $\{e_a\}$, we can form linear combinations $\hat{e}_i := E_i^a e_a$, where $E = (E_i^a)$ are matrices that . These are referred to as **non-coordinate basis** or a **frame**. Note that they transform as scalars under coordinate transformations. These are very important as frames fields will be basic variables in a formulation of GR that resembles Gauge field theories that are adopted for the quantum theory

C.11.2 Covectors

it is natural to regard $\{dx^a\}$ as a basis of $T_p^*\mathcal{M}$.

$$< dx^b, \partial/\partial_a >= \delta_a^b. \tag{C.654}$$

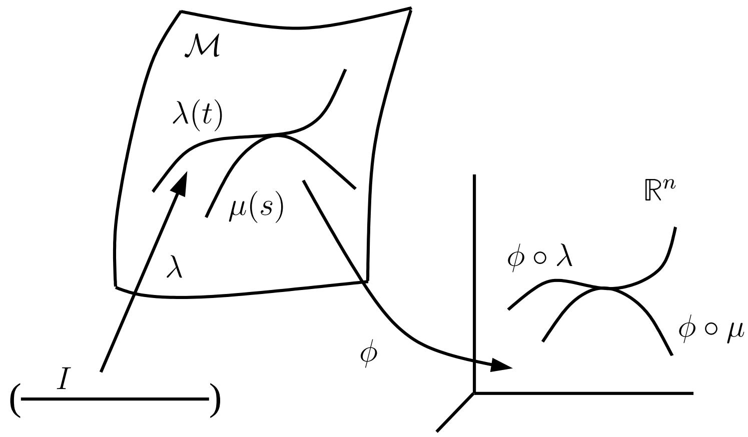


Figure C.26: Two curves $\lambda(t)$ and $\mu(t)$ are tangent at p if and only if their images are tangent at $\phi(p)$ in $\mathbb{R}^n$.

An arbitrary covector can be written

$$\omega = \omega_a dx^a \tag{C.655}$$

where ω_a are the components of ω.

Mixed tensors

The set of type (p, q). A tensor is written in terms of the coordinate basis as

$$T = T^{a_1 \ldots a_p}_{b_1 \ldots b_q} \frac{\partial}{\partial x^{a_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{a_p}} \otimes dx^{b_1} \otimes \cdots \otimes dx^{b_q}. \tag{C.656}$$

The various connections are defined by index structure, i.e. how they transform, and by restrictions placed on them.

C.11.3 Induced Metric and Other Objects on Sub-manifolds

If we denote the normal to the surface as n^a ($n^a n_a = 1$) then the induced metric can be written as

$$h_{ab} = g_{ab} - n_a n_b \tag{C.657}$$

So that h_{ab} projects out the components of a vector normal to the hypersurface. Say $\xi^b = Cn^b$

$$h^a{}_b \xi^b = \delta^a{}_b \xi^b - n^a n_b \xi^b = 0. \tag{C.658}$$

Let $\mathcal{N}$ be a $n-$dimensional manifold of an $m-$dimensional manifold . The hypersurface on which we have coordinates y^α. If $\mathcal{N}$ is a hypersurface in $\mathcal{M}$, for the point labelled by y^α corresponds to the point of labelled by $\mathcal{M}$. Thus the hypersurface is described by the equations

$$x^a(y^\alpha) \tag{C.659}$$

tangent vector e^α_a, and induced metric $h_{\alpha\beta}$

$$e^\alpha_a = \frac{\partial y^\alpha}{\partial x^a} \tag{C.660}$$

ds^2 for an infinitesimal curve lying in the hypersurface:

$$ds^2 = h_{\alpha\beta} dy^\alpha dy^\beta = h_{\alpha\beta} \frac{\partial y^\alpha}{\partial x^a} \frac{\partial y^\beta}{\partial x^b} dx^a dx^b \tag{C.661}$$

$$g_{ab} = h_{\alpha\beta} \frac{\partial y^\alpha}{\partial x^a} \frac{\partial y^\beta}{\partial x^b} \tag{C.662}$$

or

$$g_{\mathcal{M}ab}(x) = g_{\mathcal{N}\alpha\beta}(f(x)) \, \frac{\partial f^\alpha}{\partial x^a} \frac{\partial f^\beta}{\partial x^b} \tag{C.663}$$

where f^α denote the coordinates of $f(x)$.

Definition $g_{\mathcal{N}\alpha\beta}(f(x))$ is said to be the pull-back of $g_{\mathcal{M}ab}(x)$, denoted $f_* g$

$$\begin{aligned} g_{ab} dx^a \otimes dx^b &= \delta_{\alpha\beta} \, \frac{\partial f^\alpha}{\partial x^a} \frac{\partial f^\beta}{\partial x^b} \, dx^a \otimes dx^b \\ &= d\theta \otimes \theta + \sin^2\theta \, d\phi \otimes \phi. \end{aligned} \tag{C.664}$$

where

$$h_{\alpha\beta} = g_{ab} e^a_\alpha e^b_\beta \tag{C.665}$$

is the induced metric or the first fundamental form of the hypersurface. It is a scalar with respect to coordinate transformations $x^a \to x'^a$ on $\mathcal{M}$. behaves like a tensor under coordinate transformations of the manifold $\mathcal{N}$.

C.12 Active Diffeomorphisms and the Lie Derivative

Up until now we have only considered coordinate transformation, that is, passive diffeomorphisms. We now move onto active diffeomorphisms. As their formulas look very alike the two are easily mixed up. But as we have seen in chapter 1 they are quite different, active diffeomorphisms relate *distinct* spacetime geometries, whereas a coordinate transformation merely represents the same spacetime geometry in a different coordinate system.

In chapter 1 we defined an active diffeomorphism as simultaneously dragging the metric and matter fields over the spacetime manifold while keeping the coordinate lines 'attached' (fig C.12). This is called a **pushforward**.

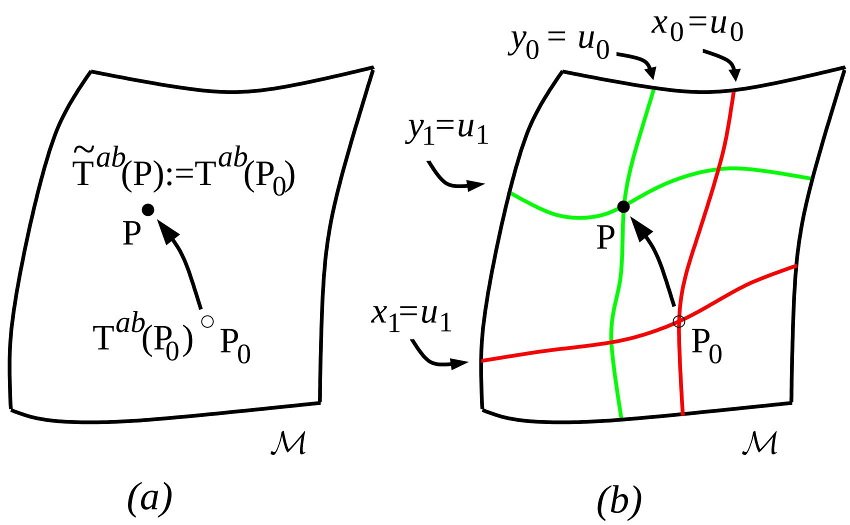


Figure C.27: activeDiffGeom. A pushforward of the tensor $T_{ab}(x)$, i.e. $T_{ab}(x) \to \tilde{T}_{ab}(y)$.

Let us slightly modify the definition of an active diffeomorphism by requiring that after we have dragged the fields across the manifold we perform a coordinate transformation back to the original coordinates. An active diffeomorphism defined this way then relates different space-time geometries and matter field configurations in the same coordinate system.

They relate $g_{ab}(x)$ to $\tilde{g}_{ab}(h(x))$ by the Jacobian matrix of the coordinate transformation $x \mapsto h(x)$,

$$\tilde{g}_{ab}(h(x)) = \Lambda^a_c \Lambda^b_d g_{ab}(x) \tag{C.666}$$

Two metrics related by an active diffeomorphism, viewed in the same coordinate system, compared at the same point also have 'transformation matrices', however, these have a different geometric interpretation!

$$\tilde{g}_{ab}(x) = \frac{\partial h^c(x)}{\partial x^a}\frac{\partial h^d(x)}{\partial x^b} g_{cd}(h(x)) \tag{C.667}$$

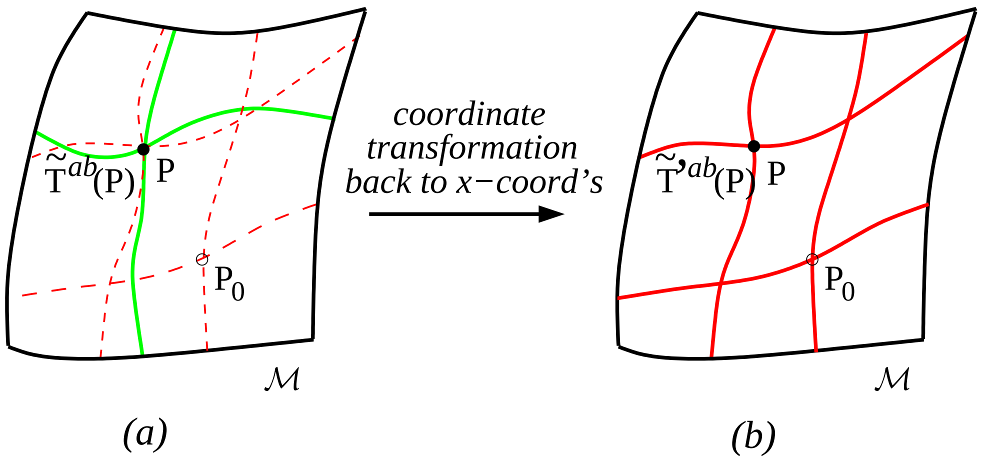


Figure C.28: activeDiffGeom1. The red dashed lines in (a) are the $x-$coordinate lines of the point P. We perform a coordinate transformation back to the original coordinate system. The pushed-forward tensor $\tilde{T}_{ab}(y)$ transforms to $\tilde{T}'_{ab}(x)$, i.e. $\tilde{T}_{ab}(y) \to \tilde{T}'_{ab}(x)$.

The fact that the coordinate values do not change, while the tensor fields do, distinguishes the active diffeomorphism from a simple coordinate transformation.

Passive diffeomorphism invariance refers to invariance under change of coordinates, i.e. the same object represented in different coordinate systems. Choose a (local) coordinate system for S in which the metirc $g_{ab}(x)$. (If the map h sends each point to the same point of the manifold $\mathcal{M}$, then in the second system S' the metric given by $\tilde{g}_{ab}(h(x))$, $f(x)$ being the coordinates on $\mathcal{M}$ of the second system.)

Any theory can be made invariant under passive diffeomorphisms because a dynamical system doesn't care which coordinate system you use to describe it. However, general relativity is the only theory invariant under active diffeomorphisms and this invariance is a property of the dynamical theory itself.

Maths Tools for Manifold Without a Metric

In the previous section we reviewed metrics on manifolds, these are important in classical general relativity and are what a physicists is most likely to be familiar with. As we have learned, in reality it is only **geometry up to active diffeomorphisms** that has physical meaning. As we have empathized, in formulating the quantum theory we prefer not to employ metrics with its direct relation to the notion of distance.

There is a rich geometric structure of the manifold without a metric defined on it. Important tools of the Lie derivative and differential forms which have nothing to do with metrics. These will be important in the quantum theory where we will avoid introducing a background metric whenever possible.

C.12.1 Mapping a Manifold to Itself Along Integral Curves

We start by considering a congruence of curves defined such that only one curve goes through each point in the manifold. Then, given any one curve of the congruence,

$$x^\mu = x^\mu(u), \tag{C.668}$$

we can use it to define the tangent vector field dx^μ/du along the curve. If we do this for every curve in the congruence, then we end up with a vector field X^μ (given by dx^μ/du at every point) defined over the whole manifold, then this can be used to define a congruence of curves in the manifold called the orbits or trajectories of X^μ.

a smooth, non-intersecting family of curves on a manifold then the tangent vectors at each point can be taken together to form a vector field on the manifold.

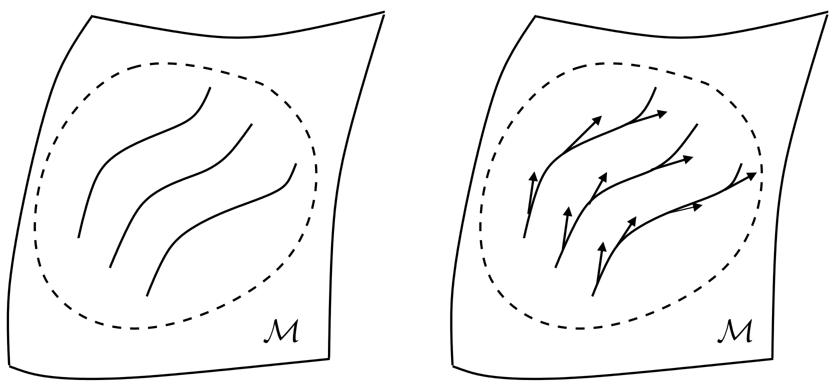


Figure C.29: The tangent vector field resulting from a congruence of curves.

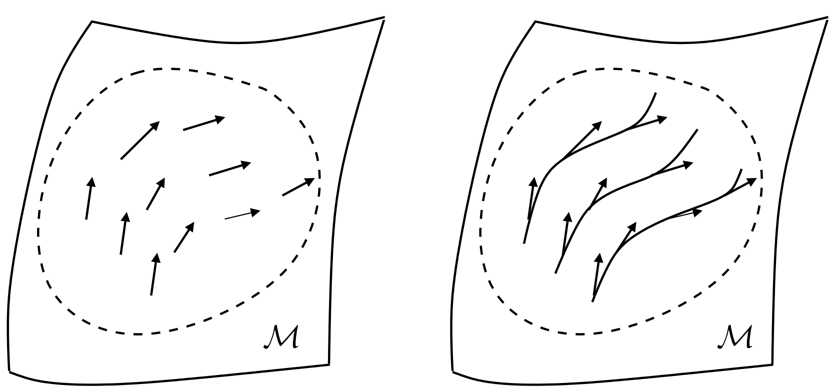


Figure C.30: The local congruence of curves resulting from vector field.

These curves are obtained by solving differential equations

$$\frac{dx^\mu}{du} = X^\mu(x(u)) \tag{C.669}$$

Let x^i be a local coordinate system and let x^i_p be the coordinates of p. The equation of the integral curve is

$$\frac{d}{dt}x^i(t) = X^i(x^m(t)),$$

with initial conditions $x^i(0) = x^i_p$. Provided X is smooth the theory of ordinary differential equations guarantees the existence and uniqueness, (at least locally, i.e., for small t), of a solution. Uniqueness implies that no two curves in the congruence intersect (at least locally).

Definition A congruence of curves is a family of curves such that precisely one curve of the family passes through each point. It is a geodesic congruence if the curves are geodesics.

Active Diffeormorphisms

C.12.2 The Lie Derivative

This is called an *active transformation*. The *passive transformation* is a coordinate transformation.

A contravariant vector flow determines a local congruence of curves,

$$x^a = x^a(u),$$

where the tangent vector field to the congruence is

$$\frac{dx^a}{du} = X^a.$$

at least locally, a vector field generates a unique integral flow about any given point p. We use this flow to take a tensor to a nearby point and hence form a derivative. This derivative is called the Lie derivative.

$$\sigma^b(\epsilon, p) = x^b(p) + \epsilon X^b(p) + \mathcal{O}(\epsilon^2) \tag{C.670}$$

The Lie derivative of a scalar field $f \in \mathcal{C}^\infty(\mathcal{M})$. Let X be a vector field on $\mathcal{M}$ we define the Lie derivative of f along X to be

$$\mathcal{L}_X f(p) = \lim_{\epsilon \to 0} \frac{f(\sigma(\epsilon, p)) - f(p)}{\epsilon} \tag{C.671}$$

which is the usual directional derivative along X.

point transform

$$x'^b = x^b(p) + \epsilon X^b(p) + \mathcal{O}(\epsilon^2) \tag{C.672}$$

We generate a new vector (with vector components in the x' coordinates). By definition its components are related to $T^a(x)$ by a pushforward

$$\tilde{T}'^a(x') := T^a\left(x^c + \epsilon X^c(x)\right) = T^a(x) + \epsilon X^c(x)\partial_c T^{ab}(x) + \mathcal{O}(\epsilon^2). \tag{C.673}$$

We now wish to transform this tensor to the $x-$coordinates so we can compare it with the original tensor $T^{ab}(x)$. Using (C.672) we have

$$\frac{\partial x^a}{\partial x'^c} = \delta^a_c - \epsilon\partial_c X^a + \mathcal{O}(\epsilon^2) \tag{C.674}$$

The parameter distance derivative of an object along the vector field is the *Lie derivative.*

$$\begin{aligned}
\tilde{T}^a(x) &= \frac{\partial x^a}{\partial x'^c}\tilde{T}'^c(x') \\
&= (\delta^a_c - \epsilon\partial_c X^a)(T^c(x) + \epsilon X^e\partial_e T^c) + \mathcal{O}(\epsilon^2) \\
&= T^a(x) + [X^e\partial_e T^a - \partial_c X^a T^c(x)]\epsilon + \mathcal{O}(\epsilon^2)
\end{aligned} \tag{C.675}$$

$$\mathcal{L}_X T^a = \lim_{\epsilon\to 0}\frac{\tilde{T}^a(x) - T^a(x)}{\epsilon} \tag{C.676}$$

$$\mathcal{L}_X T_a(x) = X^c\partial_c T_a + T_b\partial_c X^a \tag{C.677}$$

What is the Lie derivative for a tensor $T^{ab}(x)$? We generate a new tensor (with tensor components in the x' coordinates). By definition its components are related to $T^{ab}(x)$ by a pushforward

$$\tilde{T}'^{ab}(x') := T^{ab}\left(x^c + \epsilon X^c(x)\right) = T^{ab}(x) + \epsilon X^c(x)\partial_c T^{ab}(x) + \mathcal{O}(\epsilon^2). \tag{C.678}$$

We now wish to transform this tensor to the $x-$coordinates so we can compare it with the original tensor $T^{ab}(x)$. Using (C.674) again. The parameter distance derivative of an object along the vector field is the *Lie derivative.*

$$\begin{aligned}
\tilde{T}^{ab}(x) &= \frac{\partial x^a}{\partial x'^c}\frac{\partial x^b}{\partial x'^d}\tilde{T}'^{cd}(x') \\
&= (\delta^a_c - \epsilon\partial_c X^a)(\delta^b_d - \epsilon\partial_d X^b)(T^{cd}(x) + \epsilon X^e\partial_e T^{cd}) + \mathcal{O}(\epsilon^2) \\
&= T^{ab}(x) + [X^e\partial_e T^{ab} - \partial_c X^a T^{cb}(x) - \partial_d X^b T^{ad}(x)]\epsilon + \mathcal{O}(\epsilon^2)
\end{aligned} \tag{C.679}$$

$$\mathcal{L}_X T^{ab} = \lim_{\epsilon \to 0} \frac{\tilde{T}^{ab}(x) - T^{ab}(x)}{\epsilon} \tag{C.680}$$

$$\mathcal{L}_X T^{ab} = X^c \partial_c T^{ab} - T^{ac} \partial_c X^b - T^{cb} \partial_c X^a. \tag{C.681}$$

$$\mathcal{L}_X T_a(x) = X^c \partial_c T_a + T_b \partial_c X^a, \qquad \mathcal{L}_X T_{ab}(x) = X^c \partial_c T_{ab} + T_{cb} \partial_a X^c + T_{ac} \partial_b X^c$$

The first term of the Lie derivative, $X^c \partial_c$, corresponds to the pushforward, shifting the tensor to another point in the manifold. The remaining terms arise from the coordinate transformation back to the original coordinates. Is it coordinate invariant? Does it have the same form in all coordinate systems? In fact (C.677) is equivalent to:

$$X^c \nabla_c T^a - T^c \nabla_c X^a \tag{C.682}$$

since

$$\begin{aligned} \mathcal{L}_X T^a(x) &= X^c \partial_c T^a - T^c \partial_c X^a \\ &= X^c(\partial_c T^a + \Gamma^a_{dc} T^d) - T^c(\partial_c X^a + \Gamma^a_{dc} X^d) \\ &= X^c \nabla_c T^a - T^c \nabla_c X^a \end{aligned} \tag{C.683}$$

where we have used that the connection is symmetric in its lower indices. Similarly, (C.681) is equivalent to:

$$\begin{aligned} & X^c \nabla_c T^{ab} - T^{ac} \nabla_c X^b - T^{cb} \nabla_c X^a \\ = \; & X^c(\partial_c T^{ab} + \Gamma^a_{dc} T^{db} + \Gamma^b_{dc} T^{ad}) - T^{ac}(\partial_c X^b + \Gamma^b_{dc} X^d) - T^{cb}(\partial_c X^a + \Gamma^a_{dc} X^d) \\ = \; & X^c \partial_c T^{ab} - T^{ac} \partial_c X^b - T^{cb} \partial_c X^a + T^{db} X^c(\Gamma^a_{dc} - \Gamma^a_{cd}) + T^{ad} X^c(\Gamma^b_{dc} - \Gamma^b_{cd}) \\ = \; & X^c \partial_c T^{ab} - T^{ac} \partial_c X^b - T^{cb} \partial_c X^a \end{aligned} \tag{C.684}$$

In general, the partial derivatives appearing in Lie derivatives can be replaced by covariant derivatives. Hence, the combination of pushback and coordinate transformation make the Lie derivative a tensor in the tangent space at x^a.

$$\tilde{T}'(q) = T(h_\epsilon(p)) \qquad \tilde{T}(p) = h_{\epsilon *}[T(h_\epsilon(p))] \tag{C.685}$$

We can write down the coordinate free equation

$$(\mathcal{L}_X T)(p) = \frac{h_{\epsilon *}[T(h_\epsilon(p))] - T(p)}{\epsilon} \tag{C.686}$$

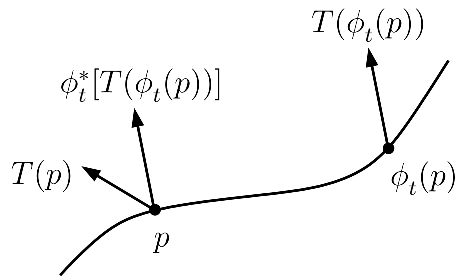


Figure C.31: .

the original tensor components at a different point. distinguishes the Lie derivative from the directional derivative.

the curve passing through P is given by x^1 varying, with x^2, x^3, x^4 all constant along the curve, and such that

$$X^\alpha \stackrel{*}{=} \delta_1^\alpha = (1, 0, 0, 0) \tag{C.687}$$

along this curve. The notation used in means that the equation holds only in a particular coordinate system. Then it follows that

$$X = X^\alpha \partial_\alpha = \partial_1, \tag{C.688}$$

and equation reduces to

$$L_X T_{\alpha\beta} \stackrel{*}{=} \partial_1 T_{\alpha\beta} \tag{C.689}$$

Thus, in this special coordinate system, Lie differentiation reduces to ordinary differentiation.

If we have a map ϕ from a manifold $\mathcal{M}$ to another manifold $\mathcal{N}$, and we choose a point $x \in \mathcal{M}$, we can *push forward* a vector from $T\mathcal{M}_x$ to $T\mathcal{N}_{\phi(x)}$, by a head-to-head and tail-to-tail map. If the vector has components X^μ and the map takes the point with coordinates x^μ to one with coordinates $\xi(x)$, the vector $\phi_* X$ has components

$$(\phi_* X)^\mu = \frac{\partial \xi^\mu}{\partial x^\nu} X^\nu. \tag{C.690}$$

This looks like the transformation formula for contravariant vector components under a coordinate transformation, but we are doing an active transformation, changing a vector into a different one.

pushforward ϕ_* (**??**)

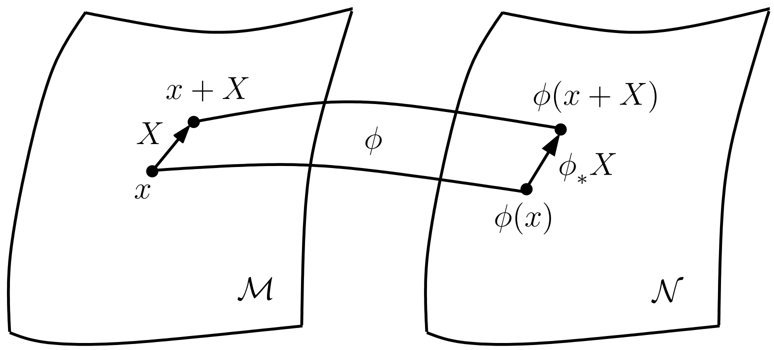


Figure C.32: pullbackDef0. Pushing forward a vector X from $T\mathcal{M}_x$ to $T\mathcal{N}_{\phi(x)}$.

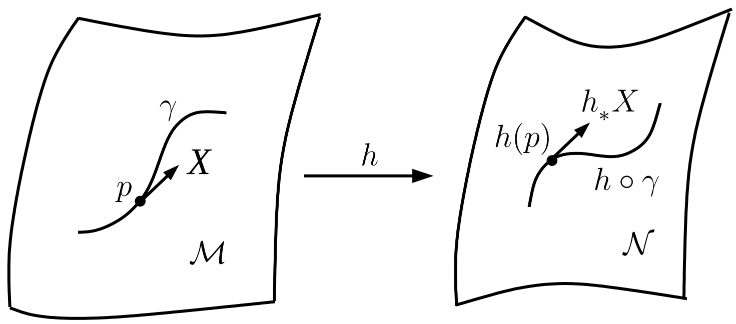


Figure C.33: The **push-forward map** h_* that maps the tangent spaces of $\mathcal{M}$ *linearly* into the tangent spaces of $\mathcal{N}$.

Recall that a one-form maps a vector to a number. Given a one-form ω on $\mathcal{N}$, we define $\phi^*\omega$ as a one-form on $\mathcal{M}$ by specifying what we get when we plug the vector X at $x \in \mathcal{M}$ into it. This we do by pushing the X forward to $T\mathcal{N}_{\phi(x)}$, plugging it into ω, and declaring the result to be the evaluation of $\phi^*\omega$ on the X. Symbolically

$$[\phi^*\omega](X) = \omega(\phi_* X). \tag{C.691}$$

or in components

$$[\phi^*\omega]_a X^a = \omega_a [\phi_* X]^a. \tag{C.692}$$

We work a coordinate system $(x^1, \ldots, x^m)$, such that x^1 is the parameter along the integral curves and the other coordinates are choosen any way. In this coordinate system and the components of the tensor pulled back from $\phi_t(p)$ to p are simply

$$\phi_{t\star}[T^{a_1 \ldots a_k}_{b_1 \ldots b_l}(\phi_t(p))] \stackrel{*}{=} T^{a_1 \ldots a_k}_{b_1 \ldots b_l}(x^1 + t, x^2, \ldots, x^n). \tag{C.693}$$

In this coordinate system the Lie derivative becomes

$$\mathcal{L}_V T^{a_1\ldots a_k}_{b_1\ldots b_l} \stackrel{*}{=} \frac{\partial}{\partial x^1} T^{a_1\ldots a_k}_{b_1\ldots b_l}, \tag{C.694}$$

Coordinate-Free Description

We will prove

$$\mathcal{L}_X Y = [X, Y]. \tag{C.695}$$

$$\sigma^b(\epsilon, p) = x^b(p) + \epsilon X^b(p) + \mathcal{O}(\epsilon^2) \tag{C.696}$$

for any f

$$\begin{aligned}
Y_{\sigma(\epsilon)} f &= \sum_a Y^b(\sigma(\epsilon)) \frac{\partial}{\partial x^a}\Big|_{\sigma(\epsilon)} f \\
&= \sum_a (Y^b + \epsilon X^c \partial_c Y^a)\Big|_{\sigma(\epsilon)} f + \mathcal{O}(\epsilon^2) \\
&= \sum_a (Y^b + \epsilon X^c \partial_c Y^a)(\partial f + \epsilon X^c \partial_a \partial_c f) + \mathcal{O}(\epsilon^2) \\
&= \sum_a (Y^b \partial_a + \epsilon X^c \partial_c Y^a \partial_a + \epsilon Y^a X^c \partial_c \partial_a) f + \mathcal{O}(\epsilon^2)
\end{aligned} \tag{C.697}$$

Therefore

$$\begin{aligned}
\sigma(-\epsilon)_* Y_{\sigma(\epsilon)} f &= Y_{\sigma(\epsilon)}(f \circ \sigma(-\epsilon)_*) \\
&= \sum_a Y^a\ (h \circ \sigma(\epsilon)) \frac{\partial}{\partial x^a}\Big|_p (h \circ \sigma(-\epsilon)) + \mathcal{O}(\epsilon^2) \\
&= \sum_a (Y^b \partial_a + \epsilon X^c \partial_c Y^a \partial_a + \epsilon Y^a X^c \partial_c \partial_a)(f - \epsilon \partial_b f X^b) + \mathcal{O}(\epsilon^2) \\
&= \sum_a Y^b \partial_a f + \epsilon(\partial_c Y^a X^c - Y^c \partial_c X^a)\partial_a f + \mathcal{O}(\epsilon^2)
\end{aligned} \tag{C.698}$$

From which

$$\begin{aligned}
\frac{\sigma(-\epsilon)_* Y_{\sigma(\epsilon)}(f) - Y(f)}{\epsilon} &= (\partial_c Y^a X^c - Y^c \partial_c X^a)\partial_a f + \mathcal{O}(\epsilon^2) \\
&= [X, Y]^a \partial_a f + \mathcal{O}(\epsilon^2) \\
&= [X, Y]^a(f) + \mathcal{O}(\epsilon^2)
\end{aligned} \tag{C.699}$$

The Lie derivative of a covariant tensors

$$L_X Y^\alpha = X^\beta \partial_\beta Y^\alpha - Y^\beta \partial_\beta X^\alpha \tag{C.700}$$

The Lie derivative of a covariant vector field Y_α is given by

$$L_X Y_\alpha = X^\beta \partial_\beta Y_\alpha + Y^\beta \partial_\alpha X^\beta \tag{C.701}$$

The Lie Derivative

there is a coordinate system in which

$$\mathcal{L}_{\vec{N}} \vec{M} = N^a \partial_a M_b - M^a \partial_a N_b \tag{C.702}$$

It satisfies the Leibniz rule

$$L_X(Y^a Z_{bc}) = Y^a(L_X Z_{bc}) + (L_X Y^a) Z_{bc}. \tag{C.703}$$

It is type-preserving; that is, the Lie derivative of s tensor of type(p,q) is again a tensor of type (p,q).

The Lie derivative of a scalar field ϕ is simply an ordinary derivative in the direction of X

$$L_X \phi = X\phi = X^a \partial_a \phi \tag{C.704}$$

Now, given the Lie derivative of a vector and a scalar, we can apply the Leibniz rule to deduce the Lie derivative of a covariant vector field Y_a: consider the Lie derivative of the scalar formed by the contraction of an arbitrary vector Z^a with an arbitrary covector Y_a.

$$L_X(Y_c Z^c) = X^b \partial_b (Y_c Z^c) = Z^c X^b \partial_b Y_c + Y_c X^b \partial_b Z^c \tag{C.705}$$

whereas the Leibniz rule gives

$$L_X(Y_c Z^c) = Y_c L_X Z^c + (L_X Y_c) Z^c. \tag{C.706}$$

$$Z^c L_X Y_c = Z^c X^b \partial_b Y_c + Z^c Y_b \partial_a X^c. \tag{C.707}$$

but as Z^c is arbitrary this means

$$L_X Y_a == X^b \partial_b Y_a + Y_b \partial_a X^b. \tag{C.708}$$

$$\bar{T} := \phi_{\Delta\lambda}T(p_0) \quad \overline{x}^a(P) = x^a(P_0) \tag{C.709}$$

$$\mathcal{L}_\xi T := \lim_{\Delta\lambda} \frac{\phi_{\Delta\lambda}T(x) - T(x)}{\Delta\lambda} \tag{C.710}$$

$$\mathcal{L}_X T^{a\ldots}_{b\ldots} = X^c \partial_c T^{a\ldots}_{b\ldots} - T^{c\ldots}_{b\ldots}\partial_c X^a - \cdots + T^{a\ldots}_{c\ldots}\partial_b X^c + \cdots \tag{C.711}$$

C.12.3 Pull-back and Lie Derivative of a co-vector

The pullback of the function f by ϕ, denoted $\phi^* f$, is defined by

$$\phi^* f = (f \circ \phi). \tag{C.712}$$

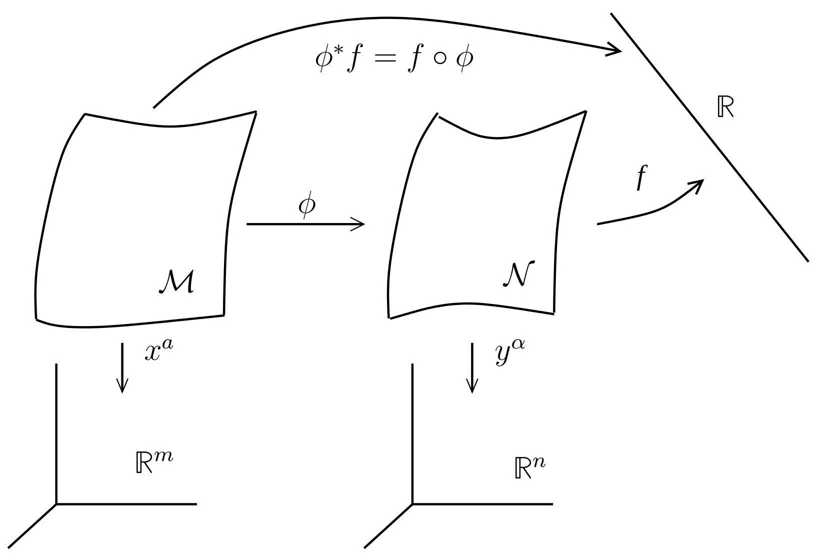


Figure C.34: The **pullback map** ϕ^* of a function f from $\mathcal{N}$ to $\mathcal{M}$ by a map $\phi : \mathcal{M} \to \mathcal{N}$ is the composition of ϕ with f.

Suppose we have a smooth map $h : \mathcal{M} \to \mathcal{N}$ as in section C.11.1. We saw that we could push-forward a vector $X_p \in T_p\mathcal{M}$ to a vector $h_* X_{h(p)} \in T_{h(p)}\mathcal{N}$ by

$$h_* X_{h(p)}(g) = X_p(g \circ h) \tag{C.713}$$

There should be a dual map which maps co-vectors in $\mathcal{N}$ to co-vectors in $\mathcal{M}$.

$$h_p^*\omega(X_p) =< h^*\omega, X_p > = < \omega, h_* X_{h(p)} > = \omega(h_* X_{h(p)}) \tag{C.714}$$

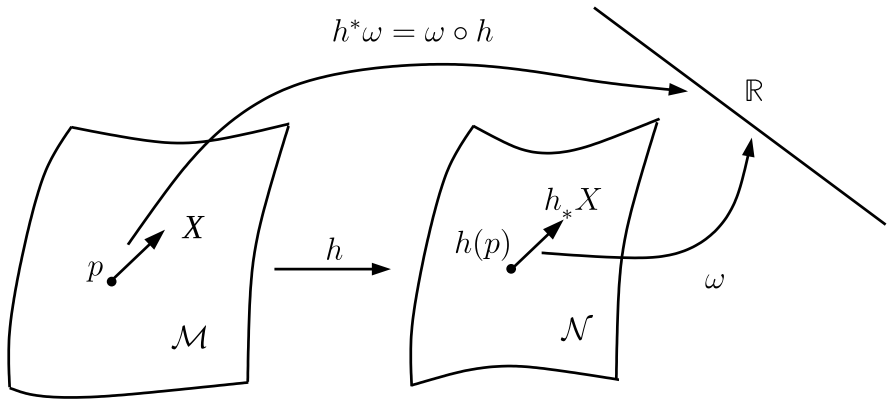


Figure C.35: pushLie. The maps the co-tangent spaces of $\mathcal{M}$ *linearly* into the co-tangent spaces of $\mathcal{N}$.

Let $h : \mathcal{M} \to \mathcal{N}$, $(y^1, \ldots y^m)$ be local coordinates on $V \subset \mathcal{N}$ and $(x^1, \ldots x^n)$ be local coordinates on $U \cap h^{-1}(V) \subset \mathcal{M}$. If

$$\omega = \sum_{b=1}^{n} \omega_b dy^b \Big|_p \tag{C.715}$$

We have

$$< \omega, h_* X_{f(p)} > \quad = \quad \sum_{b=1}^{n} \omega_b (f^* \omega X_{f(p)})^b \tag{C.716}$$

C.12.4 More on Lie Derivative

Definition A one-parameter group of diffeomorphisms. A one-parameter family of maps $\{\phi_t\}_{t \in \mathbb{R}}$ is said to be a one parameter group of diffeommorphism if:

(i) Each $\phi_t : \mathcal{M} \to \mathcal{M}$ is a diffeomorphism;

(ii) $\phi_0 = \text{id}$;

(iii) $\phi_{s+t} = \phi_s \circ \phi_t$ for all $s, t \in \mathbb{R}$.

That is we have a group action of $\mathbb{R}$ on $\mathcal{M}$.

□

Lemma C.12.1 *Let φ_t be the one parameter group of diffeomorphisms generated by the complete vector field X on the manifold $\mathcal{M}$, and $\psi : \mathcal{M} \to \mathcal{M}$ is a diffemorphism on $\mathcal{M}$. Then $\psi \circ \varphi_t \circ \psi^{-1}$ is the one-parameter group of diffeomorphisms generated by $\psi_* X$.*

Proof: We show that the tangent to the curve $\phi_p(t) := (\psi \circ \varphi_t \circ \psi^{-1})(p)$

$$\begin{aligned}
(\psi_* X)_p f &:= X_q(f \circ \psi) = \frac{d}{dt} f \circ \psi(\varphi_t(q))\Big|_{t=0} \\
&= \frac{d}{dt} f\big(\psi \circ \varphi_t \circ \psi^{-1}(\psi(q))\big)\Big|_{t=0} \\
&= \frac{d}{dt} f\big(\psi \circ \varphi_t \circ \psi^{-1}(p)\big)\Big|_{t=0} \\
&= \frac{d}{dt} f\big(\phi_p(t)\big)\Big|_{t=0} \qquad \text{(C.717)}
\end{aligned}$$

□

Corollary C.12.2 *A complete tangent vector field is invariant under a diffeomorphism $\psi : \mathcal{M} \to \mathcal{M}$ if and only if the one-parameter group of diffeomorphisms φ_t generated by X commutes with ψ.*

Proof:

□

C.12.5 Isometries and Killing Vector Fields

when the metric is the same. If you move along the direction of a Killing vector field, the metric diesn't change.

A space time possess a symmetry if there exists a coordinate system such that the components of the components $g_{ab}(x)$ of the metric are independent of at least one or more of the coordinates. Then the metric has a symmetry under translations by this coordinate holding the remaining coordinates fixed.

see M. Gockeler, T. Schucker, *Differential geometry, gauge theories, and gravity*

in general these symmetries go when we go to curved space-time with fixed metric and cannot exist in general relativity where the metric becomes a dynamical variable.

Then a transformation leaving $g_{ab}(x)$ invariant is called an **isometry**.

The Lie derivative is natural to express the invariance of a tensor under a change of position. The vector ξ^a that generates the symmetry is called a Killing vector. In the original coordinates the components of the Killing vector are simply $\xi^a = \delta^a_b$. A coordinate covariant characterization of a

Killing vector is

$$g_{ab}(x) = \frac{\partial x'^c}{\partial x^a}\frac{\partial x'^d}{\partial x^b} g'_{ab}(x') \tag{C.718}$$

will be an isometry if

$$g_{ab}(x) = \frac{\partial x'^c}{\partial x^a}\frac{\partial x'^d}{\partial x^b} g_{ab}(x') \tag{C.719}$$

There is a coordinate system in which $g_{ab}(x)$ is the same function that $g'_{ab}(x')$, where $x'^a = x^a + \epsilon K^a$

For example

$dl^2 = dr^2 + r^2\, d\theta^2$

$g_{ab}(r, \theta) = D(1, r^2)$ and $g_{ab}(1, \theta') = D(1, r'^2)$ where $\theta' = \theta + \epsilon K_\alpha$

and give curves along which the geometrical environment is unchanged.

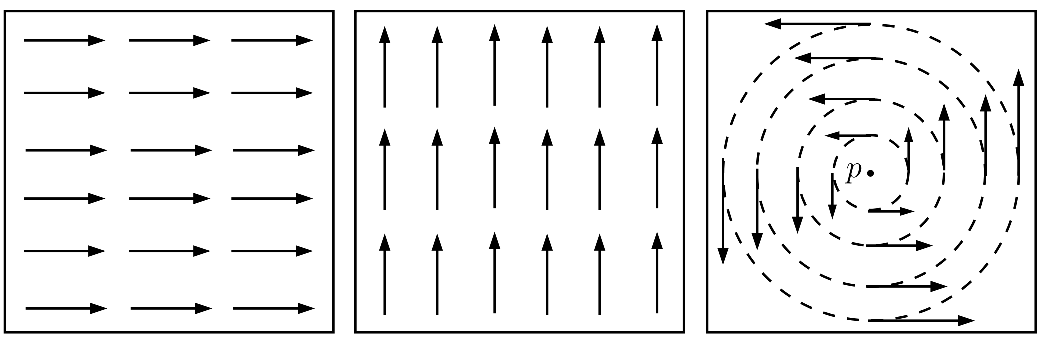


Figure C.36: The Killing vector field resulting from a congruence of curves.

Fig.(C.12.5). The first two are the "infinitesimal generators" of horizontal and vertical translations. The third is the generator of counter-clockwise rotations centered at p.

The three dimensional rotation group $O(3)$ is the isometry group for of the ordinary round sphere S^2.

$$\frac{\partial x^a}{\partial x^c}\frac{\partial x^b}{\partial x^d} \tag{C.720}$$

$$\frac{\partial x'^a}{\partial x^b} = d^\alpha_b + \epsilon \partial_b K^a \tag{C.721}$$

$$g'_{ab}(x') = g_{ab}(x) \ \text{ when } x' = xK_a. \tag{C.722}$$

$$g'_{ab}(x^a + \epsilon K^a) = g_{ab}(x) \tag{C.723}$$

$$g'_{\alpha\beta}(x') = (\delta^\gamma_\alpha + \partial_\alpha K^\gamma)(\delta^\delta_\beta + \partial_\beta K^\delta)g_{\gamma\delta}(x^\sigma + \epsilon K^\sigma) \tag{C.724}$$

$$\begin{aligned}
&= (\delta^\gamma_\alpha + \partial_\alpha K^\gamma)(\delta^\delta_\beta + \partial_\beta K^\delta)[g_{\gamma\delta}(x^\sigma) + \epsilon K^\sigma \partial_\sigma g_{\gamma\delta} + \dots] \\
&= g_{\alpha\beta}(x) + \epsilon[g_{\alpha\delta}\partial_\beta K^\delta + g_{\beta\delta} + K^\sigma \partial_\sigma g_{\alpha\beta}] + \mathcal{O}(\epsilon^2).
\end{aligned} \tag{C.725}$$

$$\mathcal{L}_X = X^\sigma \partial_\sigma g_{\alpha\beta} + g_{\alpha\delta}\partial_\beta X^\delta + g_{\beta\delta}\partial_\alpha X^\delta. \tag{C.726}$$

$$\mathcal{L}_K g_{\alpha\beta}(x) = 0 \tag{C.727}$$

$$\mathcal{L}_K g_{\alpha\beta}(x) = \nabla_\alpha K_\beta + \nabla_\beta K_\alpha = 0, \qquad \nabla_{(\beta} K_{\alpha)} = 0 \tag{C.728}$$

An isometry is generated by a Killing vector field K^α satisfying $\mathcal{L}_K g_{\alpha\beta}(x) = 0$

C.12.6 Conserved Quantities

$\nabla_{(a} k_{b)} = 0$

Consider a freely falling particle whose worldline has tangent vector X. Define the quantity $E = X^a k_a$, where k is a Killing vector. Then

$$\begin{aligned}
X^a \nabla_a (X^c k_c) &= X^a X^b \nabla_a k_b + k_b \underbrace{X^a \nabla_a X^b}_{=0} \\
&= X^a X^b \nabla_a k_b = 0 \\
&= X^a X^b \nabla_{(a} k_{b)} = 0
\end{aligned} \tag{C.729}$$

Thus E is conserved along the worldline of X. Given the energy-momentum tensor of a continuous distribution of matter, satisfying $\nabla_b T^{ac} = 0$. Define $J^a := T^{ac} k_c$. Then

$$\begin{aligned}
\nabla_c (T^{bc} k_b) &= k_b \underbrace{\nabla_c T^{bc}}_{=0} + T^{bc} \nabla_c k_b \\
&= T^{bc} \nabla_c k_b \\
&= T^{bc} \nabla_{(c} k_{b)} = 0
\end{aligned} \tag{C.730}$$

Thus $\nabla_a J^a = 0$, i.e., the current is conserved.

Notion of energy and angular momentum have played a key role in analyzing behaviour of physical theories. For theories of fields on a fixed, background spacetime, a locally conserved stress-energy tensor, T_{ab}, normally can be defined. If the background spaceime has a Killing field, k^a, then $J^a = T^a{}_b k^b$ is a locally conserved current. If Σ is a Cauchy surface, then $q = \int_\Sigma J^a d\Sigma_a$ defines a conserved quantity associated with k^a; if Σ is a timelike or null surface, then $\int_\Sigma J^a d\Sigma_a$ has the interpretation of the flux of this quantity through Σ.

However, in diffeomorphism covariant theories such as general relativity, there is no notion of the local stress-energy tensor of the gravitational field, so conserved quantities () cannot and their fluxes cannot be defined by the above procedures, even when Killing fields are present.

C.12.7 Adapted Coordinates

These symmetries are removed by active diffeomorphisms, symmetries help us find simple solutions to Einstein's equation but strictly it is only those properties shared by all the spacetimes in the symmetry class that have physical meaning.

However, if we are neglecting the dynamics of gravity and only consered with dynamical theories over curved spacetime, then the particulary simple spacetime in the equivalence class can be used to get your physical properties.

Spherically symmetric spacetimes

$$ds^2 = g_{00}dt^2 + 2g_{oi}dtdx^i + dr^2 + r^2 \sin^2\theta d\theta^2 + r^2 d\phi^2. \tag{C.731}$$

Axisually symmetric spacetimes

$$ds^2 = dr^2 + r^2 \sin^2\theta d\theta^2 + r^2 d\phi^2. \tag{C.732}$$

A gravitational field is said to be sationary when a reference frame exists in which all the components g are independent of the time coordinate g_{ab}. This coordinate, by the way, is usually referred to as coordinate time. Sationary spacetimes.

$$ds^2 = g_{00}dt^2 + 2g_{oi}dtdx^i + g_{ij}dx^i dx^j. \tag{C.733}$$

Static spacetimes we when changing $dt \to -dt$ ds should remain unchanged.

$$ds^2 = g_{00}dt^2 + g_{ij}dx^i dx^j. \tag{C.734}$$

C.12.8 Properties of Killing Fields

A very important and immediate result is the following. If $k^b k_b = 0$, then from $\nabla_a k_b = \nabla_b k_a$ we have $k^b \nabla_b k_a = -k^b \nabla_a k_b = \nabla_a (k^b k_b) = 0$, i.e., the curve to which k is tangent is a null geodesic.

$$\nabla_a \nabla_b v_c = R_{abc}{}^d v_d \tag{C.735}$$

Proof:

$$\nabla_a \nabla_b v_c - \nabla_b \nabla_a v_c = -R_{abc}{}^d v_d, \tag{C.736}$$

which on using Killing's equation, gives

$$\nabla_a \nabla_b v_c + \nabla_b \nabla_c v_a = -R_{abc}{}^d v_d \tag{C.737}$$

$$R_{abcd} + R_{adbc} + R_{acdb} = 0, \tag{C.738}$$

we have that

$$\begin{aligned} 2\nabla_b \nabla_c v_a &= -(R_{abc}{}^d + R_{bca}{}^d - R_{cab}{}^d) v_d \\ &= 2R_{cab}{}^d v_d. \end{aligned} \tag{C.739}$$

C.12.9 Diffeomorphism Gauge Group - Symmetry of GR Under Active Diffeomorphisms

It is often stated that coordinate transformations are the gauge symetries of GR. Then move onto the diffeomorphism group, however, the diff group is formed by active diffeomorphisms not coordinate transformations! The gauge symmetry referred to is GR's invariance under active diffeomorphisms!

Let us consider an infinitessimal point transformation

$$x'^a = x^a + \xi^a(x) \tag{C.740}$$

We have already proven that the metric $g_{ab}(x)$ gets mapped to the metric $g_{ab}(x) - 2D_{(a}\xi_{b)}$ under this point transformation.

Lie algebra of vector fields ξ

A vector space V with elements $x, y, z, \ldots$ and bilinear bracket $[\cdots, \cdot]$, that takes two elements of V and returns another element of V, is a Lie algerbra if the bracket is antisymmetric and the Jacobi identity holds for all elements in V.

vector field $\xi^a(x)$ generates an infinitesimal active diffeomorphism. Has a Lie algebra

$$[\xi_{(1)}, \xi_{(2)}]f = \tag{C.741}$$

C.13 Frame Fields

That is absolutely crucial to the loop quantum gravity programme is that GR can be put into a form that strongly resembles gauge theories in particle physics.

In appendix we introduced a natural basis for the tangent space T_P at a point P that were induced by the coordinates. We consider a set of basis vectors *not* derived from any coordinate system. Say we are given a time-like vector field v^α which defines a congruence of curves. For each of these curves, take any point P. We introduce a orthonormal frame of three unit space-like vectors.

$$e_I^\alpha = (e_1^\alpha, e_2^\alpha, e_3^\alpha) \tag{C.742}$$

which are orthogonal to v^α and where I is a label running from 0 to 3.

We define

$$e_0^\alpha := v^\alpha \tag{C.743}$$

orthonomality relations

$$\begin{aligned} &e_1^a e_{1a} = e_2^a e_{2a} = e_3^a e_{3a} = -e_0^a e_{0a} = 1 \\ &e_0^a e_{1a} = e_0^a e_{2a} = e_0^a e_{3a} = e_1^a e_{2a} = e_1^a e_{3a} = e_2^a e_{3a} = 0 \end{aligned} \tag{C.744}$$

The four vectors are said to form a *frame* or *tetrad* at P, and the orthonormality relations can be succinctly summarized as

$$e_i^a e_{ja} = \eta_{ij} \tag{C.745}$$

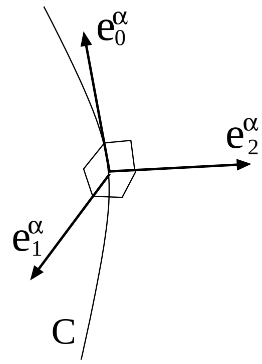


Figure C.37: Framefield or tetrad with one spatial dimension suppressed.

C.14 The Spin Connection

The triad is not a vector basis induced by a coordinate system. It turns out that the use of such frame fields brings out a different point of view on the connection and curvature, one in which GR has a strong resemblance to particle physics field theories.

Instead of a basis determined by coordinates, $\partial/\partial x^a$, we may choose any other n linearly independent vectors $\mathbf{e}_\alpha(\alpha = 1, \ldots, n)$, with components $\mathbf{e}_\alpha^\mathbf{a}(\alpha = 1, \ldots, n)$ with respect to the coordinate basis.

We have the freedom to choose a different basis. The metric $g_{ab}(x)$ is left invariant under local $SO(3,1)$ transformations such that

$$e_I^a(x^a) \rightarrow e'^a_I(x^a) = O_i^J(x^a) e_J^a(x^a), \tag{C.746}$$

where $O^j_{\ i}(x^a)$ is a matrix in $SO(3,1)$ which depends on position in space. When "Physical quantities" are left invariant, such transformations are known as *gauge transformations*, and theories invariant under them are called *gauge theories.*

Now we have introduced these frame fields we now need to know how to compare vectors in frames at different points. Put another way; a difficulty arise when one considers partial derivatives, $\partial_a V^i$. Because the matrix $O_i^j(x^a)$ depends on spacetime, it will contribute an inhomogeneous term to the transformation of the partial derivative,

$$\begin{aligned} \partial_a V'^I(x) &= \partial_a \left(O^I_{\ J}(x^a) V^J(x^a) \right) \\ &= O^I j(x^a) \partial_a V^J(x^a) + V^J(x^a) \partial_a O^I_{\ J}(x^a) \end{aligned} \tag{C.747}$$

The same sort of problem is encountered when considering the transformation of the partial differentiation of vector fields $\partial_a V^b(x)$. The solution there is to add the connection $\Gamma^a_{bc}(x)$ to correct for the inhomogeneous term in the transformation law, giving us the covariant derivative, $\nabla_a V_b = \partial_a V_b + \Gamma^c_{ab} V_c$. The same remedy is applied to $\partial_a V^i(x^a)$ and we introduce a connection

$$\omega_a{}^I{}_J(x) \tag{C.748}$$

with two tetrad indicies and one spacetime index.

$$\mathcal{D}_a V^I(x) = \partial_a V^I(x) + \omega_a{}^I{}_J V^J(x) \tag{C.749}$$

We require $\mathcal{D}_a V^i(x)$ to transform as a vector in internal space,

$$\mathcal{D}'_a V'^I(x) = O^I{}_J \mathcal{D}_a V^I(x). \tag{C.750}$$

Therefore the connection transforms as

$$\omega'_a{}^I{}_J = O^I{}_K \omega_a{}^K{}_J - \partial_a O^I{}_J \tag{C.751}$$

$$\mathcal{D}_a V^I = \partial_a V^I + \omega^I_{aJ} V^J \tag{C.752}$$

$$\mathcal{D}_a V^I_b = \partial_a V^I_b + \Gamma^c_{ab} V^I_c + \omega^I_{aJ} V^J_b \tag{C.753}$$

This covariant derivative is said to be compatible to the tetrad metric η_{IJ} if,

$$\mathcal{D}_a \eta_{IJ} = 0 \tag{C.754}$$

This implies,

$$\partial_a \eta^{IJ} + \omega^I_{aK} \eta^{JK} + \omega^I_{aK} \eta^{KJ} = 0, \tag{C.755}$$

This implying that the connection is antisymmetric in its tetrad indices,

$$\omega^{IJ}_a = -\omega^{JI}_a. \tag{C.756}$$

---

the connection $\Gamma^\alpha_{\mu\nu}$ is uniquely determined by the requirement

$$\mathcal{D}_\mu e^I_\nu(x) = 0 \tag{C.757}$$

that is,

$$\partial_\mu e^I_\nu + \Gamma^I_{\alpha J} e^j_b + \Gamma^\gamma_{\alpha\beta} e^I_\gamma = 0 \tag{C.758}$$

It is said to be compatible to the co-triad.

The connection field $\Gamma^\alpha_{\mu\nu}$ can be calculated in much the same way as the $\Gamma^\gamma_{\alpha\beta}$ was calculated.

C.14.1 Curvature Associated with the Spin Connection.

Let us work out the commutator in the case of a vector λ_I.

We will need the covariant derivative of $\mathcal{D}_\beta \lambda$. This is a tensor with one internal space index and one space-time index, so it's covariant derivative is

$$\mathcal{D}_\alpha T_{\beta I} = \partial_\alpha T_{\beta I} + \Gamma^\rho_{\alpha\beta} T_{\rho I} + \omega^J_{\beta I} T_{\beta J} \tag{C.759}$$

$$\begin{aligned} \mathcal{D}_\alpha(\mathcal{D}_\beta \lambda_I) &= \partial_\alpha(\mathcal{D}_\beta \lambda_I) + \omega^J_{\beta I}(\mathcal{D}_\beta \lambda_J) + \Gamma^\rho_{\alpha\beta}(\mathcal{D}_\rho \lambda_I) \\ &= \partial_\alpha[\partial_\beta \lambda_I + \omega^J_{\beta I}] + \omega^K_{\alpha I}[\partial_\beta \lambda_I + \omega^J_{\beta I}] + \Gamma^\rho_{\alpha\beta} \mathcal{D}_\rho \lambda_I \end{aligned} \tag{C.760}$$

$$\begin{aligned} \mathcal{D}_\alpha \mathcal{D}_\beta \lambda_I - \mathcal{D}_\beta \mathcal{D}_\alpha \lambda_I &= \Gamma^\rho_{\alpha\beta} \mathcal{D}_\rho \lambda_I - \Gamma^\rho_{\beta\alpha} \mathcal{D}_\rho \lambda_I \\ &= \partial_\alpha \partial_\beta \lambda_I - \partial_\beta \partial_\alpha \lambda_I \\ &+ \partial_\alpha(\omega^J_{\beta I} \lambda_I) - \partial_\beta(\omega^J_{\alpha I} \lambda_I) \\ &+ \omega^K_{\alpha I} \partial_\beta \lambda_K - \omega^K_{\beta I} \partial_\alpha \lambda_K \\ &+ \omega^K_{\alpha I} \omega^J_{\beta K} \lambda_J - \omega^K_{\beta I} \omega^J_{\alpha K} \lambda_J \end{aligned} \tag{C.761}$$

The first line is zero as we assume partial derivatives commute. The terms in the third line are cancelled by terms in the second line. So we obtain the result

$$\mathcal{D}_\alpha \mathcal{D}_\beta \lambda_I - \mathcal{D}_\beta \mathcal{D}_\alpha \lambda_I = R_{\alpha\beta\ I}{}^J \lambda_J \tag{C.762}$$

where $R_{\alpha\beta\ I}{}^J$ is defined by

$$R_{\alpha\beta\ I}{}^J = \partial_\alpha \omega_{\beta I}{}^J - \partial_\beta \omega_{\alpha I}{}^J + \omega_{\alpha I}{}^K \omega_{\beta K}{}^J - \omega_{\beta I}{}^K \omega_{\alpha K}{}^J \tag{C.763}$$

which can be written in the more compact form

$$R_{\alpha\beta}{}^{IJ} = \partial_{[\alpha} \omega_{\beta]}{}^{IJ} + \omega_{[\alpha}{}^{IK} \omega_{\beta]K}{}^J \tag{C.764}$$

By considering the covariant derivative of $K_l = e^I_\alpha \lambda_I$ we can find the relation between the spin curvature with the

$$\begin{aligned} R^\alpha_{\beta\gamma\rho}(e^J_\alpha \lambda_J) = R^\alpha_{\beta\gamma\rho} K_\alpha &= \mathcal{D}_\beta \mathcal{D}_\rho K_\gamma - \mathcal{D}_\rho \mathcal{D}_\beta K_\gamma \\ &= \mathcal{D}_\beta \mathcal{D}_\rho (e^I_\gamma \lambda_I) - \mathcal{D}_\rho \mathcal{D}_\beta (e^I_\gamma \lambda_I) \\ &= e^I_\gamma [\mathcal{D}_\beta \mathcal{D}_\rho \lambda_I - \mathcal{D}_\rho \mathcal{D}_\beta \lambda_I] \\ &= e^I_\gamma R_{\beta\rho I}{}^J l \end{aligned} \tag{C.765}$$

Direct product

If A_{ij} is a $m \times m$ matrix and B_{ij} is a $n \times n$ matrix, the direct product is

$$C = A \otimes B$$

where C is an $mn \times mn$ matrix with elements

$$C_{\alpha\beta} = A_{ij} B_{kl}$$

with

$$\alpha = n(i-1) + k, \quad \beta = n(j-1) + l$$

.

The determinant of $C_{\alpha\beta}$ is given by the usual formula

$$detC_{\alpha\beta} = \frac{1}{(mn)!} \sum_{\alpha_1 \beta_1} \cdots \sum_{\alpha_N \beta_N} \epsilon_{\alpha_1 \ldots \alpha_N} \epsilon_{\beta_1 \ldots \beta_N} C_{\alpha_1 \beta_1} \ldots C_{\alpha_N \beta_N}$$

.

In terms of A_{ij} and B_{kl} this becomes

$$\begin{aligned} detC_{\alpha\beta} = & \frac{1}{(mn)!} \sum_{i_1, j_1, k_1, l_1} \cdots \sum_{i_N, j_N, k_N, l_N} \epsilon_{n(i_1-1)+k_1 \cdots n(i_N-1)+k_N} \epsilon_{n(j_1-1)+l_1 \cdots n(j_N-1)+l_N} \\ & A_{i_1 j_1} B_{k_1 l_1} \ldots A_{i_N j_N} B_{k_N l_N} \end{aligned} \tag{C.766}$$

.

C.15 Differential Forms

An alternative approach to multivariable calculus, including line, surface and volume integrals, which is ultimately more powerful than vector calculus, and, being coordinate free is ideally suited to the context of differential manifolds without envoking any coordinate system.

The alagebraic approach to exterior calaulus is efficient in terms of proof but lacks intuiative content.

A first step to eleviate this problem let us first give a simple "non-calculus" example of what is known as a **form**, the main mathematical object involved in the theory.

Example:

$\square$

We give examples of differential forms.

Example:

Parametrize γ with some parameter t, so that its coordinates are given by $x^i(t)$. At time t the 'velocity' $dx(t)/dt$ is a tangent vector to $\mathcal{M}$ at $x(t)$.

we can form an integration along a curve in $\mathcal{M}$.

$$\int \left(A_i(x(t)) \frac{dx^i}{dt} \right) dt. \tag{C.767}$$

$$\int_\gamma A_i(x) dx^i, \tag{C.768}$$

or even more compact form

$$\int_\gamma A. \tag{C.769}$$

$$d\phi(x) = \frac{\partial \phi(x)}{\partial x^a} dx^a \tag{C.770}$$

$$dx^i \otimes dx^j(v^1, v^2) = dx^i(v^1) dx^2(v^2) \tag{C.771}$$

The righthand side is the multiplication of the two ordinary numbers obtained by letting the linear map dx^i act on v^1, and dx^j act on v^2.

The integral would change sign if we change the orientation of the integration, just like in the one-dimensional case. We would like to explicitly incorporate this by making the bilinear map be antisymmetric under the exchange of v^1 and v^2.

$$dx^a \wedge dx^b = \frac{1}{2!}(dx^a \otimes dx^b - dx^b \otimes dx^a) \tag{C.772}$$

$$dS_i = \frac{1}{2!}\epsilon_{ijk} dx^j \wedge dx^k \tag{C.773}$$

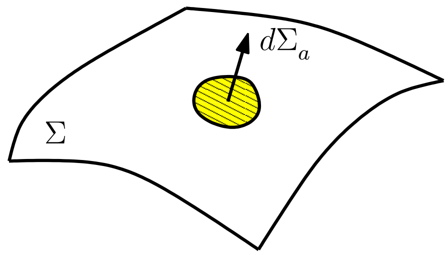


Figure C.38: surfElement.

Definition A $p-$form is defined to be a completely antisymmetric tensor of type $(0, p)$. A one-form is a $(0, 1)$ tensor and a scalar function is a zero form. The number p is the degree of the form.

$\square$

an example of where this notation is employed is in the action of a free electro-dynamical field

$$S = -\frac{1}{4}\int d^4x\sqrt{|g|}F_{\mu\nu}F^{\mu\nu} = -\frac{1}{4}\int F \wedge *F. \tag{C.774}$$

Examples:

1. If $\varphi = \varphi_1 dx^1 + \varphi_2 dx^2$ is a 1-form on an open $V \subset R^2$ then

$$\begin{aligned} d\varphi &= d\varphi_1 \wedge dx^1 + d\varphi_2 \wedge dx^2 \\ &= \left(\frac{\partial \varphi_1}{\partial x^1}dx^1 + \frac{\partial \varphi_1}{\partial x^2}dx^2\right) \wedge dx^1 + \left(\frac{\partial \varphi_2}{\partial x^1}dx^1 + \frac{\partial \varphi_2}{\partial x^2}dx^2\right) \wedge dx^2 \\ &= \left(\frac{\partial \varphi_2}{\partial x^1} - \frac{\partial \varphi_1}{\partial x^2}\right) dx^1 \wedge dx^2. \end{aligned} \tag{C.775}$$

$$\int\int \ldots\ dx^1 \wedge dx^2 \rightarrow \int\int \ldots\ dx^1 dx^2 \tag{C.776}$$

Stoke's theorem gives the well known result (Green's theorem)

$$\int\int_V \left(\frac{\partial\varphi_2}{\partial x^1} - \frac{\partial\varphi_1}{\partial x^2}\right) dx^1 dx^2 = \int_{\partial V} (\varphi_1(x)dx^1 + \varphi_2(x)dx^2). \tag{C.777}$$

2. The 1-form

$$\omega = \frac{xdy - ydx}{x^2 + y^2} \tag{C.778}$$

on $R^2 - \{0\}$. This will crop up when we consider Chern-Simons theory with magnetic source and in the black hole entropy calculation as the surface states on the horizon.

We have

$$\begin{aligned} \frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) &= \frac{1}{x^2+y^2} - \frac{2x^2}{x^2+y^2} = \frac{y^2-x^2}{x^2+y^2}, \\ \frac{\partial}{\partial y}\left(\frac{x}{x^2+y^2}\right) &= -\frac{1}{x^2+y^2} + \frac{2y^2}{x^2+y^2} = \frac{y^2-x^2}{x^2+y^2}. \end{aligned} \tag{C.779}$$

Therefore ω on $R^2 - \{0\}$ is closed.

ω is not exact because if $\alpha(t) = (\cos t \sin t)$ is the unit circle about 0 $(0 \leq t \leq 2\pi)$ then

$$\int_\alpha \omega = \int_0^{2\pi} \frac{\cos t \cdot \sin t - \sin \cdot(-\sin t)}{\cos^2 t + \sin^2 t} dt = \int_0^{2\pi} dt\ 1 = 2\pi \neq 0. \tag{C.780}$$

$$f(b) - f(a) = \int_{[a,b]} df. \tag{C.781}$$

$$\begin{aligned} \alpha &= \omega_1 \wedge \omega_2 \\ &= A_i^{(1)} dx^i \wedge A_i^{(2)} dx^j \\ &= A_i^{(1)} \wedge A_i^{(2)} dx^i \wedge dx^j \\ &= \frac{1}{2}[A_i^{(1)} A_i^{(2)} - A_i^{(2)} A_i^{(1)}] dx^i \wedge dx^j. \end{aligned} \tag{C.782}$$

A differential form of degree can be written

$$\alpha = A_{i_1 i_2 \dots i_p} dx^{i_1} \wedge dx^{i_2} \wedge dx^{i_p} \tag{C.783}$$

where $A_{i_1 i_2 \dots i_p}$ is an antisymmetric tensor.

If we change coordinate system in $\mathbb{R}^n$, we have

$$\omega = A_i dx^i = A_i \frac{\partial x^i}{\partial x'^j} \tag{C.784}$$

that

$$A'_i = \frac{\partial x^j}{\partial x^i} A_i$$

$$A'_{j_1 j_2 \ldots j_p} = \det \left[\frac{\partial x^{i_k}}{\partial x'^{j_l}} \right] A_{i_1 i_2 \ldots i_p} \tag{C.785}$$

which is the Jacobian of the transformation.

dual forms

It follows that the differential form

$$\eta = \sqrt{|g|} dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n, \tag{C.786}$$

retains the same form under coordinate transformations.

C.15.1 Exterior Calculus

If α is a $p-$form and β is a $q-$form, the exterior product $\alpha \wedge \beta$ is a $(p+q)-$form, which is given in local coordinates by

$$\begin{aligned} \alpha &= \frac{1}{p!} \alpha_{\mu_1 \ldots \mu_p} \\ \beta &= \frac{1}{q!} \beta_{\mu_1 \ldots \mu_q} \\ \alpha \wedge \beta &= \frac{1}{p!} \frac{1}{q!} \alpha_{[\mu_1 \ldots \mu_p} \beta_{\mu_{p+1} \ldots \mu_{p+q}]} \end{aligned} \tag{C.787}$$

C.15.2 Exterior Derivatives

Just as the wedge product gave us the cross-product in three dimensions, the exterior derivative gives us the curl. Consider a vector A. The exterior derivative of its associated 1-form

$$\begin{aligned}
dA &= d(A_1 dx^1 + A_2 dx^2 + A_3 dx^3) \\
&= (\frac{\partial A_1}{\partial x^2} - \frac{\partial A_2}{\partial x^1})dx^2 \wedge dx^1 + (\frac{\partial A_2}{\partial x^3} - \frac{\partial A_3}{\partial x^2})dx^3 \wedge dx^2 + (\frac{\partial A_3}{\partial x^1} - \frac{\partial A_1}{\partial x^3})dx^1 \wedge dx^3.
\end{aligned} \tag{C.788}$$

$$\omega = f dx \wedge dy + g dy \wedge dz + h dz \wedge dx \tag{C.789}$$

$$d\omega = \left(\frac{\partial g}{\partial x} + \frac{\partial h}{\partial y} + \frac{\partial f}{\partial z}\right) dx \wedge dy + \wedge dz \tag{C.790}$$

if $\mathbf{V} = (g, h, f)$, the expression $\partial g/\partial x + \partial h/\partial y + \partial f/\partial z$ is $\nabla \cdot \mathbf{V}$

$$d[d(Adx + Bdy + Cdz)] = d^2(Adx + Bdy + Cdz) = 0 \tag{C.791}$$

in component form is

$$\nabla \cdot (\nabla \times \mathbf{V}) = 0. \tag{C.792}$$

The exterior derivative is, denoted d, derivative operation on differential forms, which takes $p-$forms to the $(p+1)-$forms ($d : \Omega^p \to \Omega^{p+1}$).

Take ω in local coordinates

$$\omega = \frac{1}{p!}\omega_{\mu_1 \ldots \mu_r}\, dx^{\mu_1} \wedge \cdots \wedge\, dx^{\mu_p}, \tag{C.793}$$

the derivative operator is then

$$\begin{aligned}
d\omega &= \frac{1}{p!}\partial_\nu \omega_{\mu_1 \ldots \mu_r} d\nu \wedge\, dx^{\mu_1} \wedge \cdots \wedge\, dx^{\mu_p} \\
&= \frac{1}{(p+1)!}
\end{aligned} \tag{C.794}$$

Theorem C.15.1 $d^2 = 0$

The action of d^2 on ω is

$$d^2\omega = \frac{1}{r!}\frac{\partial^2 \omega_{\mu_1 \ldots \mu_r}}{\partial x^\lambda \partial x^\nu}\, dx^\lambda \wedge\, dx^\nu \wedge\, dx^{\mu_1} \wedge \cdots \wedge\, dx^{\mu_r}. \tag{C.795}$$

This vanishes identically because $\partial^2\omega_{\mu_1\ldots\mu_r}/\partial x^\lambda\, dx^\nu$ is symmetric with respect to λ and ν while $dx^\lambda \wedge dx^\nu$ is anti-symmetric.

Theorem C.15.2 *If* $\omega \in \Omega^p(\mathcal{M})$ *then* $d\omega \in \Omega^{p+1}(\mathcal{M})$

Proof.

Definition: A differential form $\omega \in \Omega^r(X)$ is called

(i) *closed* if $d\omega = 0$,
(ii) *exact* if $\omega = d\eta$ for some $\eta \in \Omega^{r-1}(X)$.

Note that

$$d\omega(X,Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]). \tag{C.796}$$

Proof.

$$\begin{aligned} X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]) &= X(\omega_j Y^j) - Y(\omega_j X^j) - \omega_j(X(Y^j) - Y(X^j)) \\ &= X^i\partial_i(\omega_j Y^j) - Y^i\partial_i(\omega_j X^j) - \omega_j(X^i\partial_i Y^j - Y^i\partial_i X^j) \\ &= X^i(\partial_i\omega_j)Y^j - Y^i(\partial_i\omega_j)X^j \\ &= \partial_i\omega_j(X^iY^j - Y^iX^j) \\ &= \frac{1}{2}(\partial_i\omega_j - \partial_j\omega_i)X^iY^j \\ &= d\omega(X,Y) \end{aligned} \tag{C.797}$$

Interior product

$$i_{e_x}(dx\wedge dy) = dy, \quad i_{e_x}(dy\wedge dz) = 0, \quad i_{e_x}(dz\wedge dx) = -dz. \tag{C.798}$$

$$i_X\omega(X_1,\ldots,X_{r-1}) := \omega(X,X_1,\ldots,X_{r-1}). \tag{C.799}$$

For $X = X^\mu\partial/\partial x^\mu$ and $\omega = (1/3!)\,\omega_{\mu_1\mu_2\mu_3}dx^{\mu_1}\wedge dx^{\mu_2}\wedge dx^{\mu_3}$, we have

$$\begin{aligned} i_X\omega &= \frac{1}{2!}X^\nu\omega_{\nu\mu_2\mu_3}dx^{\mu 2}\wedge dx^{\mu_3} \\ &= \frac{1}{3!}\big(X^\nu\omega_{\nu\mu_2\mu_3}dx^{\mu_2}\wedge dx^{\mu_3} - X^\nu\omega_{\mu_1\nu\mu_3}dx^{\mu_1}\wedge dx^{\mu_3} + X^\nu\omega_{\mu_1\mu_2\nu}dx^{\mu_1}\wedge dx^{\mu_2}\big) \end{aligned} \tag{C.800}$$

For $X = X^\mu\partial/\partial x^\mu$ and $\omega = (1/r!)\,\omega_{\mu_1\ldots\mu_r}dx^{\mu_1}\wedge\cdots\wedge dx^{\mu_r}$, we have

$$
\begin{aligned}
i_X\omega &= \frac{1}{(r-1)!} X^\nu \omega_{\nu\mu_2\ldots\mu_r} dx^{\mu_2} \wedge \cdots \wedge dx^{\mu_r} \\
&= \frac{1}{r!}\sum_{s=1}^{r} X^{\mu_s}\omega_{\mu_1\ldots\mu_s\ldots\mu_r}(-1)^{s-1} dx^{\mu_2} \wedge \cdots \wedge d\hat{x}^{\mu_s} \wedge \cdots \wedge dx^{\mu_r} \qquad \text{(C.801)}
\end{aligned}
$$

Now we prove $\mathrm{i}_X(\mathrm{i}_X\omega) = 0$

$$
\begin{aligned}
\mathrm{i}_X(\mathrm{i}_X\omega) &= \frac{1}{(r-1)!}\mathrm{i}_X\big(X^{\nu_1}\omega_{\nu_1\mu_2\ldots\mu_r}\, dx^{\mu_2} \wedge \cdots \wedge dx^{\mu_r}\big) \\
&= \frac{1}{(r-1)!(r-2)!} X^{\nu_1} X^{\nu_2}\omega_{\nu_1\nu_2\ldots\mu_r}\, dx^{\mu_3} \wedge \cdots \wedge dx^{\mu_r} \\
&= 0. \qquad \text{(C.802)}
\end{aligned}
$$

□

$\mathrm{i}_X(\omega \wedge \eta) = \mathrm{i}_X\omega \wedge \eta + (-1)^r \omega \wedge \mathrm{i}_X\eta.$

Proof.

$$
\mathrm{i}_X\omega = \frac{X^\nu\omega_{\nu\mu_2\ldots\mu_r}}{(r-1)!} dx^{\mu_2} \wedge \cdots \wedge dx^{\mu_r}, \qquad \mathrm{i}_X\eta = \frac{X^\nu\eta_{\nu\mu_{r+2}\ldots\mu_q}}{(q-1)!} dx^{\mu_{r+2}} \cdots \wedge dx^{\mu_{r+q}}.
$$

$$
\begin{aligned}
& (-1)^r\omega \wedge \mathrm{i}_X\eta = \\
=\ & (-1)^r\omega \wedge \frac{X^\nu\eta_{\nu\mu_{r+2}\ldots\mu_{r+q}}}{(q-1)!} dx^{\mu_{r+2}} \cdots \wedge dx^{\mu_{r+q}} \\
=\ & (-1)^r \frac{X^\nu\omega_{\mu_1\ldots\mu_r}\eta_{\nu\mu_{r+2}\ldots\mu_{r+q}}}{r!(q-1)!} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r} \wedge dx^{\mu_{r+2}} \cdots \wedge dx^{\mu_{r+q}} \\
=\ & \frac{X^\nu\omega_{\nu\mu_1\ldots\mu_{r-1}}\eta_{\mu_r\mu_{r+2}\ldots\mu_{r+q}}}{r!(q-1)!} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r} \wedge dx^{\mu_{r+2}} \cdots \wedge dx^{\mu_{r+q}} \qquad \text{(C.803)}
\end{aligned}
$$

We then replace dummy indices: $\mu_1 \to \mu_2$, $\mu_1 \to \mu_2$, $\ldots, \mu_r \to \mu_{r+1}$, and then we can write

$$
\begin{aligned}
\mathrm{i}_X\omega \wedge \eta + (-1)^r\omega \wedge \mathrm{i}_X\eta &= X^\nu\omega_{\nu\mu_2\ldots\mu_r}\eta_{\mu_{r+1}\ldots\mu_q}\Big(\frac{1}{(r-1)!q!} + \frac{1}{r!(q-1)!}\Big) \\
&\quad dx^{\mu_2} \wedge \cdots \wedge dx^{\mu_{r+q}} \\
&= (r+q)\frac{X^\nu\omega_{\nu\mu_2\ldots\mu_r}\eta_{\mu_{r+1}\ldots\mu_q}}{r!q!} dx^{\mu_2} \wedge \cdots \wedge dx^{\mu_{r+q}} \\
&= \mathrm{i}_X(\omega \wedge \eta) \qquad \text{(C.804)}
\end{aligned}
$$

□

Lie Derivative of Forms

$$
\begin{aligned}
(di_X + i_X d)\omega &= d\,(X^\mu \omega_\mu) + i_X[\frac{1}{2}(\partial_\mu \omega_\nu - \partial_\nu \omega_\mu) dx^\mu \wedge dx^\nu] \\
&= (\omega_\mu \partial_\nu X^\mu + X^\mu \omega_\nu \partial_\mu) dx^\nu + X^\mu (\partial_\mu \omega_\nu - \partial_\nu \omega_\mu) dx^\nu \\
&= (\omega_\mu \partial_\nu X^\mu + X^\mu \partial_\mu \omega_\nu) dx^\nu. \qquad \text{(C.805)}
\end{aligned}
$$

Comparing this with (C.701), we find that

$$
\mathcal{L}_X = (di_X + i_X d)\omega. \qquad \text{(C.806)}
$$

$$
\begin{aligned}
\mathcal{L}_X \omega &= X^\nu \frac{1}{r!} \partial_\nu \omega_{\mu_1 \ldots \mu_r} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r} + \\
&+ \sum_{s=1}^{r} \partial_{\mu_s} X^\nu \frac{1}{r!} \omega_{\mu_1 \ldots \mu_{s-1} \nu \mu_{s+1} \ldots \mu_r} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r} \qquad \text{(C.807)}
\end{aligned}
$$

$$
\begin{aligned}
d\mathrm{i}_X \omega &= \frac{1}{r!}\, d\big(\sum_{s=1}^{r} X^{\mu_s} \omega_{\mu_1 \ldots \mu_s \ldots \mu_r} (-1)^{s-1} dx^{\mu_1} \wedge \cdots \wedge d\hat{x}^{\mu_s} \wedge \cdots \wedge dx^{\mu_r}\big) \\
&= \frac{1}{r!} \sum_{s=1}^{r} [\partial_\nu X^{\mu_s}\ \omega_{\mu_1 \ldots \mu_s \ldots \mu_r} + X^{\mu_s}\ \partial_\nu \omega_{\mu_1 \ldots \mu_s \ldots \mu_r}] \\
&\quad (-1)^{s-1}\ dx^\nu \wedge dx^{\mu_1} \wedge \cdots \wedge d\hat{x}^{\mu_s} \wedge \cdots \wedge dx^{\mu_r} \qquad \text{(C.808)}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{i}_X d\omega &= \frac{1}{r!} i_X \big(\partial_\nu \omega_{\mu_1 \ldots \mu_r}\ dx^\nu \wedge dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r}\big) \\
&= \frac{1}{r!} \big[X^\nu \partial_\nu \omega_{\mu_1 \ldots \mu_r}\ dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r} \\
&+ \sum_{s=1}^{r} X^{\mu_s} \partial_\nu (-1)^s dx^{\mu_1} \wedge \cdots \wedge d\hat{x}^{\mu_s} \wedge \cdots \wedge dx^{\mu_r}\big]. \qquad \text{(C.809)}
\end{aligned}
$$

$$
\begin{aligned}
&(d\mathrm{i}_X + \mathrm{i}_X d)\omega \\
= \ &\frac{1}{r!} \sum_{s=1}^{r} \partial_\nu X^{\mu_s} \omega_{\mu_1 \ldots \mu_s \ldots \mu_r} (-1)^{s-1}\ dx^\nu \wedge dx^{\mu_1} \wedge \cdots \wedge d\hat{x}^{\mu_s} \wedge \cdots \wedge dx^{\mu_r} \\
+ \ &\frac{1}{r!} X^\nu \partial_\nu \omega_{\mu_1 \ldots \mu_r} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r} \qquad \text{(C.810)}
\end{aligned}
$$

We use

$$
dx^\nu \wedge dx^{\mu_1} \wedge \cdots \wedge d\hat{x}^{\mu_s} \wedge \cdots \wedge dx^{\mu_r} = (-1)^{s-1} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_{s-1}} \wedge dx^\nu \wedge \cdots \wedge dx^{\mu_r}
$$

in the last equation and then swap the indices ν and μ_s around to give (C.807).

$\square$

$\mathrm{i}_X\mathcal{L}_X\omega = \mathcal{L}_X\mathrm{i}_X\omega$

$$\mathrm{i}_X\mathcal{L}_X\omega = \mathrm{i}_X(d\mathrm{i}_X + \mathrm{i}_X d)\omega = \mathrm{i}_X d\mathrm{i}_X\omega = (d\mathrm{i}_X + \mathrm{i}_X d)\mathrm{i}_X\omega = \mathcal{L}_X\mathrm{i}_X\omega. \tag{C.811}$$

$\square$

Pullback of a $p-$form

Consider a map from a manifold $\mathcal{M}$ to $\mathcal{N}$, $f : \mathcal{M} \to \mathcal{N}$.

$$f^*\omega = \frac{1}{p!}\omega_{\nu_1\ldots\nu_p}(f(x))\frac{\partial f^{\nu_1}}{\partial x^{\mu_1}}\cdots\frac{\partial f^{\nu_p}}{\partial x^{\mu_p}}\; dx^{\mu_1}\wedge\cdots\wedge dx^{\mu}_p \tag{C.812}$$

Theorem C.15.3 *The exterior derivative and the pull back commute: $d(f^*\omega) = f^*d\omega$.*

Proof.

$$\begin{aligned} d(f^*\omega) &= \frac{1}{p!}\partial_\lambda\left(\omega_{\mu_1\ldots\nu_p}(f(x))\frac{\partial f^{\nu_1}}{\partial x^{\mu_1}}\cdots\frac{\partial f^{\nu_p}}{\partial x^{\mu_p}}\right)\; dx^\lambda\wedge dx^{\mu_1}\wedge\cdots\wedge dx^{\mu}_p \\ &= \frac{1}{p!}\frac{\partial f^\rho}{\partial x^\lambda}\partial_\rho\omega_{\nu_1\ldots\nu_p}\;\frac{\partial f^{\nu_1}}{\partial x^{\mu_1}}\cdots\frac{\partial f^{\nu_p}}{\partial x^{\mu_p}}\; dx^\lambda\wedge dx^{\mu_1}\wedge\cdots\wedge dx^{\mu}_p \\ &\quad +\frac{1}{p!}\omega_{\nu_1\ldots\nu_p}\;\frac{\partial^2 f^{\nu_1}}{\partial x^\lambda\partial x^{\mu_1}}\cdots\frac{\partial f^{\nu_p}}{\partial x^{\mu_p}}\; dx^\lambda\wedge dx^{\mu_1}\wedge\cdots\wedge dx^{\mu}_p+\ldots \\ &= \frac{1}{p!}\frac{\partial f^\rho}{\partial x^\lambda}\partial_\rho\omega_{\nu_1\ldots\nu_p}\;\frac{\partial f^{\nu_1}}{\partial x^{\mu_1}}\cdots\frac{\partial f^{\nu_p}}{\partial x^{\mu_p}}\; dx^\lambda\wedge dx^{\mu_1}\wedge\cdots\wedge dx^{\mu}_p \\ &= f^*d\omega \end{aligned} \tag{C.813}$$

The terms of third line vanish because the second partial derivatives are symmetric and so vanish, only the term of the second line survives. It means that the exterior derivative of a differential form is independent of the coordinate system in which it is computed.

$\square$

The volume form and the Hodge Dual

The volume

$$V = \frac{1}{p!}\epsilon_{\mu_1\ldots\mu_p}\sqrt{|g|}\,dx^{\mu_1}\wedge\cdots\wedge\,dx^{\mu_m} \tag{C.814}$$

$$\begin{aligned}
\omega\wedge *\eta &= (\omega_{\mu_1\mu_2}\,dx^{\mu_1}\wedge\,dx^{\mu_2})\wedge *\,(\eta_{\nu_1\nu_2}\,dx^{\nu_1}\wedge\,dx^{\nu_2})\\
&= \frac{1}{(2!)^2}\omega_{\mu_1\mu_2}\eta_{\nu_1\nu_2}\sqrt{|g|}\epsilon^{\nu_1\nu_2}{}_{\mu_3\mu_4}\,dx^{\mu_1}\wedge\,dx^{\mu_2}\wedge\,dx^{\mu_3}\wedge\,dx^{\mu_4}\\
&= \frac{1}{2!}\sum_{\mu\nu}\omega_{\mu_1\mu_2}\eta^{\nu_1\nu_2}\frac{1}{2!}\epsilon_{\nu_1\nu_2\mu_3\mu_4}\\
&\quad \epsilon_{\mu_1\mu_2\mu_3\mu_4}\times\sqrt{|g|}\,dx^1\wedge\,dx^2\wedge\,dx^3\wedge\,dx^4\\
&= \frac{1}{2!}\omega_{\mu\nu}\eta^{\mu\nu}\sqrt{|g|}\,dx^1\wedge\,dx^2\wedge\,dx^3\wedge\,dx^4
\end{aligned} \tag{C.815}$$

$$\int\omega\wedge{}^*\eta = \frac{1}{2!}\int_{\mathcal{M}}\omega_{\mu_1\mu_2}\eta_{\mu_3\mu_4}\sqrt{|g|}dx^1dx^2dx^3dx^4 \tag{C.816}$$

$$\begin{aligned}
\omega\wedge *\eta &= \frac{1}{(r!)^2}\omega_{\mu_1\ldots\mu_r}\eta_{\nu_1\ldots\nu_r}\frac{\sqrt{|g|}}{(m-r)!}\epsilon^{\nu_1\ldots\nu_r}{}_{\mu_{r+1}\ldots\mu_m}\\
&\quad\times dx^{\mu_1}\wedge\cdots\wedge\,dx^{\mu_r}\wedge\,dx^{\mu_{r+1}}\wedge\cdots\wedge dx^{\mu_m}\\
&= \frac{1}{r!}\sum_{\mu\nu}\omega_{\mu_1\ldots\mu_r}\eta^{\nu_1\ldots\nu_r}\frac{1}{r!(m-r)!}\epsilon_{\nu_1\ldots\nu_r\mu_{r+1}\ldots\mu_m}\\
&\quad\epsilon_{\mu_1\ldots\mu_r\mu_{r+1}\ldots\mu_m}\times\sqrt{|g|}\,dx^1\wedge\cdots\wedge\,dx^m\\
&= \frac{1}{r!}\omega_{\mu_1\ldots\mu_r}\eta^{\mu_1\ldots\mu_r}\sqrt{|g|}\,dx^1\wedge\cdots\wedge\,dx^m.
\end{aligned} \tag{C.817}$$

C.15.3 Maxwell's equations in differential forms

$$\int_S \mathbf{E}\cdot d\mathbf{A} = \frac{Q}{\epsilon_0} \tag{C.818}$$

$$\int_S \mathbf{B}\cdot d\mathbf{A} = 0 \tag{C.819}$$

$$\oint \mathbf{E}\cdot dl = -\frac{\partial\Phi_B}{\partial t} \tag{C.820}$$

$$\oint \mathbf{B}\cdot dl = \mu_0\left(I+\frac{\partial\Phi_E}{\partial t}\right) \tag{C.821}$$

$$\tag{C.822}$$

Reminder

source equations

$$\begin{aligned} \nabla \cdot E &= \rho && \text{(C.823)} \\ \nabla \times B - \frac{\partial E}{\partial t} &= j && \text{(C.824)} \end{aligned}$$

internal equations

$$\begin{aligned} \nabla \cdot B &= 0 && \text{(C.825)} \\ \nabla \times E - \frac{\partial B}{\partial t} &= 0 && \text{(C.826)} \end{aligned}$$

ρ is the charge density and j the current density.

Gauss's law relates the flux through a closed surface to the enclosed charge

$$\nabla \cdot \frac{\partial E}{\partial t} = \frac{\partial \rho}{\partial t}$$

and

$$-\nabla \cdot \frac{\partial E}{\partial t} = \nabla \cdot j$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot j = 0 \tag{C.827}$$

$$F^{ab} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix} \tag{C.828}$$

source

$$j^a = (\rho, j) \tag{C.829}$$

source and internal equations can be written

Consider the following forms

$$
\begin{aligned}
\mathcal{F} &= \frac{1}{2!} F_{\mu\nu} dx^\mu \wedge dx^\nu \\
\mathcal{J} &= J_\mu dx^\mu,
\end{aligned}
\tag{C.830}
$$

As $\mathcal{F}$ is a form of order 2, is a form of order 1:

$$
\mathcal{A} = A_\mu dx^\mu, \tag{C.831}
$$

Applying the exterior derivative to this we find

$$
\begin{aligned}
d\mathcal{A} &= \partial_\mu A_\nu dx^\mu \wedge dx^\nu \\
&= \frac{1}{2}[\partial_\mu A_\nu - \partial_\nu A_\mu] dx^\mu \wedge dx^\nu \\
&= \frac{1}{2!} F_{\mu\nu} dx^\mu \wedge dx^\nu \\
&= \mathcal{F}.
\end{aligned}
\tag{C.832}
$$

The field strength tensor is thus the exterior derivative of the one-form potential:

$$
\mathcal{F} = d\mathcal{A}. \tag{C.833}
$$

$$
\begin{aligned}
\mathcal{F} &= F_{0i} dx^0 \wedge dx^i + \frac{1}{2} F_{ij} dx^i \wedge dx^j \\
&= E_i dx^0 \wedge dx^i + \frac{1}{2} \epsilon_{ijk} B_k dx^i \wedge dx^j
\end{aligned}
\tag{C.834}
$$

Consider two gauge potentials $A_\mu(x)$ and $A'_\mu(x)$ related by a gauge transformation, i.e. $A_\mu(x) = A'_\mu(x) + \partial_\mu \phi(x)$. Written as a differential form this is

$$
\mathcal{A} - \mathcal{A}' = d\phi, \tag{C.835}
$$

Taking the exterior derivative of this we find

$$
d\mathcal{A} - d\mathcal{A}' = \mathcal{F} - \mathcal{F}' = d^2\phi = 0, \tag{C.836}
$$

Let us apply the operation d to the form $\mathcal{F}$,

$$
d\mathcal{F} = \frac{1}{2!} \partial_{[\lambda} F_{\mu\nu]} dx^\lambda \wedge dx^\mu \wedge dx^\nu = 0 \tag{C.837}
$$

$$\partial_{[\lambda}F_{\mu\nu]} = 0. \tag{C.838}$$

This is the so-called Bianchi identity and is identical to the source free maxwell equations.

The Hodge dual of $\mathcal{F}$

we find

$$^*\mathcal{F} = \frac{1}{2}\epsilon_{\alpha\beta\mu\nu}F^{\alpha\beta}dx^\mu \wedge dx^\nu \tag{C.839}$$

$$^*\mathcal{F} = B_i dx^0 \wedge dx^i + \frac{1}{2}\epsilon_{ijk}E_k dx^i \wedge dx^j \tag{C.840}$$

$$^*\mathcal{J} = J^0 d^3x - \frac{1}{2}\epsilon_{ijk}J_k dx^0 \wedge dx^i \wedge dx^j \tag{C.841}$$

Maxwell's equations are then

$$\begin{aligned} \mathcal{F} &= d\mathcal{A} \\ d^*\mathcal{F} &= 4\pi\ ^*\mathcal{J}\ . \end{aligned} \tag{C.842}$$

we get

$$\begin{aligned} d\mathcal{F} &= 0 \\ d^*\mathcal{F} &= 4\pi\ ^*\mathcal{J}. \end{aligned} \tag{C.843}$$

Thus the form $\mathcal{F}$ is *closed*. It is therefore locally *exact* : there exists a form $\mathcal{A}$ such that

The action for the Maxwell theory

$$S = -\frac{1}{2}\int_{\mathcal{M}} F^{\alpha\beta}F_{\alpha\beta}d^4x + \int_{\mathcal{M}} A_\alpha J^\alpha d^4x \tag{C.844}$$

From (C.815) we have

$$\mathcal{F} \wedge {}^*\mathcal{F} = F^{\alpha\beta}F_{\alpha\beta}\ d^4x \tag{C.845}$$

$$\mathcal{A} \wedge {}^*\mathcal{J} = A_\alpha J^\alpha\ d^4x \tag{C.846}$$

The action for the Maxwell theory can be written as

$$S = -\frac{1}{2}\int_{\mathcal{M}} \mathcal{F}\wedge {}^*\mathcal{F} + \int_{\mathcal{M}} \mathcal{A}\wedge {}^*\mathcal{J} \tag{C.847}$$

Holonomy-flux variables

regularization is natural geometrically

Stoke's Theorem

$$f(b) - f(a) = \int_a^b \frac{df}{dx}(x)\, dx. \tag{C.848}$$

In the language of exterior calculus this is written as

$$f(b) - f(a) = \int_{[a,b]} df. \tag{C.849}$$

$$\oint_C Pdx + Qdy = \int_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) ds \tag{C.850}$$

There is a more general result that includes Stokes theorem, Green's theorem, Gauss' theorem and the divergence theorem. If $\mathcal{M}$ is an oriented manifold of dimension n with boundary $\partial\mathcal{M}$ and ω is an $n-1$ form then Stoke's theorem says that

$$\int_X d\omega = \int_{\partial X} \imath^*\omega, \tag{C.851}$$

$$\int_{\partial D} (P\, dy\, dz + Q\, dz\, dx + R\, dx\, dy) = \int_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dx\, dy\, dz. \tag{C.852}$$

Again the integrands are related through exterior differentiation. Let φ be a 2-form

$$\varphi = P\, dy \wedge\, dz + Q\, dz \wedge\, dx + R\, dx \wedge\, dy. \tag{C.853}$$

Then () can be written

$$\int_{\partial D} \varphi = \int_D d\varphi. \tag{C.854}$$

In familiar vector calculus notation, Gauss' Theorem appears as

$$\int_{\partial D} \mathbf{A} \cdot da = \int_D d^3 r \nabla \cdot \mathbf{A}, \tag{C.855}$$

where $\mathbf{A}$ is the vector field

$$\mathbf{A} = P\mathbf{e_x} + Q\mathbf{e_y} + R\mathbf{e_z}, \tag{C.856}$$

and $d\mathbf{a}$ area element in ∂D along the outward normal.

Example.

$$\begin{aligned} \varphi &= \frac{1}{2}\sum_{i,j,k=1}^{3} \epsilon_{ijk}\varphi_i dx^j \wedge dx^k, \\ &= \varphi_1 dx^2 \wedge dx^3 + \varphi_2 dx^3 \wedge dx^1 + \varphi_3 dx^1 \wedge dx^2 \end{aligned} \tag{C.857}$$

we find

$$d\varphi \sum_{i=1}^{3} \frac{\partial \varphi_i}{\partial x^i} dx^1 \wedge dx^2 \wedge dx^3 \tag{C.858}$$

and Stoke's theorem tells us:

$$\begin{aligned} & \int_V \left(\frac{\partial \varphi_1}{\partial x^1} + \frac{\partial \varphi_1}{\partial x^2} + \frac{\partial \varphi_1}{\partial x^3} \right) dx^1 \wedge dx^2 \wedge dx^3 \\ = & \int_{\partial V} (\varphi_1(x) dx^2 \wedge dx^3 + \varphi_2(x) dx^3 \wedge dx^1 + \varphi_3(x) dx^1 \wedge dx^2). \end{aligned} \tag{C.859}$$

Example.

$$\varphi = \varphi_1 dx^1 + \varphi_2 dx^2 + \varphi_3 dx^3$$

$$\begin{aligned} d\varphi &= \left(\frac{\partial \varphi_3}{\partial x^2} - \frac{\partial \varphi_2}{\partial x^3} \right) dx^2 \wedge dx^3 + \left(\frac{\partial \varphi_1}{\partial x^3} - \frac{\partial \varphi_3}{\partial x^1} \right) dx^3 \wedge dx^1 \\ &+ \left(\frac{\partial \varphi_2}{\partial x^1} - \frac{\partial \varphi_1}{\partial x^2} \right) dx^1 \wedge dx^2, \end{aligned} \tag{C.860}$$

we get

$$\int \left[\left(\frac{\partial \varphi_3}{\partial x^2} - \frac{\partial \varphi_2}{\partial x^3} \right) dx^2 \wedge dx^3 + \left(\frac{\partial \varphi_1}{\partial x^3} - \frac{\partial \varphi_3}{\partial x^1} \right) dx^3 \wedge dx^1 \right.$$
$$+ \quad \left. \left(\frac{\partial \varphi_2}{\partial x^1} - \frac{\partial \varphi_1}{\partial x^2} \right) dx^1 \wedge dx^2 \right] = \int_\partial (\varphi_1 dx^1 + \varphi_2 dx^2 + \varphi_3 dx^3). \tag{C.861}$$

C.15.4 Integration on a Manifold

This generalization can be proved in the same way the usual Stokes' theorem is proved.

$\omega = \omega_{\mu_1} dx^{\mu_1}$ and

$d\omega_{\mu_1\mu_2} dx^{\mu_1} \wedge dx^{\mu_2} = \partial_{[\mu_1}\omega_{\mu_1]} dx^{\mu_1} \wedge dx^{\mu_2}$

$$\begin{aligned}
\int_C d\omega &= \int_0^a \int_0^b d\omega_{\mu_1\mu_2} dx^{\mu_1} dx^{\mu_2} \\
&= \int_0^a \int_0^b \partial_{[\mu_1}\omega_{\mu_2]} dx^{\mu_1} dx^{\mu_2} \\
&= \int_0^a \int_0^b \left(\partial_{\mu_1}\omega_{\mu_2} - \partial_{\mu_2}\omega_{\mu_1}\right) dx^{\mu_1} dx^{\mu_2} \\
&= \int_0^b \Big(\omega_{\mu_2}(a, x^{\mu_2}) - \omega_{\mu_2}(0, x^{\mu_2})\Big) dx^{\mu_2} - \int_0^a \Big(\omega_{\mu_1}(x^{\mu_1}, b) - \omega_{\mu_1}(x^{\mu_1}, 0)\Big) dx^{\mu_1} \\
&= \int_0^a \omega_{\mu_1}(x^{\mu_1}, 0) dx^{\mu_1} + \int_0^b \omega_{\mu_2}(a, x^{\mu_2}) dx^{\mu_2} + \int_a^0 \omega_{\mu_1}(x^{\mu_1}, b) dx^{\mu_1} + \int_b^0 \omega_{\mu_2}(0, x^{\mu_2}) dx^{\mu_2} \\
&= \int_{\partial C} \omega
\end{aligned} \tag{C.862}$$

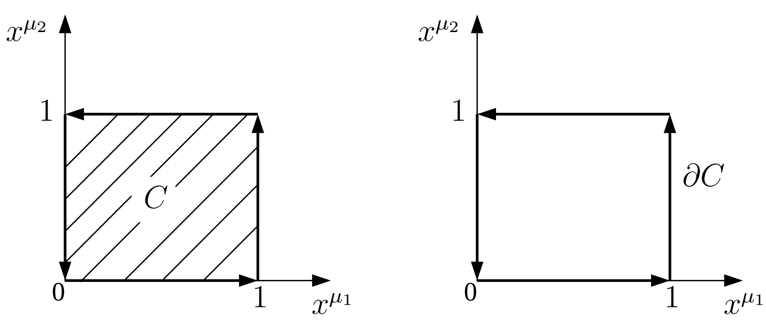


Figure C.39: stokeExam.

Proof in unit cube case: Here

$$M = \{(x_1, \ldots, x_k) : \text{ for each } i, 0 \le x_i \le 1\}.$$

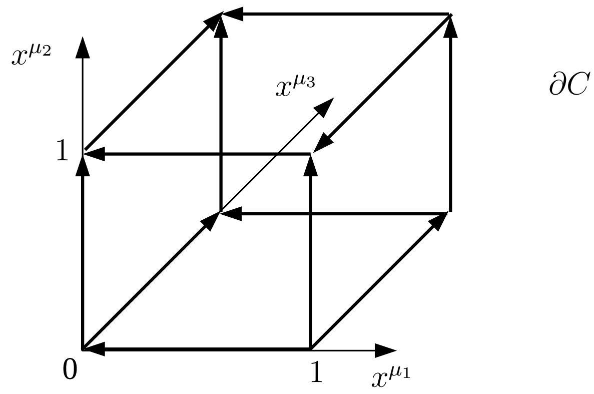


Figure C.40: stokeExam2.

The boundary ∂M of this cube consists of $2k$ unit cubes in $\mathbb{R}^{k-1}$. We will be concerned with the two boundary components

$$S_1 = \{(0, x_2, \ldots, x_k) \in M\}$$

and

$$S_2 = \{(1, x_2, \ldots, x_k) \in M\}.$$

For $\omega = f(x_1, \ldots, x_k)\, dx_2 \wedge \cdots \wedge dx_k$, we have

$$\begin{aligned} d\omega &= \sum \frac{\partial f}{\partial x_i} dx_i \wedge dx_2 \wedge \cdots \wedge dx_k \\ &= \frac{\partial f}{\partial x_1} dx_1 \wedge dx_2 \wedge \cdots \wedge dx_k \end{aligned} \tag{C.863}$$

since $dx_j \wedge dx_j = 0$.

Now integrate $d\omega$ along the unit cube M. We choose our orientation preserving parameterizing map to be the identity map. Then

$$\int_M d\omega = \int_0^1 \cdots \int_0^1 \frac{\partial f}{\partial x_1} dx_1 \ldots dx_k.$$

The integration gives

$$\begin{aligned} \int_M d\omega &= \int_0^1 \cdots \int_0^1 f(1, x_2, \ldots, dx_k) dx_2 \ldots dx_k \\ &\quad - \int_0^1 \cdots \int_0^1 f(0, x_2, \ldots, dx_k) dx_2 \ldots dx_k \end{aligned} \tag{C.864}$$

Now we compare this to the surface integral $\int_{\partial M} \omega$. Since $\omega = f(x_1, \ldots, x_k) dx_2 \wedge \cdots \wedge dx_k$, the only parts of the integral along the boundary that will not be zero will be along S_1 and S_2, both of which are unit cubes in $\mathbb{R}^{k-1}$, with coordinates given by $x_1, \ldots, x_k$. They will have opposite orientations though. (We saw this in the example for when M is a square in the plane; then S_1 is the bottom of the square and S_2 is the top of the square. Note how the orientations on S_1 and S_2 induced from the orientation of the square are indeed opposite).

Then

$$\begin{aligned} \int_{\partial M} \omega &= \int_{C_1} \omega + \int_{C_2} \omega \\ &= \int_0^1 \ldots \int_0^1 -f(0, x_2, \ldots, x_k) dx_2 \ldots dx_k \\ &= \int_0^1 \ldots \int_0^1 f(1, x_2, \ldots, x_k) dx_2 \ldots dx_k, \end{aligned} \tag{C.865}$$

which we have just shown to equal to $\int_M d\omega$, as desired.

□

The boundary of a line segment is the two end points,

$$\partial[p_1, p_2] = [p_1] - [p_2], \tag{C.866}$$

triangle

$$\partial[p_1, p_2, p_3] = [p_1, p_2] + [p_2, p_3] + [p_3, p_1]. \tag{C.867}$$

$$\begin{aligned} \partial\partial[p_1, p_2, p_3] &= \partial[p_1, p_2] + \partial[p_2, p_3] + \partial[p_3, p_1] \\ &= [p_1] - [p_2] + [p_2] - [p_3] + [p_3] - [p_1] \\ &= 0. \end{aligned} \tag{C.868}$$

$$\partial^2 = 0. \tag{C.869}$$

$$\partial[p_1, \ldots, p_{n+1}] = \sum_{i=1}^{n+1} [p_1, \ldots, \hat{p}_i, \ldots, p_{n+1}]. \tag{C.870}$$

Theorem C.15.4 (Stokes' Theorem). *If $\omega \in \Omega^{p-1}(\mathcal{M}\mathbb{R})$ and σ is a $p-$chain then*

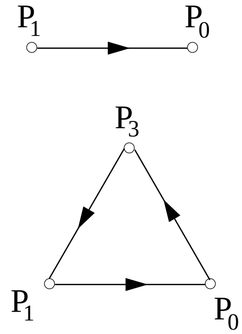


Figure C.41: boundery.

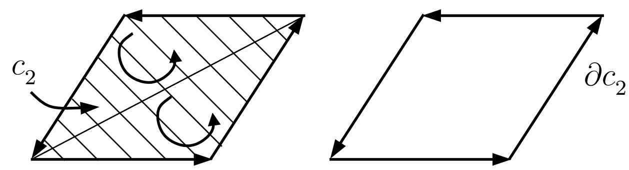


Figure C.42: boundarychain. Boundary of a chain.

$$\int_C d\omega = \int_{I_p} C^* d\omega = ??? \tag{C.871}$$

$$\begin{aligned}
\int_{I_p} d\psi &= (-1)^{p-1} \int_{I_p} \partial_p f \; dx^1 \dots dx^p \\
&= (-1)^{p-1} \int dx^1 \dots dx^{p-1} \int_0^1 \partial_p f \; dx^p \\
&= (-1)^{p-1} \int dx^1 \dots dx^{p-1} \Big(f(x^1, \dots, dx^{p-1}, 1) - f(x^1, \dots, dx^{p-1}, 0) \Big)
\end{aligned} \tag{C.872}$$

On the other hand the boundary of I_p is

$$\begin{aligned}
\partial I_p &= \sum_{i=1}^{p} (-1)^{i+1} \big(\pi_i^{(1)} - \pi_i^{(0)} \big) \\
&= \sum_{i=1}^{p} (-1)^{i+1} \Big(\{[x^1, \dots, 1, x^p) | x^j \in I_p\} - \{[x^1, \dots, 0, x^p) | x^j \in I_p\} \Big)
\end{aligned} \tag{C.873}$$

C.15.5 Cartan Structure Equations

Let $\{\hat{e}_i\}$ be the non-coordinate basis and $\{\hat{\theta}^i\}$ the dual basis. The vector fields satisfy

$$[\hat{e}_i, \hat{e}_j] = c_{ij}{}^k \hat{e}_k. \tag{C.874}$$

The connection coefficients with respect to the basis $\{\hat{e}_i\}$ by

$$\nabla_i \hat{e}_j \equiv \omega^k{}_{ij} \hat{e}_k \tag{C.875}$$

$$\omega^i{}_j := \omega^i{}_{kj} \hat{\theta}^k. \tag{C.876}$$

The 2−forms of torsion are given by the first **Cartan's equations of structure**, which read

$$T^k = d\hat{\theta} + \omega^k{}_i \wedge \hat{\theta}^i, \tag{C.877}$$

and the 2-forms of curvature by the second Cartan's equations of structure

$$\Omega^k{}_i = d\omega^k{}_i + \omega^k{}_j \wedge \omega^j{}_i. \tag{C.878}$$

where $T^i \equiv \frac{1}{2} T^i{}_{jk} \hat{\theta}^j \wedge \hat{\theta}^k$ is the torsion two-form and $R^i{}_j \equiv \frac{1}{2} R^i{}_{jk\ell} \hat{\theta}^k \wedge \hat{\theta}^\ell$ the curvature two-form

Proof. Let the LHS of (C.877) act on the basis vectors $\hat{e}_k$ and $\hat{e}_\ell$,

$$\begin{aligned}
& d\hat{\theta}^i(\hat{e}_k, \hat{e}_\ell) + [< \omega^i{}_j, \hat{e}_k >< \hat{\theta}^j, \hat{e}_\ell > - < \hat{\theta}^j, \hat{e}_k >< \omega^i{}_j, \hat{e}_\ell >] \\
= \; & \{\hat{e}_k[< \hat{\theta}^i, \hat{e}_\ell >] - \hat{e}_k[< \hat{\theta}^i, \hat{e}_k >] - < \hat{\theta}^i, [\hat{e}_k, \hat{e}_\ell] >\} + \{< \omega^i{}_\ell, \hat{e}_k > - < \omega^i{}_k, \hat{e}_\ell >\} \\
= \; & -c_{k\ell}{}^i + \omega^i_{k\ell} - \omega^i_{\ell k} = T^i_{k\ell}
\end{aligned} \tag{C.879}$$

where we have made use of (C.796). The RHS acting on $\hat{e}_k$ and $\hat{e}_\ell$ yields

$$\frac{1}{2} T^i{}_{jm}[< \hat{\theta}^j, \hat{e}_k >< \hat{\theta}^m, \hat{e}_\ell > - < \hat{\theta}^m, \hat{e}_k >< \hat{\theta}^j, \hat{e}_\ell >] = T^i{}_{k\ell}$$

which completes the proof. Equation may be proven similarly.

□

Taking the exterior derivatives of (C.877) and (C.15.5), we have the **Bianchi identities**

$$\begin{aligned}
dT^i + \omega^i{}_j \wedge T^j &= R^i{}_j \wedge \hat{\theta}^j && \text{(C.880)} \\
dR^i{}_j + \omega^i{}_k \wedge R^k{}_j - R^i{}_k \wedge \omega^k{}_j &= 0. && \text{(C.881)}
\end{aligned}$$

These are the non-coordinate basis versions of $R^a{}_{[bcd]} = 0$ and $\nabla_{[e|} R^a{}_{b|cd]} = 0$.

C.15.6 A Differential Geometry Translator

$$\begin{aligned} dA &= (\frac{\partial A_x}{\partial x}dx + \frac{\partial A_y}{\partial x}dy + \frac{\partial A_z}{\partial x}dz) \wedge dx \\ &= \frac{\partial A_y}{\partial x}\, dy \wedge dx + \frac{\partial A_z}{\partial x}\, dz \wedge dx \end{aligned} \tag{C.882}$$

The gravitational field e

$$e^I(x) = e^I_a dx^a \tag{C.883}$$

The spin connection

$$\omega^I{}_J(x) = \omega^I{}_{aJ}(x)dx^a \tag{C.884}$$

$$\mathcal{D}v^I = dv^I + \omega^I{}_J v^J \tag{C.885}$$

$$R^I{}_J = R^I{}_{J\,ab} dx^a \wedge dx^b \tag{C.886}$$

$T(X,Y) = T^a{}_{bc} X^b Y^c$

$$T^a = \frac{1}{2} T^a{}_{bc} \omega^b \wedge \omega^c \tag{C.887}$$

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] \tag{C.888}$$

$R(X,Y)Z = R^a{}_{bcd} X^c Y^d Z^b$

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z + \nabla_{[X,Y]} Z \tag{C.889}$$

Hodge-Star operation:

$$\mathcal{S}[e^I, \omega^{IJ}] = \frac{1}{4\kappa} \int_{\mathcal{M}} \epsilon_{IJKL} e^I_a e^J_b \left(R^{KL}_{\sigma\rho} - \frac{\Lambda}{6} e^K_\sigma e^L_\rho \right) dx^\mu dx^\nu dx^\sigma dx^\rho \tag{C.890}$$

$$\mathcal{S}[e^I, \omega^{IJ}] = \frac{1}{4\kappa} \int_{\mathcal{M}} \epsilon_{IJKL} (e^I \wedge e^J \wedge R[\omega]^{KL} + \frac{\Lambda}{6} e^I \wedge e^J \wedge e^K \wedge e^L) \tag{C.891}$$

C.16 More on Lie groups

For $g \in G$ we define the adjoint isomorphism $ad_g : G \to G$ by

$$ad_g(h) := ghg^{-1} = L_g \circ R_{g^{-1}} h, \tag{C.892}$$

for all $h \in G$.

It is a group homomorphism as

$$\begin{aligned} ad_{[g_1,g_2]}(g_3) &= [[g_1, g_2], g_3] \\ &= [g_1, [g_2, g_3]] + [g_2, [g_3, g_1]] \\ &= ad(g_1)(ad(g_2)(g_3)) - ad(g_2)(ad(g_1)(g_3)) \\ &= [ad(g_1), ad(g_2)](g_3) \end{aligned} \tag{C.893}$$

To see this is an isomorphism write $ad_g(h) = g'$ for any $g' \in G$, this has the solution $h = g^{-1}g'g$. This solution is unique as $ad_g(h) = ad_g(h')$ implies $h = h'$. This is called the adjoint representation of the Lie group G

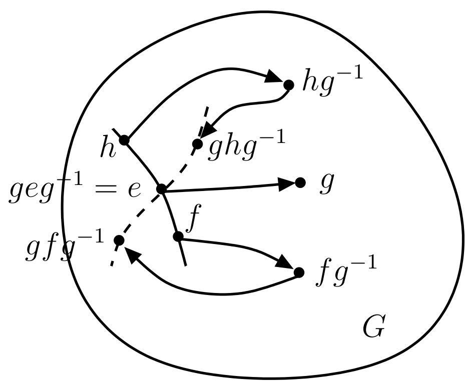


Figure C.43: A curve through e under the map $h \mapsto ghg^{-1}$, first a right action $R_{g^{-1}}$ as $h \mapsto hg^{-1}$ followed by the left action L_g as $hg^{-1} \mapsto ghg^{-1}$. The identity e is mapped to itself but points h and f near it are generally changed, so that a tangent vector at e, in $T_e(G)$, is mapped to another one in $T_e(G)$.

The map ad_g fixes the neutral element e therefore the adjoint isomorphism ad_g on the Lie group G induces an isomorphism on $T_e(G)$

$$\mathcal{A}d_g := (ad_g)_* : T_e(G) \to T_e(G)$$

or

$$\mathcal{A}d_g : \mathcal{G} \to \mathcal{G}$$

on the Lie algebra $\mathcal{G}$.

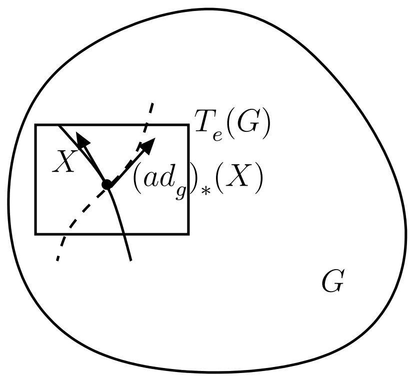


Figure C.44: A vector $X \in T_e(G)$ is mapped to another one in $\mathcal{A}d_g(X) \in T_e(G)$. Written formally as $(ad_g)_* : T_e(G) \to T_e(G)$.

This is a group homomorphism between groups

$$\begin{aligned} \mathcal{A}d(g_1 g_2) &= (ad_{g_1 g_2})_* = (ad_{g_1} \circ ad_{g_2})_* \\ &= (ad_{g_1})_* \circ (ad_{g_2})_* = \mathcal{A}d(g_1) \circ \mathcal{A}d(g_2) \end{aligned} \tag{C.894}$$

Note by injectivety that $\mathcal{A}d(g)$ maps any vector $X \in T_e(G)$ to a non-zero vector for all $g \in G$. Also we have $\mathcal{A}d(g)(X+Y) = \mathcal{A}d(g)(X) + \mathcal{A}d(g)(Y)$. We thus have the map

$$\mathcal{A}d : G \to GL(n, \mathbb{R}), \tag{C.895}$$

where $n = dim\mathcal{G}$.

This homomorphisms in turn induces a homomorphism between Lie algebras

$$Ad := \mathcal{A}d_* : \mathcal{G} \to \mathcal{GL}(n, \mathbb{R}), \quad n = dim\mathcal{G} \tag{C.896}$$

called the adjoint representation of the Lie algebra $\mathcal{G}$.

If G is a matrix group, the adjoint representation becomes a simple matrix operation.

Definition The kernel of a group homomorphism $\varphi : G \to H$ is defined by

$$\ker \varphi := \varphi^{-1}(\{e_H\}) = \{x \in G : \varphi(x) = e_H\}.$$

Proposition C.16.1 *A group homomorphism $\varphi : G \to H$ is injective if and only if its kernel is trivial, i.e., ker $\varphi = \{e_G\}$.*

Proof:

Assume φ is injective. Since we must have $e_G \in \ker \varphi$, $\ker \varphi = \{e_G\}$. Assume $\ker \varphi = \{e_G\}$. Say φ is not injective, i.e., there is $x \neq y$ such that $\varphi(x) = \varphi(y)$. Then $\varphi(xy^{-1}) = e_H$, implying $xy^{-1} = e_G$ or $x = y$.

$\square$

If G is a group, then by an automorphism of G we mean an isomorphism of G onto itself. The collection of such automorphisms, denoted $\mathrm{Aut}(G)$, is a subgroup of Sym (G).

The kernel of ad_x is the subgroup of G consisting of the elements $x \in G$ with the property that $xyx^{-1} = y$ for all $y \in G$, or, equivalently that $xy = yx$ for all $y \in G$. Thus the kernel of I equals the center $Z(G)$ of G.

C.16.1 Discrete Groups

A discrete group is a group with the discrete topology.

For example any finite group is a discrete group.

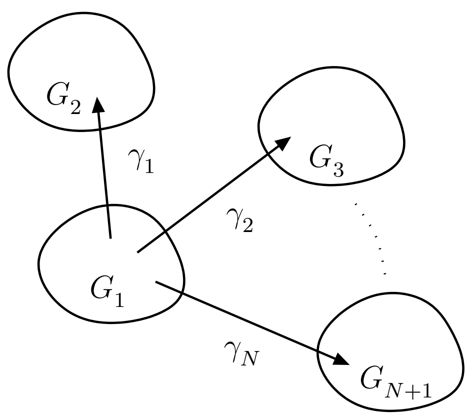


Figure C.45: If a Lie group is a direct product of the proper subgroup and some discrete subgroup then each connected component G_i is obtained from the proper subgroup G_1 by applying some discrete transformation γ_i of a discrete subgroup Γ.

C.16.2 Universal Covering Group

Two curves $g(\tau)$ and $g'(\tau)$ connecting the elements g_0 and g_1 are said to be homotomic if there exists a continuous deformation of one curve into the another, which leaves the end points g_0

and g_1 unaltered, i.e., there exists a continous function $h(\tau, s)$ of two parameters τ and s such that

$$h(0, s) = g_0, \qquad h(1, s) = g_1$$

$$h(\tau, 0) = g(\tau), \qquad h(\tau, 1) = g'(\tau).$$

A Lie group is said to be simply connected if every loop is homotopic to the null loop, i.e., every loop is contractable to one point.

A topological set not able to be partitioned into non-empty open subsets each of which has no points in common with the closure of the other. A topological space X is **connected** if $\emptyset$ and X are the only subsets of X that are both open and closed.

C.17 Group Actions on Sets

Action of a group

$$\sigma(g, x) = y \tag{C.897}$$

$$\sigma(t, x) \tag{C.898}$$

if the flow $\sigma(t, x)$ is periodic with period T.

We can construct a new action whose group is $U(1)$

$$\overline{\sigma}(\exp(2\pi i t/T)y, x) = \sigma(t, x) \tag{C.899}$$

and one whose group is $SO(2)$

$$\tilde{\sigma}(\left(\begin{array}{cc} \cos(2\pi t/T) & \sin(2\pi t/T) \\ -\sin(2\pi t/T) & \cos(2\pi t/T) \end{array}\right), x) = \sigma(t, x). \tag{C.900}$$

The action of $GL(n, \mathbb{R})$ on $\mathbb{R}^n$

$$\sigma(\mathbf{M}, x) = \mathbf{M} \cdot x \tag{C.901}$$

where $\cdot$ is the usual matrix multiplication on a vector. The action of subgroups of $GL(n, \mathbb{R})$ is defined similarly. $O(n)$ acts on $S^{(n-1)}(r)$, an $(n-1)-$sphere of radius r,

$$\sigma : O(n) \times S^{(n-1)}(r) \to S^{(n-1)}(r). \tag{C.902}$$

Formal definition of the action of a group on a manifold:

Definition Let G be a Lie group and $\mathcal{M}$ be a manifold. The action of G on $\mathcal{M}$ is a differentional map $\sigma : G \times \mathcal{M} \to \mathcal{M}$ which satisfies the conditions

(i) $\sigma(e,p) = p \qquad$ for any $p \in \mathcal{M}$

(ii) $\sigma(g_1, \sigma(g_2, p)) = \sigma(g_1 g_2, p)$.

We as well define the **orbit** of a point x of $\mathcal{M}$ as

$$orb(x) = \{gx \mid g \in G\} \tag{C.903}$$

i) The action of the group is *transitive* if any orbit is the whole of X.

ii) The action is *effective*, or *faithful*, if the trivial action on X, i.e.,if $\sigma(g,p) = p$ for all $p \in X$, implies $g = e$.

If the action is not effective, the set of g corresponding to the trivial action is an invariant subgroup H of G, and we can take G/H as having a faithful action.

iii) The action is *free* if the existence of an p such that $gp = p$ implies that $g = p$.

The **stabalizer** of x as

$$Stab(x) = \{g \in G \mid gx = x\} \tag{C.904}$$

The orbits are equivalence classes - we are often interested in the quotient space.

Isotropy group

The identity element e is obviously in $H(p)$. Now, let $g_1, g_2 \in H(p)$

$$\sigma(g_1 g_2, p) = \sigma(g_1, \sigma(g_2, p)) = \sigma(g_1, p) = p$$

$g^{-1} \in H(p)$ because

$$p = \sigma(e,p) = \sigma(g^{-1}g, p) = \sigma(g^{-1}, \sigma(g,p)) = \sigma(g^{-1}, p)$$

One can consider the quotient $G/H(p)$

Example $SO(3)/SO(2)$

we have

$$SO(3)/SO(2) \cong S^2$$

Definition A group action is *effective* if the identity element is the only element that, that is, if $\sigma(g, x) = x$ for all $x \in \mathcal{M}$, then $g = e$.

C.17.1 Transitive Actions

Properties:

i) Orbits are disjoint,

ii) $\mathcal{M}$ is the union of of the orbits.

Example. Let $G = O(n)$ and $\mathcal{M} = S^n$.

Definition A group action is *transitive* if, for any $x_1, x_2 \in \mathcal{M}$, there exists a $g \in G$ such that $\sigma(g, x_1) = x_2$.

There is only one orbit!

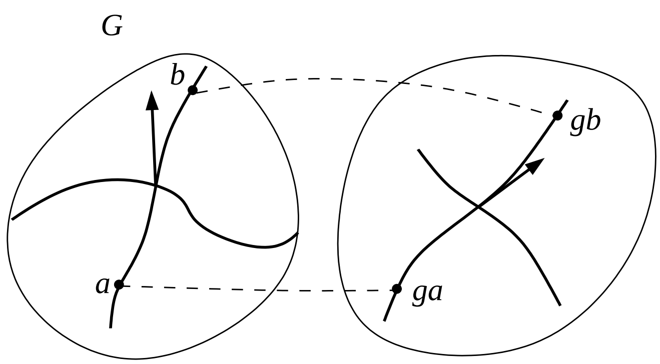


Figure C.46: leftTran. The left translation along g maps a neighbourhood of e onto one of g. There is a natural map of a vector at e to one at g.

C.17.2 Faithful Actions

The action of G on X is said to be faithful if $gx = hx$ for all $x \in X$ implies that $g = h$.

C.17.3 Free Actions

The action of G on X is said to be free if for all $g \in G$, $g \neq e$, and for all $x \in X$, $gx \neq x$.

Definition A group action is *free* if every element apart from the identity of G has no fixed points in M, that is, if there exists an element $x \in M$ such that $\sigma(g, x) = x$, then $g = e$.

C.17.4 Introduction to Gauge Invariance of the Yang-Mills Equations

$$\hat{U} = \exp\left(\frac{i}{2}\mathbf{a} \cdot \tau\right) \tag{C.905}$$

Minimal coupling term

$$\begin{aligned} \mathcal{L}_{int} &= g\sum_{i=1}^{3} \overline{\Psi}\gamma^{\mu}\frac{\tau^i}{2}A_{\mu}^{i}\Psi \\ &= g\overline{\Psi}\gamma^{\mu}\mathbf{A}_{\mu} \cdot \frac{\tau}{2}\Psi \end{aligned} \tag{C.906}$$

$$\Psi \to \Psi' = \hat{U}\Psi = \exp\left(i\mathbf{a}(x) \cdot \hat{\mathbf{T}}\right)\Psi(x) \tag{C.907}$$

$$\begin{aligned} \partial_{\mu}\Psi \to \hat{U}\partial_{\mu}\Psi &= \hat{U}\partial_{\mu}(\hat{U}^{-1}\hat{U}\Psi) \\ &= \partial_{\mu}(\hat{U}\Psi) + \hat{U}(\partial_{\mu}\hat{U}^{-1})\hat{U}\Psi \\ &= \partial_{\mu}\Psi' + \hat{U}(\partial_{\mu}\hat{U}^{-1})\Psi' \\ &= \left[\partial_{\mu} + \hat{U}(\partial_{\mu}\hat{U}^{-1})\right]\Psi' \end{aligned} \tag{C.908}$$

By adding the coupling (C.905) to the free Dirac equation the additional term $\hat{U}(\partial_{\mu}\hat{U}^{-1})$, occurring for $\partial_{\mu}\mathbf{a}$, can be absorbed by gauging the fields A_{μ}^{i} simultaneously. The Dirac Lagrangean is the cast into a gauge invariant form

$$\mathcal{L}(\Psi, \mathbf{A}_{\mu}) = i\overline{\Psi}\gamma^{\mu}\Psi + g\overline{\Psi}\gamma^{\mu}\mathbf{A}_{\mu} \cdot \hat{\mathbf{T}}\Psi \tag{C.909}$$

and the gauged density

$$\mathcal{L}' = i\overline{\Psi}'\gamma^{\mu}\Psi' + g\overline{\Psi}'\gamma^{\mu}\mathbf{A}'_{\mu} \cdot \hat{\mathbf{T}}\Psi' \tag{C.910}$$

$$
\begin{aligned}
\mathcal{L} &= i\overline{\Psi}\gamma^\mu\Psi + g\overline{\Psi}\gamma^\mu\mathbf{A}_\mu\cdot\hat{\mathbf{T}}\Psi \\
&= i\overline{\Psi}\hat{U}^{-1}\hat{U}\gamma^\mu\partial_\mu(\hat{U}^{-1}\hat{U}\Psi) + g\overline{\Psi}\hat{U}^{-1}\hat{U}\gamma^\mu\mathbf{A}_\mu\cdot\hat{\mathbf{T}}\hat{U}^{-1}\hat{U}\Psi \\
&= i\overline{\Psi}'\hat{U}\gamma^\mu\partial_\mu(\hat{U}^{-1}\Psi') + g\overline{\Psi}'\hat{U}\gamma^\mu\mathbf{A}_\mu\cdot\hat{\mathbf{T}}\hat{U}^{-1}\Psi' \\
&= i\overline{\Psi}'\gamma^\mu\hat{U}\partial_\mu(\hat{U}^{-1}\Psi') + g\overline{\Psi}'\gamma^\mu\hat{U}\mathbf{A}_\mu\cdot\hat{\mathbf{T}}\hat{U}^{-1}\Psi' \\
&= i\overline{\Psi}'\gamma^\mu\partial_\mu\Psi' + i\overline{\Psi}'\gamma^\mu[\hat{U}(\partial_\mu\hat{U}^{-1})]\Psi' + g\overline{\Psi}'\gamma^\mu(\hat{U}\mathbf{A}_\mu\cdot\hat{\mathbf{T}}\hat{U}^{-1})\Psi' \\
&= i\overline{\Psi}'\gamma^\mu\partial_\mu\Psi' + g\overline{\Psi}'\gamma^\mu\left[\hat{U}\mathbf{A}_\mu\cdot\hat{\mathbf{T}}\hat{U}^{-1} + \frac{i}{g}\hat{U}(\partial_\mu\hat{U}^{-1})\right]\Psi' \qquad \text{(C.911)}
\end{aligned}
$$

for this to be identical to the original action we must have:

$$
\mathbf{A}'_\mu\cdot\hat{\mathbf{T}} = \hat{U}\mathbf{A}_\mu\cdot\hat{\mathbf{T}}\hat{U}^{-1} + \frac{i}{g}\hat{U}(\partial_\mu\hat{U}^{-1}). \qquad \text{(C.912)}
$$

We are forced to incorporate the term $\hat{U}(\partial_\mu\hat{U}^{-1})$, which is generated by gauging the kinematic energy of the field Ψ into the gauge transformation of the $\mathbf{A}_\mu$ fields. Let us look at the significance of this term in electrodynamics. In that case $\hat{U}$ is just $\hat{U} = \exp(ia(x))$. Then the gauge transformation reads

$$
A'_\mu(x) = A_\mu(x) + \frac{1}{g}\partial_\mu a(x).
$$

We can write the Lagrangean in the concise form

$$
\begin{aligned}
\mathcal{L} &= i\overline{\Psi}(\partial_\mu - ig\mathbf{A}_\mu\cdot\hat{\mathbf{T}})\Psi \\
&\equiv i\overline{\Psi}\nabla_\mu\Psi \qquad \text{(C.913)}
\end{aligned}
$$

Here we have introduced tha covariant derivative

$$
\nabla_\mu = \partial_\mu - ig\mathbf{A}_\mu\cdot\hat{\mathbf{T}} \qquad \text{(C.914)}
$$

The gauge transformation properties of $\mathbf{A}_\mu$ can be summarised as follows

$$
\nabla_\mu \to \nabla'_\mu = \hat{U}\nabla_\mu\hat{U}^{-1} \qquad \text{(C.915)}
$$

The kinetic energy term of the $\mathbf{A}_\mu$ fields is missing. By analogy with electrodynamics we write

$$
\mathbf{F}_{\mu\nu} = \nabla_\mu\mathbf{A}_\nu - \nabla_\nu\mathbf{A}_\mu \qquad \text{(C.916)}
$$

$$[\nabla'_\mu, \nabla'_\nu] = \hat{U}[\nabla_\mu, \nabla_\nu]\hat{U}^{-1}$$

$$\begin{aligned}
[\nabla_\mu, \nabla_\nu] &= [\partial_\mu - ig\mathbf{A}_\mu \cdot \hat{\mathbf{T}}, \partial_\nu - ig\mathbf{A}_\nu \cdot \hat{\mathbf{T}}] \\
&= [\partial_\mu, \partial_\nu] + \partial_\mu(-ig\mathbf{A}_\mu \cdot \hat{\mathbf{T}}) - (-ig\mathbf{A}_\mu \cdot \hat{\mathbf{T}})\partial_\mu \\
&\quad + \partial_\nu(-ig\mathbf{A}_\nu \cdot \hat{\mathbf{T}}) - (-ig\mathbf{A}_\nu \cdot \hat{\mathbf{T}})\partial_\nu \\
&\quad + (-ig)^2[\mathbf{A}_\mu \cdot \hat{\mathbf{T}}, \mathbf{A}_\nu \cdot \hat{\mathbf{T}}] \\
&= -ig\{(\partial_\mu A_\nu \cdot \hat{T}) - (\partial_\nu A_\mu \cdot \hat{T}) - ig[\mathbf{A}_\mu \cdot \hat{\mathbf{T}}, \mathbf{A}_\nu \cdot \hat{\mathbf{T}}]\} \qquad \text{(C.917)}
\end{aligned}$$

Notice that in the last line the derivatives act only on the gauges fields.

$$\mathbf{F}_{\mu\nu} = \frac{i}{g}[\nabla_\mu, \nabla_\nu] \qquad \text{(C.918)}$$

Then $\mathbf{F}_{\mu\nu}$ transforms as

$$\mathbf{F}'_{\mu\nu} = \hat{U}\mathbf{F}_{\mu\nu}\hat{U}^{-1} \qquad \text{(C.919)}$$

We obtain a gauge invariant Lagrangian by performing the trace over the internal indices

$$\begin{aligned}
\mathcal{L}'_A &= -\frac{1}{2}Tr\{(\mathbf{F}'_{\mu\nu} \cdot \hat{\mathbf{T}})(\mathbf{F}'^{\mu\nu} \cdot \hat{\mathbf{T}})\} \\
&= -\frac{1}{2}Tr\{\hat{U}(\mathbf{F}'_{\mu\nu} \cdot \hat{\mathbf{T}})\hat{U}^{-1}\hat{U}(\mathbf{F}'^{\mu\nu} \cdot \hat{\mathbf{T}})\hat{U}^{-1}\} \\
&= -\frac{1}{2}Tr\{\hat{U}^{-1}\hat{U}(\mathbf{F}'_{\mu\nu} \cdot \hat{\mathbf{T}})(\mathbf{F}'^{\mu\nu} \cdot \hat{\mathbf{T}})\} \\
&= \mathcal{L}_A \qquad \text{(C.920)}
\end{aligned}$$

C.18 Principle Bundles and Connections

A principle bundle is a fibre bundle $\pi : P \to E$ with fibre F equal to the structure group G and having the property that for all U_a and U_b with $U_a \cap U_b \neq \emptyset$,

$$\varphi_{ba} : U_a \cap U_b \to Left(F) \subset Diff(F),$$

where $Left(F) = \{L_g | L_g(h) = gh, \text{ for all } h \in G, g \in G\}$. In other words, changing coordinates corresponds to multiplying the fibre on the left by some element of G.

Lemma C.18.1 *For every $G-$principle bundle, G acts naturally on P on the right.*

Given $u \in P$, we want to define ug, for each $g \in G$. Let U be a neighbourhood about $\pi(u)$ that has a trivialization. Using these coordinates, represent u as $(\pi(u), h)$ where $h \in G$. Then define ug to be a point of P that has the coordinates $(\pi(u), hg)$. It is not hard to check that this definition is independent of coordinates, and then it is clear that it is a right action.

To place gauge theory in a more general perspective, it is helpful to consider fibre bundle formulism. The idea in gauge theory is to consider group bundles, where each fibre is a copy of the internal symmetry group, and where the base space corresponds to spacetime.

think about gauge theory geometrically - to understand the gauge field as a connection on the principal bundle.

We have already encountered one fibre bundle, from general relativity: the tangent space at a point in spacetime is a fibre, with the fibre bundle

the gauge fields play the same role in gauge theory as the Christoffel symbols play in the tangent in general relativity.

A principal fibre bundle allows one to simultaneously view the physical space, $\mathcal{M}$, referred to as the base space, and the bundle space E, where the bundle space generally reflects the symmetry group of the theory by associating with each point $x \in \mathcal{M}$ a fibre in E diffeomorphic to some Lie Group $\mathcal{G}$, refined to as the gauge group or structure group.

Bundles

A principal bundle $P(\mathcal{M}, G)$, where G is a Lie group and $\mathcal{M}$ is a compact manifold.

covering $\mathcal{M}$ with topological trivial open subsets U_a and giving a set of transition functions

$$t_{ab} : U_a \cap U_b \to G. \tag{C.921}$$

The transition functions are to satisfy three conditions:

$$
\begin{aligned}
t_{aa}(x) &= 1, \\
t_{ab}(x)t_{ba}(x) &= \mathbf{1}, \\
t_{ab}(x)t_{bc}(x)t_{ca}(x) &= \mathbf{1},
\end{aligned} \tag{C.922}
$$

for all points $x \in \mathcal{M}$ where the functions are all defined.

A gauge transformation is defined to be a collection of maps

$$
\lambda_a : U_a \to G. \tag{C.923}
$$

acting on transition functions in the following way,

$$
t_{ab}(x) \longmapsto \lambda_a^{-1}(x)t_{ab}(x)\lambda_b(x). \tag{C.924}
$$

Gauge-transformed transition functions define the same bundle. A bundle is characterized by a gauge equivalence class of transition functions satisfying (C.922).

A connection ω is defined globally is represented by a connection one-form ω_a on each U_a, with values in the Lie algebra of G, where, on each overlap $U_a \cap U_b$,

$$
\omega_a = \mathrm{Ad}_{t_{ab}^{-1}}\omega_b + t_{ab}^{-1}dt_{ab} \tag{C.925}
$$

where Ad is the adjoint representation.

$$
[T_a, T_b] = C^{abc}T_c, \tag{C.926}
$$

C^{abc} forms a representation of (C.926), the matrix $\mathcal{T}^a$ with components given by

$$
(\mathcal{T}^a)_{bc} = iC^{abc} \tag{C.927}
$$

This representation is called the *regular* representation or the *adjoint* representation.

C.19 Fibre Bundles

We have seen that a manifold has an associated with it a tangent space at *each* point p in the manifold. One can combine these spaces with the underlying manifold $\mathcal{M}$ to make one big manifold. Mathematically this combining (union) is expressed as,

$$T\mathcal{M} = \bigcup_{p\in\mathcal{M}} T_p\mathcal{M}. \tag{C.928}$$

This is known as the **tangent bundle space**. In fact we can combine spaces other than the tangent space with the manifold $\mathcal{M}$ to make a bigger topological space. These other spaces need not be vector spaces as the the tangent space is. The other space is, generally, referred to as the **fibre**. The whole space is known as a **fibre bundle**.

The simplest example of a fibre bundle, which is not a vector bundle, is the cylinder formed by choosing $\mathcal{M}$ as the real line R^1 and A as the circle S^1; one constructs the cylinder as a product between these two spaces, (actually, one should really say: *one constructs the cylinder as a product over the base manifold* R^1). This **product bundle**, also referred to as a **trivial bundle**, is denoted $(\mathcal{M} \times A, \mathcal{M}, \pi)$.The first entry refers to the fibre bundle, and the second refers to the base manifold. The third is a $\pi : \mathcal{M} \times A \to \mathcal{M}$. The projection map $\pi : E \to \mathcal{M}$ associates each fibre with a point in the base manifold. The inverse of the projection map π^{-1} associates a copy of the internal symmetry group (a fibre) with each point in the base manifold.

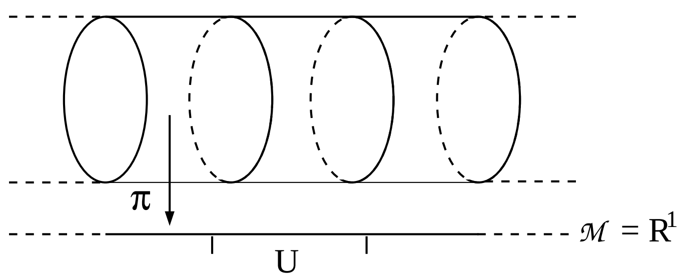


Figure C.47: Fibre bundle, $TS^1 = S^1 \times R$. The base manifold $\mathcal{M}$ (the real line R^1). The circle is the fibre. The fibre bundle consists of a manifold and a projection map π. $\pi^{-1}(U)$ is the local product space.

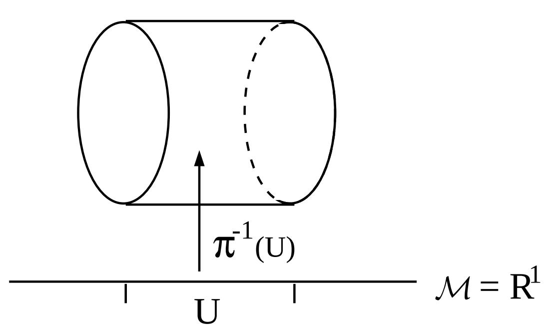


Figure C.48: The inverse map $\pi^{-1}(U)$ is the local product space.

A fibre bundle is *locally a product bundle.* A fibre bundle which is not a product bundle known as a ***non-trivial bundle***. We give an example which shows that if a manifold is locally the direct product of two other manifolds, it is nevertheless not, in general, a product manifold. Möbius strip

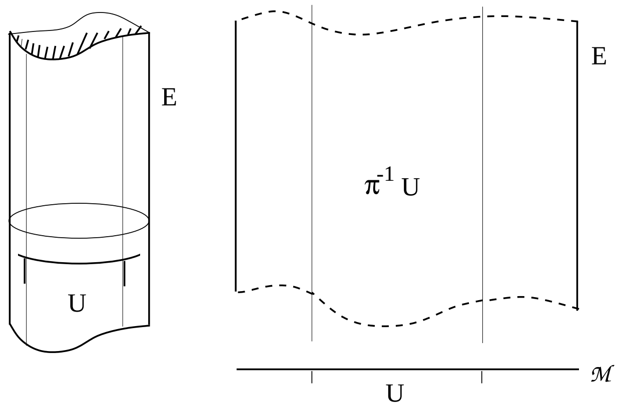


Figure C.49: tangent bundle, $TS^1 = S^1 \times R$. The base manifold $\mathcal{M}$ (the circle S^1) consists of a manifold and a projection map π. $\pi^{-1}(U)$ is the local product space.

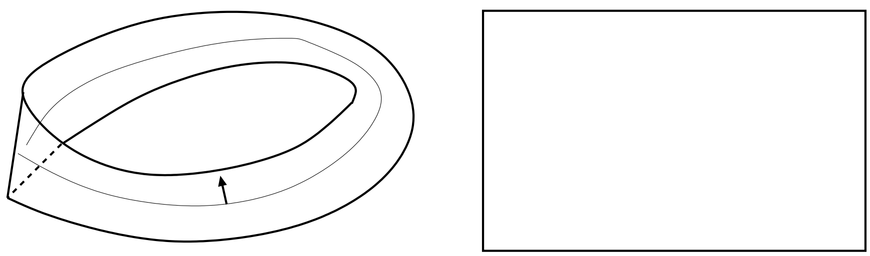


Figure C.50: Möbius.

The tangent bundle $T\mathcal{M}$ is locally of the form $U \times \mathcal{M}$. A ***section*** of this tangent bundle is simply a vector field $\tilde{\mathbf{w}}$ on $\mathcal{M}$. In the coordinate patch $(U, u_1, \ldots, u^n)$, $\tilde{\mathbf{w}}$ is given by its component functions $w_U^1(u), \ldots, w_U^n(u)$ with respect to the coordinate basis $\partial/\partial u$, but of course these functions are defined on U only, not on all of $\mathcal{M}$. In another patch V, the same field is described by another set of component functions $w_V^1(u), \ldots, w_V^n(u)$. At a point p in the overlap $U \cap V$ these two sets of vector component functions transform as

$$w_V^a(p) = [c_{UV}(p)]^a_b w_U^b(p) \tag{C.929}$$

where $c_{UV} = \partial v/\partial u$ the Jacobian matrix.

cotangent fibre bundle

$$T^*\mathcal{M} = \bigcup_{p \in \mathcal{M}} T_p^*\mathcal{M}. \tag{C.930}$$

Phase space is the cotangent fibre bundle.

$$w_V^a(p) = [c_{UV}(p)^T]^a_b w_U^b(p) \tag{C.931}$$

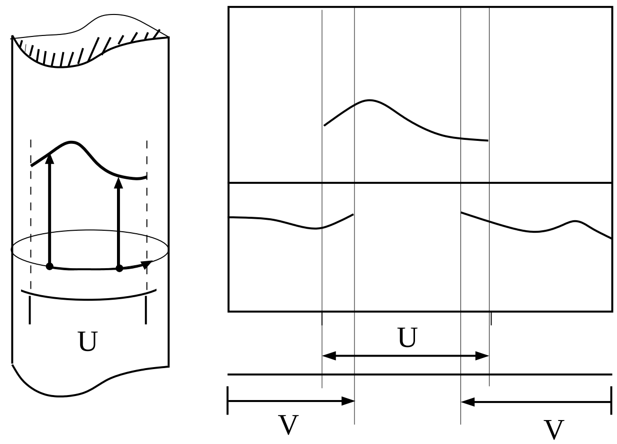


Figure C.51: tangent bundle, $TS^1 = S^1 \times R$. The curved line is a section.

Some examples:

$$\begin{aligned} \pi &\approx R^n \\ S^{n-1} &\subset \pi^{-1}(p) \end{aligned} \quad (C.932)$$

$T_0\mathcal{M}$ is atypical, since it is a *subbundle* of the vector bundle $T\mathcal{M}$.

Let $M(n \times n)$ be the set of all $n \times n$ real matrices. We associate to the matrix $\underline{\underline{x}}$ the point in the n^2-dimensional Euclidean space whose coordinates are $x_{11}, x_{12}, \ldots x_{nn}$. The topological of the parameter space is R^n. The ***general linear group*** $Gl(n, R)$ is the group of all real $n \times n$ matrices $x = (x_{ij})$ with determinate $\det x \neq 0$.

$$X^i \mathbf{e_i} = x_{11}\mathbf{e_{11}} + x_{12}\mathbf{e_{12}} + x_{21}\mathbf{e_{21}} + x_{22}\mathbf{e_{22}} + \cdots + x_{nn}\mathbf{e_{nn}}. \quad (C.933)$$

It is clear from $(xy)_{ij} = \sum_k x_{ik}y_{kj}$ that the product matrix has coordinates that are smooth functions of the coordinates of $\underline{\underline{x}}$ and $\underline{\underline{y}}$

a vector tangent to G is itself a matrix

Definition The ***left (right) translation*** is $G \mapsto G$, $l_a(g) = ag$, and the ***right translation***

$$\varphi_U : U \times GL(R, n) \to \pi^{-1}(U) \quad (C.934)$$

a local trivialization, such for any $x \in U$ and $(X_i^k) \in GL(R, n)$,

$$\varphi_U(x, (X_i^k)) = (x; e_1, \ldots, e_n) \quad (C.935)$$

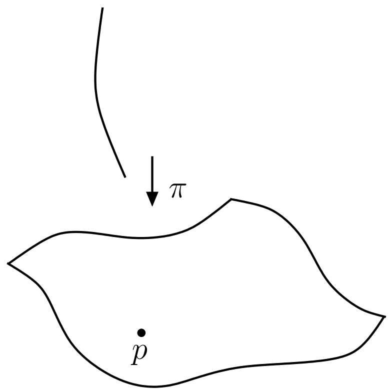


Figure C.52: Travelling up fibre.

$$y^a = y^a(x^1, \ldots, x^n), \quad i = 1, \ldots, n \tag{C.936}$$

$$\varphi_V : V \times GL(R, n) \to \pi^{-1}(V) \tag{C.937}$$

$$\mathbf{e_i} = X_i^k \frac{\partial}{\partial x^k} = Y_i^k \frac{\partial}{\partial y^k} \tag{C.938}$$

or

$$Y_i^k = X_i^k \left(\frac{\partial y^k}{\partial x^j} \right)_x \tag{C.939}$$

by () the right translation of X is given by the Jacobian matrix $\left(\frac{\partial y^k}{\partial x^j} \right)_x$

one can identify a family of transition functions

$$c_{UV} : U \cap V \to GL(R, n) \tag{C.940}$$

given by

$$c_{UV}(x) = \varphi_V^{-1} \circ \varphi_U \tag{C.941}$$

In fact $\varphi_V^{-1} \circ \varphi_U$ is precisely the right translation of X

Details

$$L_{g_1} \circ L_{g_2} = L_{g_1 g_2} \tag{C.942}$$

or

$$\begin{aligned} L_{g_1} \circ L_{g_2} &= L_{g_1}(x; (g_1 e)_1, \ldots, (g_1 e)_n) \\ &= (x; (g_1(g_2 e))_1, \ldots, (g_1(g_2 e))_n) \\ &= (x; (g_1 g_2 e)_1, \ldots, (g_1 g_2 e)_n) \end{aligned} \tag{C.943}$$

The left translation along g maps a neighborhood of e onto one of g. There is a natural map of a vector at e to one on g. a vector tangent to G is itself a matrix

The one-parameter subgroup generated by any matrix A is the integral curve through e of the left-invariant vector field whose tangent at e is A.

As matrices

$$g(t+s) = g(t)g(s), \quad \Leftrightarrow \quad g_{ij}(t+s) = \sum_k g_{ik}(t)g_{kj}(s) \tag{C.944}$$

$$g_A(t + \Delta t) = g_A(t)g_A(\Delta t) \tag{C.945}$$

$$\frac{dg_A}{dt} = g_A(t)A \tag{C.946}$$

Differentiate both sides with respect to s and put $s = 0$

$$g'(t) = g(t)g'(0). \tag{C.947}$$

Since $g'(0)$ is a constant matrix, the solution to this is

$$g(t) = g(0)\exp(tg'(0)) \tag{C.948}$$

$$1 + tg + \frac{1}{2!}t^2g^2 + \ldots \tag{C.949}$$

is the most general form for a 1-parameter subgroup of a matrix group G.

$$\Phi(x, y) = \exp_p(x\mathbf{e} + y\mathbf{f}) \tag{C.950}$$

$$\frac{\partial}{\partial x}\Phi(x,y)\bigg|_{(x,y)} = \frac{\partial}{\partial x}\exp_p(x\mathbf{e})\bigg|_{(x,y)} = \mathbf{e}. \tag{C.951}$$

Theorem Let X, Y be a pair of left (right) invariant vector fields. Then $[X,Y]$ is left (right) invariant.

Proof:

$$\begin{aligned}
[X',Y']_i &= X'_j\partial'^j Y'_i - Y'_j\partial'^i X'_i \\
&= X_j\partial^j Y'_i - Y_j\partial^j X'_i \\
&= X_j\partial^j\left(\frac{\partial y_i}{\partial x_k}Y_k\right) - Y_j\partial^j\left(\frac{\partial y_i}{\partial x_k}Y_k\right) \\
&= \frac{\partial y_i}{\partial x_k}(X_j\partial^j Y_k - Y_j\partial^j X_k) + \frac{\partial^2 y_i}{\partial x_j\partial x_k}(X_jY_k - Y_jX_k).
\end{aligned} \tag{C.952}$$

C.19.1 The Structure Group of a Bundle

In a vector bundle each $c_{UV} \in Gl(n)$. We have seen that for a Riemannian manifold $\mathcal{M}$ of dimension n, we may choose $c_{UV}(x) \in O(n)$ by using orthonormal frames. In a general bundle it may be possible to choose the $c_{UV}(x)$ such that they all lie in a specific Lie group G

$$c_{UV} : U \cap V \to G \tag{C.953}$$

We then say that G is the ***structure group*** of the bundle.

$$c_{UV} = \begin{pmatrix} \cos\alpha(x) & -\sin\alpha(x) \\ \sin\alpha(x) & \cos\alpha(x) \end{pmatrix} \in SO(2). \tag{C.954}$$

The orthonormal frames have allowed us to ***reduce*** the structure group from $Gl(2,R)$ to $SO(2)$.

C.19.2 Frame Bundle

Frame bundles are fundamental because from this we can construct the tangent bundle, and by a similar construction the cotangent bundle and all the tensor bundles.

If $\{v_1,\ldots,v_n\}$ is a local field of linear frames on a neighbourhood U in $\mathcal{M}$, and $\{e_1,\ldots,e_n\}$ is a frame at $x \in U$, then $\lambda_a^b V_{bx}$, where the number λ_a^b are the entries of a non-singular matrix λ. Thus relative to the local field, each linear frame determines an element of $GL(n,\mathbb{R})$; and each element of $GL(n,\mathbb{R})$ determines a linear frame. We therefore take $GL(n,\mathbb{R})$ as the typical fibre bundle.

The frame bundle is clearly not a vector bundle. Its typical fibre has instead the structure of a group.

C.19.3 The Idea of a Principal Bundle

If we have a connection in a principle bundle $P \to \mathcal{M}$ then a choice of a horizontal subspace at u is equivalent to a choice of projection

$$proj_u : T_u P \to V_u.$$

Then the vertical space (i.e. a fiber of the vertical bundle) may be viewed as the tangent space to the fibre, G. In turn, this an be viewed as g, the Lie algebra of G, so we really have

$$proJ_u : T_u P \to g.$$

Therefore a connection is equivalent to having a g valued 1-form on P which is invariant under the right action of G.

Let G be a Lie group and $\mathcal{M}$ a smooth manifold. A ***principle G bundle*** is over $\mathcal{M}$ is a manifold which locally looks like $\mathcal{M} \times G$.

We have already found the transition functions for this fibre bundle. We consider the intersection of two coordinate patches, and a point p

The transition functions are given by the Jacobian matrix. They are elements of the group $GL(n, \mathbf{R})$, the general linear transformations on an n dimensional real vector space.

The bundles $\mathcal{T}(\mathcal{M})$ and $\mathcal{T}^*(\mathcal{M})$ are vector bundles. a tangent bundle and a co-tangent bundle. Phase space is an example of a co-tangent bundle.

A principal bundle is a bundle whose fibres are the transition functions themselves - points in the fibre are elements of the structure group. We have just considered a principal fibre bundle, that of a frame space.

Definition The right action of G on $\pi^{-1}(U)$ is defined by $\phi_i^{-1}(ua) = (p, g_i a)$, that is, (fig C.19.3)

$$ua := \phi_i(p, g_i a) \tag{C.955}$$

for any $a \in G$ and $u \in \pi^{-1}(p)$.

□

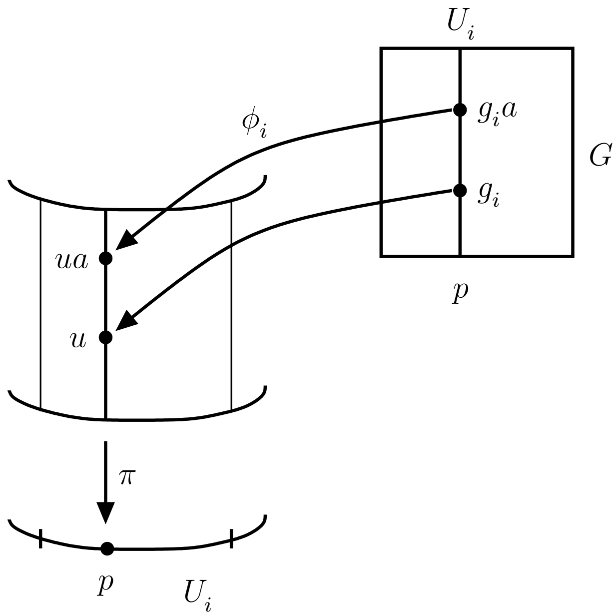


Figure C.53: Defintion of the right action of G on the principal fibre bundle P.

Definition Principal fibre bundle.

i) A differentiable manifold E called the total space;

ii) A differentiable manifold $\mathcal{M}$ called the base space;

iii) A surjection $\pi : E \to \mathcal{M}$ called the projection. The inverse image $\pi^{-1}(x) \equiv G_x \simeq G$ called the fibre at x;

iv) A Lie group G called the structure group, which acts on the fibre G on the left;

v) An open covering $\{U_I\}$ of $\mathcal{M}$ with a diffeomorphism $\phi_i : U_i \times G \to \pi^{-1}(U_i)$ such that $\pi\phi_i(x, g) = x$. The map is called the local gauge or local trivialisation since ϕ_i^{-1} maps $\pi^{-1}(U_i)$ onto the direct product $U_i \times G$;

vi) If we write $\phi_i(x, g) = \phi_{i,x}(g)$, the map $\phi_{i,x} : G \to G_x$ is a diffeommorphism. On $U_i \cap U_j \neq \emptyset$, we require that $t_{ij}(x) \equiv \phi_{i,x}^{-1}\phi_{j,x} : G \to G$ be an element of the structure group G. The ϕ_i and ϕ_j are related by a smooth map $t_{ij} : U_i \cap U_j \to G$ such that $\phi_j(x, g) = \phi_i(x, t_{ij}(x)g)$. $\{t_{ij}\}$ are the transition functions.

□

C.19.4 Principal Bundle

We are now ready to state the precise definition of a principal bundle.

We will say that a fibre bundle

$$\{P, \mathcal{M}, \pi, F, G\} \tag{C.956}$$

is a **principal bundle** *if the fibre F is the same as the group G* and if the transition functions $c_{UV}(x)$ act on $F = G$ by *left translations.*

and

C.19.5 Action of the structure Group on a Principal Bundle

In practical situations transition functions are the gauge transformations required for pasting local charts together.

C.19.6 Connections on Principal Bundles

When we make comparison of objects in two different spaces is made by a prescribed mapping, and the mappings that connect the various spaces are called connections.

The first definition of a connection has a clear geometric meaning,

1. a $G-$invariant horizontal distribution $H_u P \subset T_u P$.

An equivalent, less geometric definition of a connection consistent with the first definition is:

2. a one-form $\omega \in \Omega^1(P; \mathfrak{g})$ satisfying $\omega(s(X)) = X$.

We come to the final definition of the connection on a principal bundle, expressed in the form of $\mathcal{G}-$valued one-forms on the base manifold $\mathcal{M}$, instead of one-forms on the total space P. This definition is most suitable in physics applications.

3. a family of one-forms $A_i \in \Omega^1(P; \mathfrak{g})$ satisfying equation (for simplicity here in the case of matrix groups G):

$$A_j = t_{ij} A_i t_{ij}^{-1} + t_{ij}^{-1} dt_{ij}.$$

Definition (1): A connection on a principal bundle $\pi : P \to \mathcal{M}$ with group G is a smooth assignment to each $u \in P$ of an $n-$dimensional subspace $\mathcal{H}(p) \subset T_u P$ of the tangent space to P ($n = dim\ \mathcal{M}$) such that

$$T_u P = \mathcal{H}(p) \oplus \mathcal{V}(p), \tag{C.957}$$

where $\mathcal{V}(p)$ is the subspace

$$\mathcal{V}(p) := \{X \in T_uP : \pi_*(X) = 0\}. \tag{C.958}$$

The subspace field $\mathcal{H}$ is also required to be invariant under the left action L_g $(g \in G)$:

$$(L_g)_*\mathcal{H}(p) = \mathcal{H}(gp), \tag{C.959}$$

for all $g \in G$. $\mathcal{V}(p)$ and $\mathcal{H}(p)$ are called the vertical and horizontal subspace of T_uP, respectively.

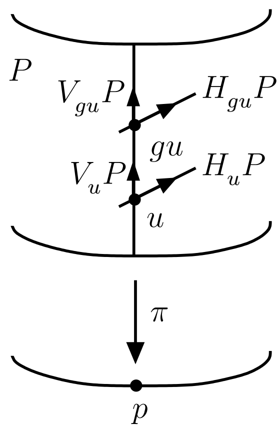


Figure C.54: The horizontal subspace $H_{gu}P$, defining by the connection in definition (1), is obtained from H_uP by the left action.

$\square$

Definition (2): Suppose $\mathcal{G}$ is the Lie algebra of a Lie group G, which is the structure group of a principal fibre bundle $\pi : P \to \mathcal{M}$. Let $A \in \mathcal{G}$, and $A^\sharp$, called a fundamental field, be the vector field on P defined by

$$A^\sharp(p) := \frac{d}{dt}\exp(At)p\Big|_{t=0}. \tag{C.960}$$

A connection on the principal bundle $\pi : P \to \mathcal{M}$ is a $\mathcal{G}-$valued one for satisfying the following two properties:

$$(i) \qquad \omega(A^\sharp(p)) = A \tag{C.961}$$
$$(ii) \qquad (L_g)^*(\omega) = \mathcal{A}d(g)(\omega). \tag{C.962}$$

$\square$

Definition (3): A connection on a principal bundle is an assignment to each local trivialisation $\phi_i : \pi^{-1}(U_i) \to U_i \times G$ (a choice of gauge in physics terms) a Lie algebra one-form ω_i on U_i which satisfies the following rule between different local trivialisations,

$$\omega_j(X_p) = (R_{t_{ij}^{-1}(p)})_*((t_{ij})_*(X_p)) + \mathcal{A}d(t_{ij}(p))(\omega_i(X_p)), \tag{C.963}$$

for all $X_p \in T_p\mathcal{M}$ and $p \in U \cap V$.

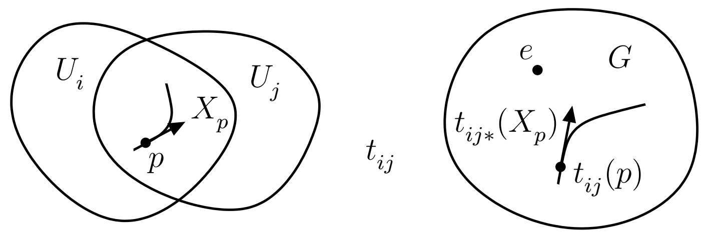


Figure C.55:

□

Since a non-trivial principal bundle does not admit a global section, the pull-back $A_i = s_i^*\omega$ exists locally but not necessarily globally.

Lemma C.19.1 *The two definitions (1) and (2) of a connection on a principal fibre bundle are equivalent.*

Proof:

(2) ⇒ (1):

Suppose we are given a $\mathcal{G}-$valued connection one-form ω, as in definition (2). Consider the field of subspaces defined by

$$\mathcal{H}(u) = \{X \in T_uP : \omega_u(\mathbb{X}) = 0\}.$$

By (), $\omega(\mathbb{Y})A \in \mathcal{G}$ for any $\mathbb{Y} \in \mathcal{V}(p)$. Hence $\mathcal{V}(p) \cap \mathcal{H}(p) \neq \emptyset$ and $T_pP = \mathcal{V}(p) \oplus \mathcal{H}(p)$. We must prove that

$$L_{g*}\mathcal{H}(u) = \mathcal{H}(gu).$$

Fix a point $u \in P$ and define $\mathcal{H}(u)$ as above. Take $\mathbb{X} \in \mathcal{H}(u)$ and construct $L_{g*}\mathbb{X} \in T_{gu}$. We find

$$\omega_{gu}(L_{g*}\mathbb{X}) = L_g^*\omega_u(\mathbb{X}) = g\omega_u(\mathbb{X})g^{-1} = 0$$

since $\omega(\mathbb{X}) = 0$. Therefore, $L_{g*}\mathbb{X} \in \mathcal{H}(gu)$. Note that L_{g*} is an invertible linear map. Hence any vector $\mathbb{Y} \in \mathcal{H}(gu)$ is expressed as $\mathbb{Y} = L_{g*}\mathbb{X}$ for some $\mathbb{X} \in \mathcal{H}(u)$. Thus $L_{g*}\mathcal{H}(u) = \mathcal{H}(gu)$.

(1) ⇒ (2):

We define a $\mathcal{G}-$valued one-form ω by

$$\omega(A^\sharp + X) = A \tag{C.964}$$

Note $0 \in \mathcal{H}(u)$ and setting $\mathbb{X} = 0$, we have $\omega(A^\sharp) = A$, and as the connection is linear this implies $\omega(\mathbb{X}) = 0$ for all $\mathbb{X} \in \mathcal{H}(u)$.

To complete the proof we must establish property (ii) (C.962). Any $\mathbb{X} \in T_uP$ can be written $\mathbb{X} = a\mathbb{H} + b\mathbb{V}$, where $H \in \mathcal{H}(u)$ and $V \in \mathcal{V}(u)$. This allows us to need only verify (C.962) separetly for arbitrary horizontal and vertical vectors. Suppose $\mathbb{H} \in \mathcal{H}(u)$. Then $\omega(\mathbb{H}) = 0$. By (C.959), $\omega((L_g)_*\mathbb{H}) = 0$ and (C.962) is verified for $\mathbb{H} \in \mathcal{H}(u)$. Consider $\mathbb{V} = A^\sharp(u) \in \mathcal{V}(u)$, for some $A \in \mathcal{G}$. We compute

$$\begin{aligned}
\omega_{gp}((L_g)_*A^\sharp(u)) &= \omega_{gp}\left((L_g)_*\frac{d}{dt}(\exp(At)u)\Big|_{t=0}\right) \\
&= \omega_{gu}\left(\frac{d}{dt}(g\exp(At)u)\Big|_{t=0}\right) \\
&= \omega_{gu}\left(\frac{d}{dt}(g\exp(At)g^{-1}\cdot gu)\Big|_{t=0}\right) \\
&= \omega_{gu}\left(\frac{d}{dt}(ad_g(\exp(At))\ gu)\Big|_{t=0}\right) \\
&= \omega_{gu}\left(\frac{d}{dt}(\exp(\mathcal{A}d(g)At)\ gu)\Big|_{t=0}\right) \\
&= \omega_{gu}(\ (\mathcal{A}d(g)A)^\sharp\ gu)
\end{aligned} \tag{C.965}$$

As $\omega(A^\sharp) = A$

$$\omega_{gu}(\ (\mathcal{A}d(g)A)^\sharp\ gu) = \mathcal{A}d(g)A$$

and then

$$\mathcal{A}d(g)A = (\mathcal{A}d(g))(\omega(A^\sharp(u)))$$

therefore we have

$$\omega_{gp}((L_g)_*A^\sharp(u)) = (\mathcal{A}d(g))(\omega(A^\sharp(u))),$$

that is, we have verified (C.962) for vertical vectors $A^\sharp(u)$ as well.

□

In the next section provide the proof of the equivalence between of definitions (2) and (3).

C.19.7 Gauge Fields

We now make contact with the more familiar notion of gauge fields as used in physics, which live on $\mathcal{M}$ instead of P.

Theorem C.19.2 *Given a $\mathfrak{g}-$valued one-form A_i on U_i and a local section $s_i : U_i \to \pi^{-1}(U_i)$, there exists a connection one-form ω such that $A_i = s_i^*\omega$ on U_i*

Proof:

(2) $\Rightarrow$ (3):

.....

Let ω be a $\mathcal{G}-$valued connection one-form as given in definition (2), and $\phi_i : \pi^{-1}(U) \to U \times G$ be a local trivialisation. Associated with ϕ_i there is a local section given by $s_i(p) = \phi_i^{-1}(p, e)$. We define a $\mathcal{G}-$valued one-form A_i on U by

$$A_i = s_i^*\omega,$$

then we show that it is the connection one-form in the sense of definition (3). To do this we have to show that the gauge transformation

$$\omega_i \to \omega_j = s_j^*\omega$$

is given by (C.963).

....

Recall that we have local sections $s_i : U_i \to \pi^{-1}(U_i)$ associated canonically to the trivialisation of the bundle. g_i is the canonical local trivialisation defined by $\phi_i(u) = (p, g_i)$ for $u = s_i(p)g_i$. Let us define a $\mathfrak{g}-$valued one-form ω_i on P by

$$\omega_i = g_i^{-1}\pi^* A_i g_i + g_i^{-1} d_P g_i \tag{C.966}$$

where d_P is the exterior derivative on P.

We first show that $s_i^*\omega_i = A_i$. For $X \in T_pM$. Note

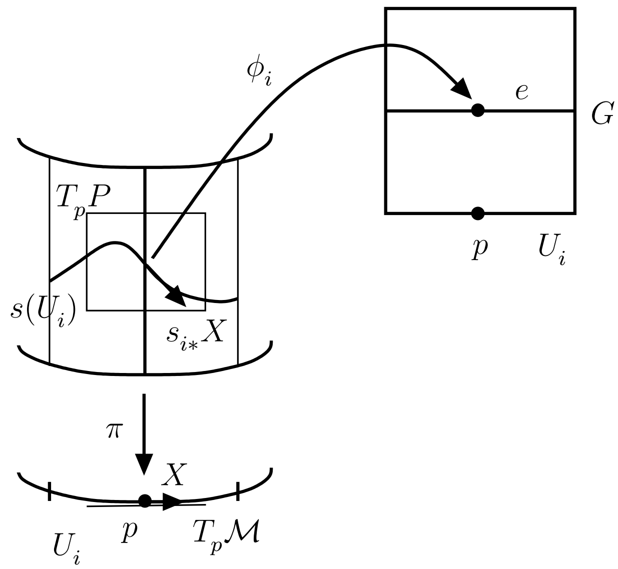


Figure C.56:

$$s_i^*\omega(X) = \omega_i(s_{i*}X)$$

As $g_i = e$ at s_i

$$\begin{aligned} s_i^*\omega(X) &= \pi^* A_i(s_{i*}X) + d_P g_i(s_{i*}X) \\ &= A_i(\pi_{i*}s_{i*}X) + d_P g_i(s_{i*}X) \end{aligned} \tag{C.967}$$

We have

$$d_P g_i(s_{i*}X) = \frac{d}{dt} g_i(s_i(t))\Big|_{t=0} = 0$$

as $g \equiv e$ along $s_i(t)$. Thus we have obtained $s_i^*\omega_i(X) = A_i(X)$.

Next we show that ω_i satisfies the axioms of a connection one-for given in definition...

We consider the subspace of T_uP tangent to the fibres. Let $X = A \in V_uP$, $A \in \mathfrak{g}$. It follows from $\pi_* X = 0$. Now we have from $g(u\exp(tA)) = g(u)\exp(tA)$

$$\begin{aligned} \omega_i(A^\sharp) &= g_i^{-1}(d_P g_i)(A^\sharp) \\ &= g_i(u)^{-1}\frac{dg(u\exp(tA))}{dt}\Big|_{t=0} \\ &= g_i(u)^{-1}g_i(u)\frac{dg\exp(tA)}{dt}\Big|_{t=0} = A. \end{aligned}$$

Take $X \in T_P P$ and $h \in G$. We have

$$\begin{aligned} R_h^* \omega_i(X) &= \omega_i(R_{h*}X) \\ &= g_i^{-1}(uh) A_i(\pi_{i*} R_{i*} X) g_i(uh) + g_i^{-1}(uh)(d_P g_i(uh))(R_{h*}X) \end{aligned}$$

Since $g_i(uh) = g_i(u)h$ and $\pi_{i*} R_{i*} X = \pi_{i*} X$ the first term above can be written

$$h^{-1} g_i^{-1}(u) A_i(\pi_{i*} X) g_i(u) h$$

For the second term we work out

$$\begin{aligned} g_i^{-1}(uh)(d_P g_i(uh))(R_{h*}X) &= g_i^{-1}(uh) \frac{d}{dt} g_i(\gamma(t)h)\Big|_{t=0} \\ &= h^{-1} g_i^{-1}(u) \frac{d}{dt} g_i(\gamma(t))h \\ &= h^{-1} g_i^{-1}(u)(d_P g_i(u))(X)h. \end{aligned}$$

Here $\gamma(t)$ is a curve through $u = \gamma(0)$, whose tangent vector at u is X. Therefore

$$R_h^* \omega_i(X) = h^{-1} \omega_i(X) h.$$

Hence the $\mathfrak{g}-$valued one-form ω_i defined by () satisfies $A_i = s_i^* \omega$ and the axioms of a connection one-form.

Global one-form on P?

Suppose we are given local one-forms A_i, A_j, etc., on $\mathcal{M}$ that satify the gauge transformation given in definition (3). First we will construct local one-forms ω_U, $\omega_V, \dots,$ on P from the local one-forms A_U,$A_V, \dots,$ on $\mathcal{M}$. We then need to show that $\omega_U = \omega_V$ on $\pi^{-1}(U \cap V)$, in order for the local one-forms ω_U, ω_V, etc., to collectively define a global connection one form ω as in definition (2).

.....

(3) $\Rightarrow$ (2):

We first define the map $\tilde{\omega}_i : T_u P \to \mathcal{G}$ by

$$\tilde{\omega}_i((s_i)_* X_p + A^\sharp) = A_i(X_p) + A. \tag{C.968}$$

A local $\mathcal{G}-$valued one-form on all of $\pi^{-1}(U_i)$ can be constructed from this map via

$$\omega_i(\mathbb{X}_{gu}) = (\mathcal{A}d(g))\ \tilde{\omega}_i\ ((L_{g^{-1}})_*\mathbb{X}_{gu}) \tag{C.969}$$

for $\mathbb{X}_{gu} \in T_{gu}P$. This one-form reduces to the one form $\tilde{\omega}_i$ on T_uP for $g = e$. The one-form ω_i satifies condition (ii) of definition 2 on the restricted bundle $\pi^{-1}(U)$

$$\begin{aligned}
\omega_{igu}((L_g)_*\mathbb{X}_u) &= \mathcal{A}d(g))\ \tilde{\omega}_{iu}\ ((L_{g^{-1}})_*((L_g)_*\mathbb{X}_u)) \\
&= \mathcal{A}d(g))\ \tilde{\omega}_{iu}\ (\mathbb{X}_u) \\
&= \mathcal{A}d(g))\ \omega_{iu}\ (\mathbb{X}_u),
\end{aligned}$$

and thus is a connection in the sense of definition (2) on this restricted bundle.

Suppose s_j is another local trivialisation, where $U_i \cap U_j \neq \emptyset$. We can similarly define a local connect ω_j on $s_{U_j}(U_i \cap U_j)$, as then they would agree on all of $\pi^{-1}(U \cap V)$ by virtue of (C.969). That is, we wish to show that

$$\tilde{\omega}_i((s_j)_*X_p + A^\sharp) = \tilde{\omega}_j((s_j)_*X_p + A^\sharp)$$

Now since $\omega_i(A^\sharp) = A = \omega_j(A^\sharp)$, it is sufficient to check that $\tilde{\omega}_i((s_j)_*X_p) = \tilde{\omega}_j((s_j)_*X_p)$, for $x \in U \cap V$ and $X_p \in T_p\mathcal{M}$.

suppose $\gamma(t)$ is a curve in $\mathcal{M}$ with $p = \gamma(0)$ and $X_p = \gamma'(0)$. We compute $(s_j)_*(X_p)$ by

$$\begin{aligned}
(s_j)_*(X_p) &= \frac{d}{dt}s_j(\gamma(t))\Big|_{t=0} \\
&= \frac{d}{dt}\big(t_{ij}(\gamma(t))s_j(\gamma(t))\big)\Big|_{t=0} \\
&= \frac{d}{dt}\big(t_{ij}(p)s_j(\gamma(t))\big)\Big|_{t=0} + \frac{d}{dt}\big(t_{ij}(\gamma(t))s_j(p)\big)\Big|_{t=0} \\
&= (L_{t_{ij}(p)})_*(s_i)_*(X_p) + \frac{d}{dt}\big(t_{ij}(\gamma(t))t_{ij}^{-1}(p)s_j(p)\big)\Big|_{t=0} \\
&= (L_{t_{ij}(p)})_*(s_i)_*(X_p) + \big((R_{t_{ij}^{-1}(p)})_*(t_{ij})_*(X_p)\big)^\sharp_{s_j(p)}.
\end{aligned}$$

Using this and by using the gauge transformation formula we then have

$$\begin{aligned}
\tilde{\omega}_i((s_j)_*X_p) &= \tilde{\omega}_i\Big(\big((R_{t_{ij}^{-1}(p)})_*(t_{ij})_*(X_p)\big)^\sharp_{s_j(p)} + (L_{t_{ij}(p)})_*(s_i)_*(X_p)\Big) \\
&= (R_{t_{ij}^{-1}(p)})_*(t_{ij})_*(X_p) + \mathcal{A}d(t_{ij})\ (A_i(X_p)) \\
&= \tilde{\omega}_i((s_j)_*X_p) = A_j(X_p) \qquad (C.970)
\end{aligned}$$

By () $\tilde{\omega}_j((s_j)_*X_p) = A_j(X_p)$ the proof is complete.

$\square$

If the principal bundle is non-trivial, it is not generally possible to describe the connection ω on P in terms of a single Yang-Mills $A_\mu(x)$ field on $\mathcal{M}$. Instead, one must cover $\mathcal{M}$ with local trivalising charts, and then local Yang-Mills fields associated with any pair of overlapping charts U_i, U_j will be related on $U_i \cap U_j$ by

$$A_\mu^{(j)}(x) = \Omega(x)^{-1} A_\mu^{(i)}(x) \Omega(x) + \Omega(x)^{-1} \partial_\mu \Omega(x)$$

with the corresponding local gauge function $\Omega_{ij}(x)$ satifying the $s_i(x) = s_j(x)\Omega_{ij}(x)$. Note that these functions $\Omega_{ij} : U_i \cap U_j \to G$ are precisely the bundle transition functions.

Let $P(\mathcal{M}, G)$ be a principal bundle over $\mathcal{M}$ and U a chart of $\mathcal{M}$. Take two local sections s_1 and s_2 over U such that $s_2(p) = s_1(p)g(p)$. The corresponding local forms A_1 and A_2 are related as

$$A_2 = g^{-1} A_1 g + g^{-1} dg. \tag{C.971}$$

In components, this reads

$$A_{2\mu}(p) = g^{-1}(p) A_{1\mu}(p) g(p) + g^{-1}(p) \partial_\mu g(p). \tag{C.972}$$

which is simply the gauge transformation ().

Example Electrodynamics:

Let P be a U(1) bundle over $\mathcal{M}$. Take overlapping charts U_i and U_j. Let A_i (A_j) be a local connection form on U_i (U_j). The transition function $t_{ij} : U_i \cap U_j \to U(1)$ is given by

$$t_{ij}(p) = \exp(i\theta(p)) \qquad \theta(p) \in \mathbb{R}. \tag{C.973}$$

A_i and A_j are related by

$$\begin{aligned} A_j(p) &= t_{ij}(p)^{-1} A_i(p) t_{ij}(p) + t_{ij}(p)^{-1} dt_{ij}(p) \\ &= A_i(p) + i d\theta(p) \end{aligned} \tag{C.974}$$

In components, we have

$$A_{j\mu} = A_i\mu + i\partial_\mu \theta. \tag{C.975}$$

$\square$

Example General relativity:

The important example is a connection in the principal $GL(n,\mathbb{R})-$bundle $\mathbf{B}(\mathcal{M})$ of frames on a $n-$dimensional manifold $\mathcal{M}$. Any local coordinate chart (U,φ) on $\mathcal{M}$ provides a local section $s: U \to \mathbf{B}(\mathcal{M})$ by associating with $x \in U \subset \mathcal{M}$, the local frame $(\partial_1, \partial_2, \ldots, \partial_n)_x$. If ω is a connection one-form on $\mathbf{B}(\mathcal{M})$, let $\Gamma := s^*\omega$ denote the associated $\mathcal{L}(GL(n,\mathbb{R}))-$valued one-form on U, and consider the relation between Γ and the local one-form Γ' associated with another coordinate chart (U',φ') such that $U \cap U' \neq \emptyset$.

The local section of $\mathbf{B}(\mathcal{M})$ associated with s' is $(\partial/\partial x'^1, \ldots, \partial/\partial x'^m)$, and the transition function $J : U \cap U' \to GL(n,\mathbb{R})$ is just the Jacobian of the coordinatre transformations:

$$(\partial_{\mu'})_x = (\partial_\nu)_x J^\nu{}_\mu(x)$$

where

$$J^\nu{}_\mu(x) := \partial x^\nu / \partial x'^\mu.$$

Then,

$$\begin{aligned}
\Gamma'_\mu(x) &= (s^*\omega)_x(\partial_{\mu'}) \\
&= J^\alpha_{\ \mu}(x)(s'^*\omega)_x(\partial_\alpha) \\
&= J^\alpha_{\ \mu}(x)\big(J^{-1}(x)\Gamma_\alpha(x)J(x) + J^{-1}(x)\partial_\alpha J(x)\big)
\end{aligned} \tag{C.976}$$

where we have used the result () for a connection in a bundle whose srtucture group is a matrix group. The second term reads

$$\begin{aligned}
J^\alpha_{\ \mu}(x)(J^{-1}(x)\partial_\alpha J(x))^\epsilon{}_\delta &= \frac{\partial x^\alpha}{\partial x'^\mu}\left(\frac{\partial x'^\epsilon}{\partial x^\lambda}\frac{\partial}{\partial x^\alpha}\left(\frac{\partial x^\lambda}{\partial x'^\delta}\right)\right) \\
&= \frac{\partial x'^\epsilon}{\partial x^\lambda}\frac{\partial^2 x^\lambda}{\partial x'^\mu \partial x'^\delta}.
\end{aligned}$$

If $G_\lambda{}^\rho$ is some basis for the Lie algebra $M(n,\mathbb{R})$ (the set of all $n \times n$ real matrices) of $GL(n,\mathbb{R})$ then we can write the matrix-valued one-form Γ_μ as

$$(\Gamma_\mu)^\epsilon{}_\delta = \Gamma^\lambda_{\mu\rho}(G_\lambda{}^\rho)^\epsilon{}_\delta.$$

The first term in () is then

$$
\begin{aligned}
J^{\alpha}{}_{\mu}(x)(J^{-1}(x)\Gamma_{\alpha}(x)J(x))^{\epsilon}{}_{\delta} &= \frac{\partial x^{\alpha}}{\partial x'^{\mu}}\left(\frac{\partial x'^{\epsilon}}{\partial x^{\beta}}(\Gamma_{\alpha})^{\beta}{}_{\lambda}(x)\frac{\partial x^{\lambda}}{\partial x'^{\delta}}\right) \\
&= \frac{\partial x^{\alpha}}{\partial x'^{\mu}}\left(\frac{\partial x'^{\epsilon}}{\partial x^{\beta}}(G_{\gamma}{}^{\rho})^{\beta}{}_{\lambda}\frac{\partial x^{\lambda}}{\partial x'^{\delta}}\right)\Gamma^{\gamma}_{\alpha\rho}(x).
\end{aligned}
$$

In particular, if we pick the natural basis set $(G_{\gamma}{}^{\rho})^{\beta}{}_{\lambda} = \delta^{\beta}_{\gamma}\delta^{\rho}_{\lambda}$ then () becomes the well known transformation law for the components $\Gamma^{\epsilon}_{\mu\delta}$ of an affine connection on $\mathcal{M}$:

$$
\Gamma'^{\epsilon}_{\mu\delta}(x) = \frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\rho}}{\partial x'^{\delta}}\frac{\partial x'^{\epsilon}}{\partial x^{\gamma}}\Gamma^{\gamma}_{\alpha\rho}(x) + \frac{\partial x'^{\epsilon}}{\partial x^{\lambda}}\frac{\partial^2 x^{\lambda}}{\partial x'^{\mu}\partial x'^{\delta}}.
$$

□

C.19.8 Parallel Transport in a Principal Bundle

Parallel transport was defined as transport without change.

What is parallel transport of an element of a principal bundle along a curve in $\mathcal{M}$?

We use the notion of a horizontal lift.

Horizontal vector fields are fields whose flow lines move from one fibre into another.

Definition Let P be a principal fibre bundle and let $\gamma[0,1] \to \mathcal{M}$ be a curve in $\mathcal{M}$. A curve $\tilde{\gamma}[0,1] \to P$ is said to a horizontal lift of γ if $\pi\tilde{\gamma} = \gamma$ and the tangent vector to $\tilde{\gamma}(t)$ always belongs to $\mathcal{H}_{\tilde{\gamma}(t)}P$

□

C.19.9 Curvature on a Principal Bundle

Suppose a connection is given on a principal bundle $\pi : P \to \mathcal{M}$, with group G. Then $X \in T_uP$ can be written uniquely as the sum of a vertical and a horizontal vector

$$
D\varphi \equiv (d\varphi)^H
$$

where

$$
(d\varphi)^H(\mathbb{X}_1,\dots,\mathbb{X}_{i+1}) \equiv d\varphi(\mathbb{H}_1,\dots,\mathbb{H}_{i+1})
$$

and $\mathbb{H}_1,\dots,\mathbb{H}_{i+1}$ are the horizontal vectors of $\mathbb{X}_1,\dots,\mathbb{X}_{i+1} \in T_uP$.

Lemma C.19.3 *If $A, B \in \mathcal{G}$, the map $\sharp$ preserves the Lie algebra structure:*

$$[A^\sharp, B^\sharp] = [A, B]^\sharp \tag{C.977}$$

Proof:

$$[A^\sharp, B^\sharp]_p = \lim_{t\to 0} \frac{1}{t}\{B^\sharp - (L_{g(t)})_* B^\sharp\}$$

$$\begin{aligned} [A^\sharp, B^\sharp]_p &= \lim_{t=0}\{(\sharp)_* B - (\sharp)_* \mathcal{A}d(g(t))B\} \\ &= (\sharp)_* \left(\lim_{t=0} \frac{1}{t}\{B - \mathcal{A}d(g(t))B\}\right) \end{aligned} \tag{C.978}$$

$$\mathcal{A}d(g(t))B = (L_{g(t)})_*(R_{g^{-1}(t)})_* B.$$

by $(R_{g^{-1}(t)})_* B = B$,

$$\mathcal{A}d(g(t))B = (L_{g(t)})_* B$$

$$[A^\sharp, B^\sharp]_p = (\sharp_p)_* \left(\lim_{t\to 0} \frac{1}{t}\{B - (L_{g(t)})_* B\}\right) = (\sharp_p)_*([A, B]) = [A, B]^\sharp(p),$$

□

C.19.10 Extension and Reduction of Principal Bundles

Given any bundle E we can construct a principal bundle $P(E)$, by replacing the fibres in E with the transition functions, while keeping the transition functions the same.

Let $H \subset G$ be a closed subgroup. P is reduced to a principal H bundle Q if:

(i) $Q \subset P$ is a submanifold,

(ii) $qh \in Q$ for all $q \in Q$, $h \in H$, such that

(1) $\pi(Q) = M$,

(2) H acts transitively in each fibre $Q_x = \pi^{-1}(x) \cap Q$, (remember $\pi^{-1}(x) \subset P$); by acts transitively we mean given any two points $p, q \in Q_x$ there exists at least one $h \in H$ such that $hp = q$.

$$\{v \in H(P)| < \omega|v >= 0\} \tag{C.979}$$

$$< dx^\nu|\frac{\partial}{\partial x^\mu} >= \delta_{\mu\nu}, \quad < dg_{kj}|\frac{\partial}{\partial g_{lm}} >= \delta_{kl}\delta_{jm}, \quad < dx^\nu|\frac{\partial}{\partial g_{lm}} >=< dg_{kj}|\frac{\partial}{\partial x^\mu} >= 0. \tag{C.980}$$

$$\begin{aligned} < \omega_{ij}|X_\mu > &= < i(g^{-1})_{ik}dg_{kj} + (g^{-1}A^a_\nu T^a g)_{ij}dx^\nu|\frac{\partial}{\partial x^\mu} + iB_{\mu lm}\frac{\partial}{\partial g_{lm}} > \\ &= (ig^{-1})_{ik} < dg_{kj}|\frac{\partial}{\partial g_{lm}} > iB_{\mu lm} + (g^{-1}A^a_\nu T^a g)_\nu < dx^\nu|\frac{\partial}{\partial x^\mu} > \\ &= (g^{-1}A^a_\mu T^a g - g^{-1}B_\mu)_{ij} \end{aligned} \tag{C.981}$$

Hence

$$B_{\mu ij} = \left(A^b_\mu T^b g\right)_{ij}. \tag{C.982}$$

$$D_\mu = \frac{\partial}{\partial x^\mu} + i\left(A^b_\mu T^b g\right)_{ij}\frac{\partial}{\partial g_{ij}} \tag{C.983}$$

$$D_\mu(gWg^{-1}) = \frac{\partial}{\partial x^\mu}\tilde{W} + iA^b_\mu T^b gWg^{-1} - igWg^{-1}A^b_\mu T^b = [\partial_\mu + A^a_\mu T^a, \tilde{W}], \tag{C.984}$$

$$ig^{-1}\mathbf{d}g + g^{-1}\mathbf{A}g = ig^{-1}h^{-1}\mathbf{d}(hg) + g^{-1}h^{-1}\mathbf{A}'hg. \tag{C.985}$$

This gives

$$\mathbf{A}' = -i\mathbf{d}hh^{-1} + h\mathbf{A}h^{-1} \tag{C.986}$$

$$[D_\mu, D_\nu] = i(\partial_\mu A_\nu - \partial_\nu A_\mu)g\frac{\partial}{\partial g} - A^a_\mu A^b_\nu\left(T^a g\frac{\partial}{\partial g}T^b g\frac{\partial}{\partial g} - T^b g\frac{\partial}{\partial g}T^a g\frac{\partial}{\partial g}\right). \tag{C.987}$$

$$[D_\mu, D_\nu] = i(\partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu, A_\nu])g\frac{\partial}{\partial g} = iF_{\mu\nu}g\frac{\partial}{\partial g} \tag{C.988}$$

Thus we see that the curvature of the principal bundle is associated with the field strength tensor.

C.19.11 The Complex Line Bundle

$$z = a + ib \quad \leftrightarrow \quad a\mathbf{e_1} + b\mathbf{e_2}, \text{ or in component form } \begin{pmatrix} a \\ b \end{pmatrix} \tag{C.989}$$

We multiply a complex number $z = a + ib$ by i, $-b + ia$

$$Je_1 = e_2, \quad Je_2 = -e_1, \quad J^2 = -I \tag{C.990}$$

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{C.991}$$

$$iz = -b + ia \quad \leftrightarrow \quad \begin{pmatrix} -b \\ a \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \tag{C.992}$$

$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y) \tag{C.993}$$

Of course, this can be expressed entirely in real terms

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \circ \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 x_2 - y_1 y_2 \\ x_1 y_2 + x_2 y_1 \end{pmatrix} \tag{C.994}$$

C.20 Summary of Differential Geometry

- Coordinate independent formalism.
- Anti-symmetric objects are important - unifying notations.

C.21 Summary

- Tensor Calculus
- Group theory

C.22 Biblioliographical notes

In this chapter I have relied on the following references:

Dana P. Williams notes on the Spectral Theorem for bounded normal operators

Introduction to topology and modern analysis G.F Simons.

Manifolds at Dr Neil Lambert Department of Mathematics King's College

C.23 Worked Exercises and Details

Commutativity of exterior derivative and $\mathcal{L}_\omega$.

Use the fact that $X^a_{;\,b}$ transforms as a tensor to deduce the transformation law of the connection given by (C.582).

Proof. $\Gamma^a_{cb} X^c = X^a_{;\,b} - X^a_{,b}$ we get

$$\Gamma'^a_{cb} X'^c = \frac{\partial x'^a}{\partial x^d} \frac{\partial x^e}{\partial x'^b} \Gamma^d_{fe} X^f - \frac{\partial x'^b}{\partial x^e} \frac{\partial^2 x^a}{\partial x'^f \partial x'^b} X^f. \tag{C.995}$$

Expressing X'^a in terms of X^f and since X^f is an arbitrary vector field, we have

$$\Gamma'^a_{cb} \frac{\partial x'^c}{\partial x^f} = \frac{\partial x'^a}{\partial x^d} \frac{\partial x^e}{\partial x'^b} \Gamma^d_{fe} - \frac{\partial x'^b}{\partial x^e} \frac{\partial^2 x^a}{\partial x'^f \partial x'^b}. \tag{C.996}$$

Multiplying through by $\partial x^c / \partial_g$ and rearranging the indices, we arrive at

$$\Gamma'^a_{cb} = \frac{\partial x'^a}{\partial x^d} \frac{\partial x^e}{\partial x'^b} \frac{\partial x^e}{\partial x'^c} \Gamma^d_{fe} - \frac{\partial x'^b}{\partial x^e} \frac{\partial x^e}{\partial x'^c} \frac{\partial^2 x^a}{\partial x'^f \partial x'^b}. \tag{C.997}$$

Commutativity of exterior derivative and $\mathcal{L}_\omega$.

C.23.1 Dynamical and Non-Dynamical Symmetries

Dynamical symmetries constrain the solutions of the equations of motion. *Non-Dynamical symmetries* are redundantcies of the mathematical formulation of the theory - these are often referred to as *Gauge symmetries*.

Example:

(a) Dynamical symmetry: Time translation invariance.

The Lagrangian does not explicitly depend on time

$$\mathcal{L} = \mathcal{L}(q_n, \dot{q}_n). \tag{C.998}$$

From this it follows that

$$\frac{d\mathcal{L}}{dt} = \frac{\partial \mathcal{L}}{\partial q_n}\dot{q}_n + \frac{\partial \mathcal{L}}{\partial \dot{q}_n}\ddot{q}_n \tag{C.999}$$

Now, at this point we make use of *the Euler-Lagrange equations of motion*

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_n}\right) - \frac{\partial \mathcal{L}}{\partial q_n} = 0 \tag{C.1000}$$

one finds

$$\begin{aligned} \frac{d\mathcal{L}}{dt} &= \dot{q}_n \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_n}\right) + \frac{\partial \mathcal{L}}{\partial \dot{q}_n}\ddot{q}_n \\ &= \frac{d}{dt}\left(\dot{q}_n \frac{\partial \mathcal{L}}{\partial \dot{q}_n}\right) \end{aligned} \tag{C.1001}$$

or

$$\frac{d}{dt}\left(\dot{q}_n \frac{\partial \mathcal{L}}{\partial \dot{q}_n} - \mathcal{L}\right) \tag{C.1002}$$

That is, the Hamiltonian is conserved in time

$$\mathcal{H} = \frac{\partial \mathcal{L}}{\partial \dot{q}^i}\dot{q}^i - \mathcal{L}. \tag{C.1003}$$

This reflects the conservation of energy E for isolated systems. Note that, as we needed to use the equations of motion this condition only holds for trajectories that extremalize the action.

(b) Non-dynamical symmetry: reparameterization invariance.

$$\begin{aligned}
S[q(t+\epsilon)] - S[q(t)] &= \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial t}\epsilon + \frac{\partial \mathcal{L}}{\partial q^i}\dot{q}^i\epsilon + \frac{\partial \mathcal{L}}{\partial \dot{q}^i}\frac{d}{dt}(\dot{q}^i\epsilon)\right] dt \\
&= \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial t}\epsilon + \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^i}\right)\frac{d\epsilon}{dt}\right] dt \\
&= [\mathcal{L}\epsilon]_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^i} - \mathcal{L}\right)\frac{d\epsilon}{dt}\, dt
\end{aligned} \tag{C.1004}$$

Parameterization-invariance means that the integral must vanish for arbitrary $d\epsilon/dt$, so that we have

$$\mathcal{H} = \frac{\partial \mathcal{L}}{\partial \dot{q}^i}\dot{q}^i - \mathcal{L} = 0. \tag{C.1005}$$

Even if the action is not extremal for some trajectory, it is still invariant under reparameterization of that trajectory.

Example: Diff (S^1) - the Virasoro algebra

As will be explained in chapter 5.24, in closed string theory there is an invariance under active diffeomorphisms acting on a circle, the corresponding group is denoted Diff (S^1). If $\theta \in (0, 2\pi]$ is a coordinate on S^1,

$$\tau(\theta) \to \tau'(\theta) = \tau(\theta) + f(\tau(\theta)) \tag{C.1006}$$

where f is periodic, i.e., $f(\tau + 2\pi) = f(\tau)$. There is a complete Fourier series expansion

$$f(\tau(\theta)) = \sum_{n=0}^{\infty} A_n \cos n\tau(\theta) + B_n \sin n\tau(\theta)$$

For an infinitesmal form transformation

$$\tau(\theta) \to \tau'(\theta) = \tau(\theta) + V(\theta)\frac{\partial}{\partial\theta}(\tau(\theta)) \tag{C.1007}$$

where $V(\theta)$ is the vector generating the infinitesmal diffeomorphism. As $V(\theta) = V(\theta + 2\pi)$ we can expand $V(\theta)$,

$$
\begin{aligned}
V(\theta)\frac{\partial}{\partial\theta} &= (\sum_{n=0}^{\infty} b_n \cos n\theta + c_n \sin n\theta)\frac{\partial}{\partial\theta} \\
&= \sum_{n=0}^{\infty}\big[\frac{(b_n - ic_n)}{2}e^{in\theta} + \frac{(b_n + ic_n)}{2}e^{-in\theta}\big]\frac{\partial}{\partial\theta} \\
&= \sum_{n=-\infty}^{\infty} a_n i e^{in\theta}\frac{\partial}{\partial\theta} \qquad \text{(C.1008)}
\end{aligned}
$$

where $a_n = (c_n - ib_n)/2$ and $a_{-n} = a_n^*$ for $n > 0$.

$$
\tau'(\theta) = \tau(\theta) + \sum_{n\in\mathbb{Z}} a_n i e^{in\theta}\theta^\alpha\partial_\alpha\tau \qquad \text{(C.1009)}
$$

Therefore one has a basis indexed by $n \in \mathbb{Z}$

$$
D_n = ie^{in\theta}\frac{\partial}{\partial\theta} \qquad \text{(C.1010)}
$$

The Lie algebra of Diff (S^1) is

$$
\begin{aligned}
[D_n, D_m]f &= [ie^{in\theta}\theta^\alpha\partial_\alpha ie^{im\theta}\theta^\beta\partial_\beta - ie^{im\theta}\theta^\alpha\partial_\alpha ie^{in\theta}\theta^\beta\partial_\beta]f \\
&= e^{i(n+m)\theta}\big[(in - im)\theta^\alpha\partial_\alpha f + i(\theta^\alpha\partial_\alpha\theta^\beta\partial_\beta - \theta^\beta\partial_\beta\theta^\alpha\partial_\alpha)f\big] \\
&= (n-m)D_{m+n}f \qquad \text{(C.1011)}
\end{aligned}
$$

or

$$
[D_n, D_m] = (n-m)D_{m+n}. \qquad \text{(C.1012)}
$$