Notes

Martin Kjøllesdal Johnsrud

December 5, 2022

Contents

Todo list

1 Differential geometry

Integrate old and new notes, fix citations

1.1 Old Notes

Topology

Based on Carroll and In general relativity, space-time is modeled as a *differentiable manifold* $\mathcal{M}$. A differentiable manifold is a mathematical space which we may draw coordinates on, enabling us to do calculus. It generalizes the familiar notion of euclidean space. Manifolds are a special kind of topological space, so this is the first structure we need to define.

A topological space is a set of points $p \in S$, as well as a topology $\mathcal{T}$. The topology contains *open sets* of S. Both S itself, and the empty set $\emptyset = \{\}$ must be part of the topology $\mathcal{T}$. Open sets must obey the rule that the union of two open sets again is an open set, i.e. $\{U_i\}_i \subseteq \mathcal{T}, \implies \cup_i U_i \in \mathcal{T}$. All *finite intersections* of open sets be open sets, i.e. $\{U_i\}_{i=1}^N \subseteq \mathcal{T}, \implies \cap_i U_i \in \mathcal{T}$. The structure of the topology allows one to talk about a neighborhood of a point $p \in S$, and define a homeomorphism. [1]

Definition. *A* neighbourhood *of a point p is a open set U which contain p.*

As S is an open set, all points have a neighbourhood. to define homeomorphism, we need a notion of a continuous function. The definition of continuous coincide with the regular ϵ, δ definition from calculus when the topological space in question is $\mathbb{R}$ and the topology is open line intervals. However, the topology allows for a more general definition of continuous function:

Definition. *A function between two topological spaces $f : S \to T$ is* continuous *if and only if, for every open set $V \subseteq T$, the pre-image $f^{-1}(V) = U \subseteq S$ is open.*

Definition. *A* homeomorphism *$f : S \to T$ is then a continuous, bijective function whose inverse f^{-1} also is continuous.*

Manifolds and tensors

A manifold $\mathcal{M}$ is a topological space with additional structure. All points $p \in S$ are locally *homeomorphic* to euclidian space $\mathbb{R}^n$. A open set $U \in \mathcal{T}$ and a corresponding homeomorphism

$$x : U \longrightarrow \mathbb{R}^n,$$

called a coordinate function, together form a chart (x, U). This chart takes points $p \in U$ and gives them coordinates $x(p) = (x_1(p), \dots x_n(p)) \in \mathbb{R}^n$. A set $\mathcal{A} = \{(x^i, U_i)\}$ such that $\{U_i\}$ cover S is an atlas. A manifold $\mathcal{M}$ is thus a topological space together with the maximal atlas, i.e. the larges possible atlas. To get a differentiable manifold, we demand that the transition functions

$$\tau^{i,j} = x^j \circ (x^i)^{-1} : x^i(U_i \cap U_j) \longrightarrow x^j(U_i \cap U_j)$$

of the atlas are infinitely differentiable, or smooth.

Let $\gamma : \mathbb{R} \to \mathcal{M}$ be a curve in the manifold. This may be used to define a directional derivative of functions $f : \mathcal{M} \to \mathbb{R}$ on the manifold. Let λ be the parameter associated with γ. The derivative is then

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda} = \frac{\mathrm{d}\lambda}{\mathrm{d}(} f \circ \gamma)(\lambda).$$

For a given point $p \in \mathcal{M}$, and form an equivalence class of curves $[\gamma]_p$ that gives the same directional derivative at p.[2] The set of all such derivative operators form a vector space, the tangent space

$$T_p = \left\{ \frac{\mathrm{d}}{\mathrm{d}\lambda}(\_ \circ \gamma)(\lambda_0) \,\middle|\, [\gamma]_p(\lambda_0) = p \right\}.$$

[1] Not to be confused with homomorphism.
[2] Called a germ

Using the chain rule, the coordinate functions x^μ induce a basis for the tangent space:

$$\frac{\mathrm{d}f(\gamma(\lambda))}{\mathrm{d}\lambda} = \left(\frac{\mathrm{d}}{\mathrm{d}\lambda}\right)_{(f\circ x^{-1})(x\circ\gamma)}(\lambda) = \frac{\partial x^\mu}{\partial(}f\circ x^{-1})\frac{\mathrm{d}}{\mathrm{d}\lambda}(x\circ\gamma)^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\partial_\mu f.$$

Thus, $V^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$ are the components of the vector, and $e_\mu = \partial_\mu$ the basis. If we change coordinates on the manifold, $x \to x'$, we get a transition function $\tau = x \circ (x')^{-1}$. This coordinate change induces a basis change in the tangent space, yielding new basis vectors

$$\partial'_\mu = \frac{\partial}{\partial x'^\mu} = \frac{\partial \tau^\nu(x')}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} = \frac{\partial}{\partial x^\nu}x'^\mu\partial_\nu.$$

This again induces a change of the components,

$$V'^\mu = \frac{\partial x'^\mu}{\partial x^\nu}V^\nu,$$

ensuring that

$$V'^\mu e'_\mu = V^\nu\frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\rho}{\partial x'^\mu}\partial_\rho = V^\mu e_\mu.$$

We saw that a function $f : \mathcal{M} \to \mathbb{R}$ has, at $p \in \mathcal{M}$, the directional derivative

$$\left.\frac{\mathrm{d}f}{\mathrm{d}\lambda}\right|_{\lambda_0} = V^\mu\frac{\partial f}{\partial x^\mu}$$

This means, with some more work, we have a map from the tangent space to the real numbers. Such a map forms, due to Riez's representation theorem, a basis for the to the vector space. Let $\omega = [f]$ be an equivalence class of all functions with the same directional derivative at p in all directions, i.e. the same gradient, parametrized such that $\gamma(\lambda_0) = p$. This gives the map $\omega : T_p \to \mathbb{R}$, defined by

$$\omega(V) = \frac{\partial f}{\partial x^\mu}\left.\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\right|_{\lambda_0}$$

meaning we have found an element of the dual space of the tangent space, $\omega \in T_p^*$. This element is denoted as the differential $\omega = \mathrm{d}f$, and is called a covector or covariant vector.

We then add a inner product to the tangent space [3]

$$\begin{aligned}
&\langle\cdot,\cdot\rangle : T_p \otimes T_p \longrightarrow \mathbb{R},\\
&\langle\lambda v + u, w\rangle = \lambda\langle v, w\rangle + \langle u, w\rangle,\ \lambda \in \mathbb{R},\\
&\langle v, u\rangle = \langle u, v\rangle
\end{aligned}$$

This induces a basis on our dual vector space, namely

$$\mathrm{d}x^\mu(\partial_\nu) = \langle\cdot,\partial_\mu\rangle(\partial_\nu) = \delta^\mu{}_\nu \implies \omega_\mu\mathrm{d}x^\mu(V) = V^\nu\omega_\mu\mathrm{d}x^\mu(\partial_\nu) = V^\nu\omega_\nu$$

With $V^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$, the components must be $\omega_\mu = \partial\mu f)$, and we get

$$\mathrm{d}f = \frac{\partial f}{\partial x^\mu}\mathrm{d}x^\mu, \quad \mathrm{d}f\left(\frac{\mathrm{d}}{\mathrm{d}\lambda}\right) = \frac{\partial f}{\partial x^\mu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$$

Linearity gives us the transformation rule of the covectors,

$$\mathrm{d}x'^\mu(\partial'_\nu) = \left\langle\frac{\partial x^\rho}{\partial x'^\nu}\partial_\rho, A^\mu{}_\sigma\partial_\sigma\right\rangle = \delta^\mu{}_\nu \implies \mathrm{d}x'^\mu = \frac{\partial x'^\mu}{\partial x^\rho}\mathrm{d}x^\rho, \quad \omega'_\rho = \frac{\partial x^\rho}{\partial x'^\mu}\omega_\rho$$

With a vector space and its dual, we may create new vector spaces through the tensor product. The tensor

$$T \in T_p \otimes T_p \otimes ...T_p^* \otimes T_p^* \otimes ... = \left(\bigotimes_{i=1}^{n} T_p\right)\left(\bigotimes_{i=1}^{m} T_p^*\right) = T_p^{(n,m)}$$

is an (n, m) rank tensor, which is a map that takes n vectors and m covectors, i.e. $T : T_p^{(m,n)} \to \mathbb{R}$. The components of a tensor in the coordinate basis are

$$T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\ldots} = T(\mathrm{d}x^{\mu_1}, \mathrm{d}x^{\mu_1}, ...\partial_{\nu_1}, \partial_{\nu_2}, ...)$$

[3] No requirement of poistive definitness is enforced here.

The most important tensor is the metric tensor, $g \in T_p^* \otimes T_p^*$, corresponding to the inner product.

$$g_{\mu\nu} = g(\partial_\mu, \partial_\nu) = \langle \partial_\nu, \partial_\mu \rangle, \quad ds^2 := g = g_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu$$

This tensor gives a correspondence between the components of vectors and covectors which inspires the definition of raising and lowering of indices:

$$g(A, B) = g_{\mu\nu} A^\mu B^\nu := A_\mu \mathrm{d}x^\mu(B), P \quad A_\mu = g_{\mu\nu} A^\nu,\ A^\nu = \delta^\nu_\mu A^\mu = g^{\nu\mu} A_\mu.$$

The anti-symmetric part of the $(0, p)$-tensor space, corresponding to a vector space V (in our case T_p), are denoted $\Lambda^p(V)$, and are called p-forms. Firstly, we define anti-symmetrization of a tensor:

$$T_{[\mu_1, \ldots \mu_p]} = \sum_{\sigma \in S_p} (-1)^{\mathrm{sgn}(\sigma)} T_{\mu_{\sigma(1)}, \ldots \mu_{\sigma(p)}}.$$

Thus, all p-forms are written as an anti-symmetrized tensor. The wedge-product is defined

$$A \wedge B = \frac{(p+q)!}{p!q!} A_{[\mu_1 \ldots \mu_p} B_{\mu_{p+1} \ldots \mu_{p+q}]}$$

The exeterior derivative acts on p-forms, and is defined

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

This leads to the properties

$$\mathrm{d}(\mathrm{d}\omega) = 0, \quad \mathrm{d}(\omega \wedge \eta) = (\mathrm{d}\omega) \wedge \eta + (-1)^p \omega \wedge (\mathrm{d}\eta), \quad \omega \in \Lambda^p,\ \eta \in \Lambda^q.$$

The Levi-Civita symbol is defined by the permutations $\sigma \in P(n)$

$$\epsilon_{\mu_{\sigma(1)} \cdots \mu_{\sigma(n)}} = \mathrm{sign}(\sigma),$$

and is zero elsewhere. It defines the determinant, as

$$\epsilon_{\nu_1 \ldots} \det(M^\mu{}_\nu) = \epsilon_{\mu_1 \ldots} M^{\mu_1}{}_{\nu_1} \ldots \implies \det(M^\mu{}_\nu) = \epsilon_{\mu_1 \ldots} \epsilon^{\nu_1 \cdots} M^{mu_1}{}_{\nu_1} \ldots$$

as the determinant of the metrix $\det(g) = |g|$ transforms as $|g(x')| = |\frac{\partial *}{\partial x} x'|^2 |g(x)|$, we can define the Levi-Civita tensor as

$$\epsilon = \sqrt{|g|} \epsilon_{\mu_1 \ldots} \mathrm{d}x^{\mu_1} \otimes \ldots = \sqrt{|g|} \mathrm{d}x^0 \wedge \mathrm{d}x^1 \ldots \wedge \mathrm{d}x^{n-1} = \sqrt{|g|} d^n x,$$

which you might recognize as the integration measure in nD.

Curvature

A vector field on the manifold takes a point $p \in \mathcal{M}$, and returns a vector in the tangent space of that point, $V(p) = V^\mu(p) e_\mu(p) \in T_p$. This is usually thought of as a function of the coordinates through $V^\mu(x) = (V^\mu \circ x^{-1})(x^\mu)$, which means we may take the partial derivative of it: $\partial_\nu V^\mu(x)$. However, when changing coordinates it becomes clear that this is not a tensor as it does not follow the same transformation rule, due to the fact that the transformation matrix $\frac{\partial x'^\mu}{\partial x^\nu}$ itself is a function of the coordinates. We define the covariant derivative of a vector and covector field as

$$\nabla_\nu V^\mu = \partial_\nu V^\mu + \Gamma^\mu_{\nu\lambda} V^\lambda, \quad \nabla_\nu \omega_\mu = \partial_\nu \omega_\mu - \Gamma^\lambda_{\mu\nu} \omega_\lambda$$

If we demand a torsion free ($\Gamma^\mu_{\nu\lambda} = \Gamma^\mu_{(\nu\lambda)}$), metric compatible $\nabla_\lambda g_{\nu\mu} = 0$ connection, we get a formula for the Christoffel symbols

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

This gives us a notion of the parallel transport of a vector along a path $\frac{\mathrm{d}*}{\mathrm{d}x^\mu} \lambda = (x^\mu \circ \gamma)(\lambda)$, and thus relating vectors in nearby tangent spaces, using the demand

$$\frac{\mathrm{d}\lambda}{\mathrm{d}V}^\mu (x(\lambda)) = \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} \nabla_\nu V^\mu = 0.$$

Setting $\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$ as the vector, this gives us a differential equation for a special set of curves, geodesics

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\nu\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\rho}{\mathrm{d}\lambda} = 0.$$

The Riemann curveature tensor may equivalently defined as either encoding the change δV^μ in a vector V^μ transported parallel transported along the vecotrs S^μ, B^ν and back, or from the commutator of the covariant derivative: [4]

$$\delta^\rho{}_\sigma = R^\rho{}_{\sigma\mu\nu}V^\sigma A^\mu B^\nu, \quad [\nabla_\mu, \nabla_\nu]V^\rho = R^\mu{}_{\nu\rho\sigma}V^\sigma.$$

These lead to the equation for the tensor,

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma} = \partial_{[\mu}\Gamma^\rho_{\nu]\sigma} + \Gamma^\rho_{\lambda[\mu}\Gamma^\lambda_{\nu]\sigma}$$

Evaluating it in normal coordinates shows the Riemann tensor has the following symmetries:

$$R_{[\rho\sigma]\mu\nu} = R_{\rho\sigma[\mu\nu]} = R_{(\{\rho\sigma\}\{\mu\nu\})} = R_{\rho\sigma\mu\nu}, \quad R_{\rho[\sigma\mu\nu]} = \nabla_{[\lambda}R_{\rho\sigma\mu]\nu} = 0.$$

Curly brackets indicates to interchange of the indices within as a group. The information of the trace of the tensor is contained in the Ricci tensor and scalar, also contained in the Einstein tensor:

$$R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu}, \quad R = R^\mu{}_\mu, \quad G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$$

The last of the identities, the Bianchi identity, gives the inspiration for the field equation of gravity as

$$\nabla^\mu G_{\mu\nu} = 0.$$

Removing all trace while retaining the symmetries gives the Weyl tensor, in n dimensions:

$$C_{\rho\sigma\mu\nu} = R_{\rho\sigma\mu\nu} - \frac{4}{n-2}g_{\rho]\lbrack\mu}R_{\nu]\lbrack\sigma} + \frac{2R}{(n-1)(n-2)}g_{\rho[\mu}g_{\nu]\sigma}.$$

1.2 Differential Geometry (from master)

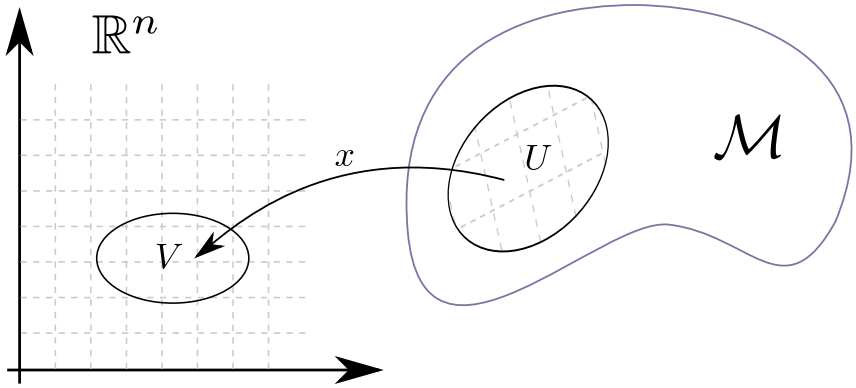


Figure 1: The coordinate function x maps a neighborhood U in the manifold $\mathcal{M}$ to a neighborhood V in $\mathbb{R}^n$.

This section is taken from my matster's thesis, and is based on

Differential geometry generalizes n-dimensional calculus to more general spaces than the usual $\mathbb{R}^n$, such as curved spacetime or the more abstract space of symmetries of a quantum field theory. The most important objects in differential geometry are *smooth manifolds*. An n-dimensional manifold, $\mathcal{M}$, is a set of points, locally homeomorphic to $\mathbb{R}^n$. That is, for all points $p \in \mathcal{M}$, there exists a neighborhood U around p, together with a corresponding set of continuous, bijective functions that map U to a neighborhood V in $\mathbb{R}^n$,

$$x : U \subseteq \mathcal{M} \longmapsto V \subseteq \mathbb{R}^n, \tag{1}$$

$$p \longmapsto x^\mu(p), \quad \mu \in \{0, \ldots, n-1\}. \tag{2}$$

[4] We consider a torsion free connection.

This is illustrated in Figure 1. We call $x(p) = (x^0(p), \ldots, x^{n-1}(p))$ a coordinate function of $\mathcal{M}$. The inverse of x, x^{-1}, obeys $x^{-1}(x(p)) = p$, for all $p \in U$. A smooth manifold is one in which the coordinate functions are infinitely differentiable. To define differentiability on manifolds, consider two coordinate functions, x, and x'. The corresponding domains U and U' may or may not overlap. We then define the transition function, a function between subsets of $\mathbb{R}^n$ by mapping via $\mathcal{M}$, as

$$f_{x' \to x} = x \circ {x'}^{-1} : \mathbb{R}^n \mapsto \mathbb{R}^n. \tag{3}$$

This map is illustrated in Figure 2.[5] A set of coordinate functions $\mathcal{A} = \{x_i\}$ whose domain cover $\mathcal{M}$ is called an *atlas* of $\mathcal{M}$. If the transition function between all pairings of coordinate functions in the atlas is smooth—that is, infinitely differentiable—we call the atlas smooth. We then define a smooth manifold as the topological manifold $\mathcal{M}$ together with a *maximal* smooth atlas $\mathcal{A}$. A smooth atlas is maximal if no coordinate function can be added while the atlas remains smooth.[6]

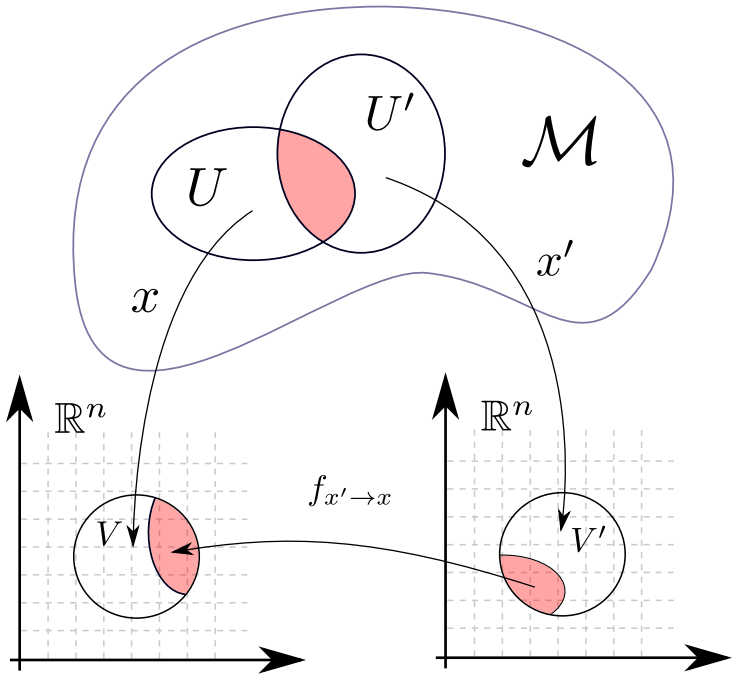


Figure 2: The transition map $f_{x' \to x}$ between two coordinate functions, x' and x, maps between the images of these function, via the manifold $\mathcal{M}$. The function's domain and image are restricted to a (possibly empty) subset of the images of x' and x. This is illustrated by the shaded regions in V' and V.

Consider two m- and n-dimensional smooth manifolds $\mathcal{M}$ and $\mathcal{N}$. Let x denote the coordinates on $\mathcal{M}$, while y denotes the coordinates on $\mathcal{N}$. We can define smooth functions between these manifolds similarly to how we define smooth coordinates. Consider the function

$$F : \mathcal{M} \longmapsto \mathcal{N}. \tag{4}$$

It is said to be smooth if, for all points $p \in \mathcal{M}$, there is a set of local coordinates x around p and y around $F(p)$ such that the map $\tilde{F} = y \circ F \circ x^{-1}$ is smooth. This map may be illustrated by a diagram,

$$\begin{array}{ccc} \mathcal{M} & \xrightarrow{F} & \mathcal{N} \\ \downarrow{\scriptstyle x} & & \downarrow{\scriptstyle y} \\ \mathbb{R}^m & \xrightarrow{\tilde{F}} & \mathbb{R}^n \end{array} \tag{5}$$

We will not be careful with the distinction between F, the function between the abstract manifolds, and $\tilde{F}$, the function of their coordinates, but rather denote both by $F(x)$. We may take the partial derivative of such a function with respect to the coordinates x, $\frac{\partial F}{\partial /} x^\mu$. However, this is dependent on our choice of

[5] To be rigorous, one has to restrict the domains and image of the coordinate function when combining them. This is illustrated in Figure 2.

[6] The maximal condition ensures that two equivalent atlases correspond to the same differentiable manifold. A single manifold can be combined with different maximal atlases of smooth coordinates or differentiable structures. A set of examples are *exotic spheres*, smooth manifolds *homeomorphic* to S^n, but not *diffeomorphic*.

coordinates, as a set of local coordinates can always be scaled arbitrarily. Any physical theory must be independent of our choice of coordinates, so our next task is to define the properties of a smooth manifold in a coordinate-independent way.

1.3 Vectors and tensors

A curve γ through $\mathcal{M}$ is a function from $\mathbb{R}$ to $\mathcal{M}$,

$$\gamma : \mathbb{R} \longmapsto \mathcal{M}, \tag{6}$$
$$\lambda \longmapsto \gamma(\lambda). \tag{7}$$

Such curves are often denoted only by their coordinates and the parameter λ, $x^\mu(\lambda) = (x^\mu \circ \gamma)(\lambda)$. With this curve, we can take the directional derivative of a real-valued function on the manifold, $f : \mathcal{M} \mapsto \mathbb{R}$. Assume $\gamma(\lambda = 0) = p$. As we are always taking the derivative of functions between $\mathbb{R}^n$, for different n, we can use the chain rule. The directional derivative of f at p, given by this curve γ, is then

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} f(x(\lambda))\bigg|_p = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\bigg|_{\lambda=0} \frac{\partial}{\partial x^\mu} f(x)\bigg|_p . \tag{8}$$

The set of all such directional derivatives, $\frac{\mathrm{d}}{\mathrm{d}\lambda}\lambda$ at p, form a vector space, $T_p\mathcal{M}$, called the *tangent space*. The tangent space is illustrated in Figure 3.

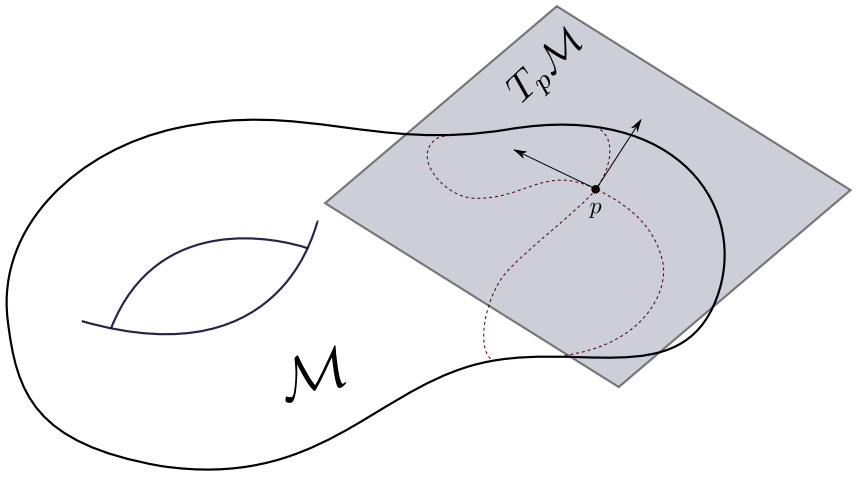


Figure 3: The tangent space $T_p\mathcal{M}$, the shaded rectangle, is the sett of all directional derivatives at $p \in \mathcal{M}$. A directional derivative is defined in terms of a curve that passes through p.

The coordinates x^μ induce a basis of this vector space, namely partial derivatives with respect to the coordinate functions at p,

$$e_\mu = \frac{\partial}{\partial x^\mu}\bigg|_p = \partial_\mu|_p, \quad \mu \in \{0, ...n-1\}. \tag{9}$$

Any element $v \in T_p\mathcal{M}$ can therefore be written

$$v = v^\mu \partial_\mu|_p = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\bigg|_{\lambda=0} \frac{\partial}{\partial x^\mu}\bigg|_p . \tag{10}$$

Here, λ is the parameter of the curve corresponding to the directional derivative v.[7] The evaluation at $\lambda = 0$ and p will often be implicit for ease of notation. This directional derivative acts on functions $f : \mathcal{M} \mapsto \mathbb{R}$ as

$$v(f) = v^\mu \partial_\mu f. \tag{11}$$

A map F between two manifolds $\mathcal{M}$ and $\mathcal{N}$ also induces a map between the tangent spaces of these manifolds. This is the *differential* of F at p,

$$\mathrm{d}F_p : T_p\mathcal{M} \longmapsto T_p\mathcal{N}, \tag{12}$$
$$v \longmapsto \mathrm{d}F_p(v). \tag{13}$$

[7]There is not only one curve corresponding to any directional derivative but rather an equivalence class. We will gloss over this technicality, as it does not affect our work.

$\mathrm{d}F_p(v)$ is an element of $T_p\mathcal{N}$, i.e., it is a directional derivative on $\mathcal{N}$. It is defined by how it acts on functions $g : \mathcal{N} \mapsto \mathbb{R}$,

$$\mathrm{d}F_p(v)(g) = v(g \circ F), \tag{14}$$

It thus acts on functions on $\mathcal{N}$ by "extending" the derivative v. This is a linear map between vector spaces and may be written in component form by considering the differentials of the coordinate functions. Denote the coordinates of $\mathcal{N}$ by y^μ, and $y^\mu \circ F = F^\mu$. Then,

$$\mathrm{d}F_p(\partial_\mu)(g) = \partial_\mu(g \circ F)|_p = \frac{\partial F^\nu}{\partial x^\mu}\Big|_p \frac{\partial g}{\partial y^\nu}\Big|_{F(p)}, \tag{15}$$

or more suggestively

$$\mathrm{d}F\left(\frac{\partial}{\partial x^\mu}\right) = \frac{\partial F^\nu}{\partial x^\mu}\frac{\partial}{\partial y_\nu}. \tag{16}$$

This is a linear map of vectors between two vectors by the matrix $A_\mu{}^\nu = \partial_\mu F^\nu$. The differential is thus a generalization of the Jacobian. In the case of a real valued function, $f : \mathcal{M} \mapsto \mathbb{R}$, and $g : \mathbb{R} \mapsto \mathbb{R}$, we get

$$\mathrm{d}f(v)(g) = v(g \circ f) = (v^\mu \partial_\mu f)\,\frac{\mathrm{d}g}{\mathrm{d}y}. \tag{17}$$

$\mathrm{d}f$ is thus a map from $T_p\mathcal{M}$ to $T_{f(p)}\mathbb{R}$, which is isomorphic to $\mathbb{R}$. Let g be the identity function, so that $\frac{\mathrm{d}g}{\mathrm{d}/}y = 1$. Then, the differential of a scalar function, also called a 1-form, is a map from vectors v to real numbers,

$$\mathrm{d}f(v) := v^\mu \partial_\mu f. \tag{18}$$

The set of all linear maps from a vector space V to the real numbers is called the *dual space* of V, denoted V^*. This is a new vector space with the same dimensionality as V. We denote the dual of $T_p\mathcal{M}$ as $T_p^*\mathcal{M}$. We can regard each coordinate function as a real-valued function with a corresponding differential. This differential obeys

$$\mathrm{d}x^\mu(\partial_\nu) = \frac{\partial x^\mu}{\partial x_\nu} = \delta^\mu_\nu. \tag{19}$$

The differentials of the coordinate functions thus form a basis for $T_p^*\mathcal{M}$, called the dual basis. Any differential $\mathrm{d}f$ can thus be written as $\mathrm{d}f = \omega_\mu \mathrm{d}x^\nu$ for some components ω_μ. We finde the components by applying the differential to the coordinate basis, $\mathrm{d}f(\partial_\mu) = \partial_\mu f = \omega_\mu$. In other words, we recover the classical expression

$$\mathrm{d}f = \frac{\partial f}{\partial x^\mu}\mathrm{d}x^\mu, \tag{20}$$

however we now interpret it as a covector-field instead of an "infinitesimal displacement".

Linear maps from vectors to real numbers is generalized by *tensors*. Given a vector space V, a general (n, m) tensor T is a multilinear map, which associates n elements from V and m from its dual V^* to the real numbers, i.e.,

$$T : V \times V \times \cdots \times V^* \times \ldots \longmapsto \mathbb{R}, \tag{21}$$

$$(v, u \ldots; \omega, \ldots) \longmapsto T(v, u, \ldots; \omega, \ldots). \tag{22}$$

Multilinear means that T is linear in each argument. The set of all such maps is the tensor product space $V \otimes V \otimes \cdots \otimes V^* \otimes \ldots$, a $\dim(V)^{n+m}$-dimensional vector space. If $\{e_\mu\}$ and $\{e^\mu\}$ are the basis for V and V^*, then we can write the basis of this of the tensor product space as $\{e_\mu \otimes \cdots \otimes e^\nu \otimes \ldots\}$. The tensor can thus be written

$$T = T^{\mu\nu\ldots}{}_{\rho\ldots}\, e_\mu \otimes e_\nu \otimes \ldots e^\rho \otimes \ldots, \quad T^{\mu\nu\ldots}{}_{\rho\ldots} = T(e^\mu, e^\nu, \ldots; e_\rho, \ldots). \tag{23}$$

We often want to decompose a tensor down into its symmetric and antisymmetric parts. To do this, we introduce the symmetrization of a tensor T,

$$T_{(\mu_1\ldots\mu_n)} = \frac{1}{n!}\sum_{\sigma\in S_n} T_{\mu_{\sigma(1)}\cdots\mu_{\sigma(n)}}, \tag{24}$$

where S_n is the set of all permutations of n objects. The antisymmetrization of a tensor is defined as

$$T_{[\mu_1\ldots\mu_n]} = \frac{1}{n!}\sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) T_{\mu_{\sigma(1)}\cdots\mu_{\sigma(n)}}. \tag{25}$$

The function $\sigma = \pm 1$, depending on if σ is a even or odd permutation. We may now write

$$T_{\mu\nu} = T_{(\mu\nu)} + T_{[\mu\nu]}. \tag{26}$$

1.4 Geometry and the metric

The metric is a symmetric, non-degenerate $(0,2)$ tensor

$$\mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu. \tag{27}$$

It defines the geometry of the manifold $\mathcal{M}$ and is the main object of study in general relativity. As it is invertible, we can define $g^{\mu\nu} = (g^{-1})_{\mu\nu}$, which is the components of a $(2,0)$ tensor. We use this to raise and lower indices, as is done with the Minkowski metric $\eta_{\mu\nu}$ in special relativity.

Up until now, we have only considered the tangent space $T_p\mathcal{M}$ at a point p and the corresponding tensor-product spaces. We are, however, more interested in *fields* of vectors, covectors, or tensors. For each point $p \in \mathcal{M}$, a tensor field T "picks out" a tensor $T(p)$ from each tensor product space corresponding to the tangent space at p, $T_p\mathcal{M}$. A vector field can be written as

$$v(p) = v^\mu(p)\partial_\mu|_p. \tag{28}$$

We will mostly be working with the components v^μ, which are functions of $\mathcal{M}$. For ease of notation, we write the vector as a function of the coordinates x. The vector field $v(x)$ is unchanged by a coordinate-transformation $x^\mu \to x'^\mu$; the coordinates are only a tool for our convenience. However, with a new set of coordinates, we get a new set of basis vectors, ∂'_μ:

$$v = v^\mu\partial_\mu = v^\mu\frac{\partial x'^\nu}{\partial x^\mu}\partial'_\nu = v'^\mu\partial'_\mu, \tag{29}$$

This gives us the transformation rules for the components of vectors,

$$v'^\mu = \frac{\partial x'^\mu}{\partial x^\nu}v^\nu. \tag{30}$$

Tangent vectors are also called *contravariant* vectors, as their components transform contra to the basis vectors. For covectors, it is

$$\omega'_\mu = \omega_\nu\frac{\partial x^\nu}{\partial x'^\mu}, \tag{31}$$

which is why covectors also are called *covariant* vectors.

The gradient of a scalar function f, $\mathrm{d}f = \partial_\mu f\mathrm{d}x^\mu$, is a coordinate-independent derivative, as $\partial_\mu f$ follows the transformation law for covectors. To generalize this, we define the covariant derivative, ∇, as a map from (n,m) tensor fields to $(n,m+1)$ tensor fields, as $f \to \mathrm{d}f$ maps a $(0,0)$ tensor, a scaalar, to a $(0,1)$- tensor The components of a covariant derivative, $\nabla_\rho T^{\mu_1\cdots}{}_{\nu_1,\ldots}$, must follow the tensor transformation law. However, this is not strong enough to uniquely define ∇. In addition to $\nabla f = \partial f$, we further assume the derivative is linear, $\nabla(T+S) = \nabla T + \nabla S$, and follow the product rule: $\nabla(T\otimes S) = (\nabla T)\otimes S + T\otimes(\nabla S)$. Lastly, we assume the derivative of the Kronecker delta gives zero, $\nabla_\mu\delta^\rho_\nu = 0$. With this, we can, in general, write the covariant derivative for vectors and covectors as [**carrollSpacetimeGeometryIntroduction2019**]

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\mu_{\nu\rho}v^\rho, \tag{32}$$

$$\nabla_\mu\omega_\nu = \partial_\mu\omega_\nu - \Gamma^\rho_{\mu\nu}\omega_\rho. \tag{33}$$

$\Gamma^\mu_{\nu\rho}$ are called *Christoffel symbols*. The generalization for higher-order tensors is straightforward,

$$\nabla_\mu T^{\nu\cdots}{}_{\rho\ldots} = \partial_\mu T^{\nu\cdots}{}_{\rho\ldots} + \Gamma^\mu_{\nu\lambda}T^{\lambda\cdots}{}_{\rho\ldots} + \cdots - \Gamma^\lambda_{\mu\rho}T^{\mu\cdots}{}_{\lambda\ldots} - \ldots. \tag{34}$$

This is still not enough to uniquely determine the covariant derivative. We will furthermore assume $\Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{\nu\mu}$ and $\nabla_\mu g_{\nu\rho} = 0$. With these assumptions, we find an explicit formula of the Christoffel symbols in terms of the metric,

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}g^{\rho\sigma}(\partial_\mu g_{\nu\sigma} - \partial_\sigma g_{\mu\nu} + \partial_\nu g_{\sigma\mu}). \tag{35}$$

Using the notion of a covariant derivative, we may also generalize *parallel transport* to curved spaces. The notion of parallel transport of a vector in flat $\mathbb{R}^n$ is intuitive—given a line $x^\mu(\lambda)$, a vector v^μ at $x^\mu(\lambda_0)$ is parallel transported to v'^μ at $x^\mu(\lambda_1)$ if you carry it along the line without "turning it". To make this more precise, a vector field v^μ is parallel transported along $x^\mu(\lambda)$ if $\frac{\mathrm{d}}{\mathrm{d}\lambda}v^\mu = \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\partial_\nu v^\mu = 0$. We generalize

this to curved spaces by replacing the partial derivative with a covariant derivative, and so the criterion for parallel transport is

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\nabla_\mu v^\nu = 0. \tag{36}$$

With this, we can imagine creating a special class of paths, called *geodesics*, namely those which parallel transport their tangent vectors $\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$. We imagine following an arrow we are holding without turning it as we walk. Using the definition of parallel transport Equation 36, together with the covariant derivative Equation 32, we get the geodesic equation,

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\rho\sigma}\frac{\mathrm{d}x^\rho}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\sigma}{\mathrm{d}\lambda} = 0. \tag{37}$$

In a flat space, where the Christoffel symbols vanish, this reduces to the familiar criterion for straight lines, $\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\lambda^2} = 0$.

The curvature of a manifold $\mathcal{M}$, with the metric $g_{\mu\nu}$, is encoded in the Riemann tensor. It is defined by

$$[\nabla_\mu, \nabla_\nu]v^\rho = R^\rho{}_{\sigma\mu\nu}v^\sigma, \tag{38}$$

which, in our case, gives the explicit formula

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}. \tag{39}$$

This form of the Riemann tensor allows us to derive several useful identities, such as

$$R_{\rho\sigma\mu\nu} = R_{[\rho\sigma]\mu\nu} = R_{\rho\sigma[\mu\nu]} = R_{\mu\nu\rho\sigma}. \tag{40}$$

In addition, the properties of the commutator imply the Jacobi identity,

$$[\nabla_\mu, [\nabla_\nu, \nabla_\sigma]] + [\nabla_\sigma, [\nabla_\mu, \nabla_\nu]] + [\nabla_\nu, [\nabla_\sigma, \nabla_\mu]] = 0. \tag{41}$$

If we apply this on δ^μ_ν, we get the differential Bianchi identity, compactly written

$$\nabla_{[\mu}R_{\nu\rho]\sigma\eta} = 0. \tag{42}$$

Although the Christoffel symbols are not tensors, the Riemann tensor is, due to its definition using covariant derivatives. We can therefore contract some of its indices to get other tensorial quantities. We define the Ricci tensor and Ricci scalar as

$$R_{\mu\nu} = R^\rho{}_{\mu\rho\nu}, \tag{43}$$

$$R = R^\mu{}_\mu = g^{\mu\nu}R_{\mu\nu}. \tag{44}$$

To interpret the Riemann tensor, we define the parallel propagator P. We want this object to take a vector at one point and parallel transport it to another point. A vector that is transported along a curve parametrized by λ should then obey

$$v^\mu(\lambda) = P^\mu{}_\nu(\lambda)v^\nu. \tag{45}$$

Inserting this into the equation for parallel transport, Equation 36, this operator must obey

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}P^\mu{}_\nu = -\Gamma^\mu_{\rho\sigma}\frac{\mathrm{d}x^\rho}{\mathrm{d}\lambda}P^\sigma{}_\nu. \tag{46}$$

This has the same form as the definition of the unitary time-evolution operator in quantum mechanics, and we could therefore write down a solution involving an exponential and a path ordering operator, $\mathcal{P}$, analogous to the time ordering operator from quantum mechanics. We may rewrite the equation on an integral form,

$$P^\mu{}_\nu(\lambda) = \delta^\mu_\nu - \int_0^\lambda \mathrm{d}\lambda' \Gamma^\mu_{\rho\sigma}V^\rho P^\sigma{}_\nu, \tag{47}$$

where we denote $\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} = V^\mu$. This allows us to solve the equation iteratively. If $\lambda \leq \epsilon \ll 1$, we expect this to converge as long as the g is well-behaved. Starting with the zeroth-order solution $P^\mu{}_\nu = \delta^\mu_\nu$ and iterating twice gives us

$$P^\mu{}_\nu(\lambda) = \delta^\mu_\nu - \int_0^\lambda \mathrm{d}\lambda'\, \Gamma^\mu_{\rho\nu}V^\rho + \int_0^\lambda \mathrm{d}\lambda' \int_0^{\lambda'} \mathrm{d}\lambda''\, \Gamma^\mu_{\rho\sigma}\Gamma^\sigma_{\eta\nu}V^\rho V^\eta + \mathcal{O}(\epsilon^3). \tag{48}$$

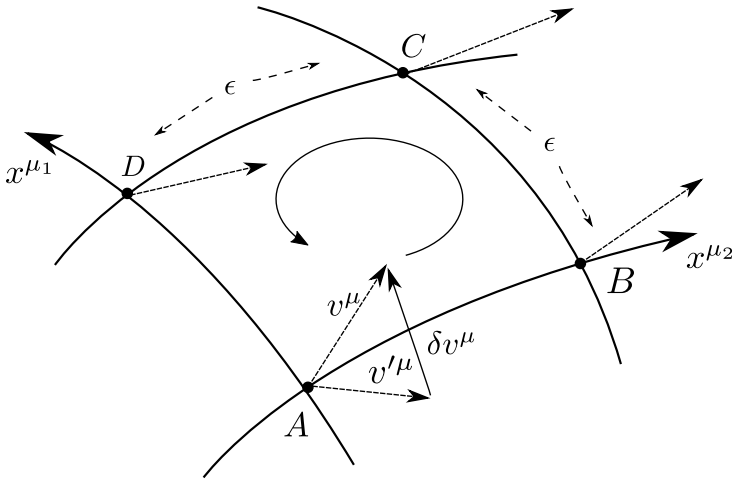


Figure 4: A vector v^μ is parallel transported in a small, closed loop, defined by the coordinate functions x^{μ_1} and x^{μ_2}. As a consequence of the curvature, it has changed by δv^μ by the time it arrives back at A.

With this, we will investigate how much a vector v^μ is changed by being parallel transported around in a small loop, as illustrated in Figure 4.

We transport v^μ along the coordinate lines. These are lines where either of the coordinate functions x^{μ_1} or x^{μ_2} are equal to 0 or ϵ. Here, the indices μ_1 and μ_2 are not free but identify the two coordinate functions which define this loop. They will therefore break summation rules; such indices may appear only on one side of an equation. The line from A to B, defined by $x^{\mu_1} = 0$, is parametrized by $x^\mu(\lambda) = \lambda\delta^\mu_{\mu_2}$, so $V^\mu = \delta^\mu_{\mu_2}$. The Christoffel symbol along this line is

$$\Gamma^\mu_{\nu\rho}(\lambda) = \Gamma^\mu_{\nu\rho}|_A + \lambda\partial_{\mu_2}\Gamma^\mu_{\nu\rho}|_A + \mathcal{O}(\epsilon^2). \tag{49}$$

Inserting this into Equation 48, we get

$$P^\mu{}_\nu(\epsilon) = \delta^\mu_\nu - \epsilon\Gamma^\mu_{\nu\mu_2}|_A + \frac{1}{2}\epsilon^2\left(\Gamma^\mu_{\mu_2\sigma}\Gamma^\sigma_{\mu_2\nu}|_A - \partial_{\mu_2}\Gamma^\mu_{\nu\mu_2}|_A\right) + \mathcal{O}(\epsilon^3). \tag{50}$$

Next, from B to C, the line is $x^\mu(\lambda) = \epsilon\delta^\mu_{\mu_2} + \lambda\delta^\mu_{\mu_1}$, so $V^\mu = \delta^\mu_{\mu_1}$, and the Christoffel symbols are $\Gamma^\mu_{\nu\rho} = \Gamma^\mu_{\nu\rho}|_B + \lambda\partial_{\mu_1}\Gamma^\mu_{\nu\rho}|_B$ to fist order in λ. Here, we have to expand once more to evaluate the symbols at A. Then, we get

$$\Gamma^\mu_{\nu\rho} = \Gamma^\mu_{\nu\rho}|_A + \epsilon\partial_{\mu_2}\Gamma^\mu_{\nu\rho}|_A + \lambda\partial_{\mu_1}\Gamma^\mu_{\nu\rho}|_A + \mathcal{O}(\epsilon^2), \tag{51}$$

The parallel propagator from B to C is then

$$P^\mu{}_\nu(\epsilon) = \delta^\mu_\nu - \epsilon\Gamma^\mu_{\nu\mu_1}|_A + \frac{1}{2}\epsilon^2\left(\Gamma^\mu_{\sigma\mu_1}\Gamma^\sigma_{\nu\mu_1}|_A - \partial_{\mu_1}\Gamma^\mu_{\nu\mu_1}|_A - 2\partial_{\mu_2}\Gamma^\mu_{\nu\mu_1}|_A\right) + \mathcal{O}(\epsilon^3), \tag{52}$$

Which gives the combined propagator from A to C, to and including second order in ϵ, as

$$\begin{aligned}
P_{AC}{}^\mu{}_\nu &= \left[\delta^\mu_\sigma - \epsilon\Gamma^\mu_{\sigma\mu_1} + \frac{1}{2}\epsilon^2\left(\Gamma^\mu_{\eta\mu_1}\Gamma^\eta_{\sigma\mu_1} - \partial_{\mu_2}\Gamma^\mu_{\sigma\mu_2} - 2\partial_{\mu_2}\Gamma^\mu_{\sigma\mu_1}\right)\right] \\
&\quad\times\left[\delta^\sigma_\nu - \epsilon\Gamma^\sigma_{\nu\mu_2} + \frac{1}{2}\epsilon^2\left(\Gamma^\sigma_{\eta\mu_2}\Gamma^\eta_{\nu\mu_2} - \partial_{\mu_2}\Gamma^\sigma_{\nu\mu_2}\right)\right] \\
&= \delta^\mu_\nu - \epsilon\left(\Gamma^\mu_{\nu\mu_1} + \Gamma^\mu_{\nu\mu_2}\right) \\
&\quad + \epsilon^2\frac{1}{2}\left(2\Gamma^\mu_{\sigma\mu_1}\Gamma^\sigma_{\nu\mu_2} + \Gamma^\mu_{\sigma\mu_1}\Gamma^\sigma_{\nu\mu_1} + \Gamma^\mu_{\sigma\mu_2}\Gamma^\sigma_{\nu\mu_2} - 2\partial_{\mu_2}\Gamma^\mu_{\nu\mu_1} - \partial_{\mu_1}\Gamma^\mu_{\nu\mu_1} - \partial_{\mu_2}\Gamma^\mu_{\nu\mu_2}\right).
\end{aligned} \tag{53}$$

The parallel propagator for CDA is the propagator for ADC with its signs flipped. The ADC propagator is the same as ABC, only with the μ_1 and μ_2 indices switched. It is thus

$$\begin{aligned}
P_{CA}{}^\mu{}_\nu &= \delta^\mu_\nu + \epsilon\left(\Gamma^\mu_{\nu\mu_2} + \Gamma^\mu_{\nu\mu_1}\right) \\
&\quad + \epsilon^2\frac{1}{2}\left(2\Gamma^\mu_{\sigma\mu_2}\Gamma^\sigma_{\nu\mu_1} + \Gamma^\mu_{\sigma\mu_2}\Gamma^\sigma_{\nu\mu_2} + \Gamma^\mu_{\sigma\mu_1}\Gamma^\sigma_{\nu\mu_1} + 2\partial_{\mu_1}\Gamma^\mu_{\nu\mu_2} + \partial_{\mu_2}\Gamma^\mu_{\nu\mu_2} + \partial_{\mu_1}\Gamma^\mu_{\nu\mu_1}\right).
\end{aligned} \tag{54}$$

The full propagator, from A to A, is $P^\mu{}_\nu = P_{CA}{}^\mu{}_\rho P_{AC}{}^\rho{}_\nu$. The terms linear in ϵ vanish, and the same with the terms with two equal μ_i-indices. The change in the vector as it is rotated around the loop is

therefore, to second order in ϵ,

$$\delta v^\mu = P^\mu{}_\nu v^\nu - v^\mu = \epsilon^2 \left(\Gamma^\mu_{\sigma\mu_1}\Gamma^\sigma_{\nu\mu_2} - \Gamma^\mu_{\sigma\mu_2}\Gamma^\sigma_{\nu\mu_1} + \partial_{\mu_1}\Gamma^\mu_{\nu\mu_2} - \partial_{\mu_2}\Gamma^\mu_{\nu\mu_1}\right) v^\nu. \tag{55}$$

Comparing with Equation 39, we see that this is the Riemann curvature tensor. In other words, the Riemann tensor encodes how a vector is transformed when parallel transported in a small, closed loop.

1.5 Integration on manifolds

The integral of a scalar function on a manifold is not a coordinate-independent notion. To obtain this, we must introduce n-forms. A n-form ω is an antisymmetric $(0, n)$ tensor. This means that it has coordinates which obey $\omega_{\mu_1\ldots\mu_n} = \omega_{[\mu_1\ldots\mu_n]}$. The n-forms are ubiquitous objects in mathematics and physics, one example is the electromagnetic field-strength tensor $F_{\mu\nu}$, and they allow for the definitions of coordinate independent integration and derivation. We will define two important maps between n-forms. The wedge product, $\wedge$, is a product that maps two n- and m-forms to an $n+m$-form, and is defined as

$$(A \wedge B)_{\mu_1\ldots\mu_{n+m}} = \frac{(n+m)!}{n!m!} A_{[\mu_1\ldots\mu_n} B_{\mu_{n+1}\ldots\mu_{n+m}]}. \tag{56}$$

Furthermore, we define the exterior derivative, a map from n-forms to $n+1$-forms, defined by

$$(\mathrm{d}T)_{\mu_1\ldots\mu_{n+1}} = (n+1)\partial_{[\mu_1} T_{\mu_2\ldots\mu_{n+1}]}. \tag{57}$$

We are interested in a coordinated independent quantity that we can integrate over. To that end, we define

$$\mathrm{d}^n x := \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1} = \varepsilon_{\mu_1\ldots\mu_n} \mathrm{d}x^{\mu_1} \otimes \cdots \otimes \mathrm{d}x^{\mu_n}, \tag{58}$$

where $\varepsilon_{\mu_1\ldots\mu_n}$ is the Levi-Civita symbol. Given a new set of coordinates, x'^μ, we may similarly define a new n-form, $\mathrm{d}^n x'$. These two n-forms are related by

$$\mathrm{d}^n x = \det\left(\frac{\partial x}{\partial x'}\right) \mathrm{d}^n x', \tag{59}$$

where we have used the relation $\varepsilon_{\mu_1\ldots\mu_n}\det(A) = \varepsilon_{\nu_1\ldots\nu_n} A^{\nu_1}{}_{\mu_1}\ldots A^{\nu_n}{}_{\mu_n}$. We define $|g| = |\det(g_{\mu\nu})|$, where $|\cdot|$ denote the absolute value. By the transformation properties of tensors, this transforms as

$$\sqrt{|g'|} = \left|\det\left(\frac{\partial x'}{\partial x}\right)\right| \sqrt{|g'|}, \tag{60}$$

This means we can use $|g|$ to compensate for the transformation of $\mathrm{d}^n x$, and get a volume form with a coordinate independent definition,

$$\mathrm{d}V = \sqrt{|g|}\,\mathrm{d}^n x = \sqrt{|g'|}\,\mathrm{d}^n x'. \tag{61}$$

With this, we can integrate scalars in a well-defined way by mapping them to a corresponding n-form, $f \to f\mathrm{d}V$. We define the integral of a scalar function f on a manifold $\mathcal{M}$ with a metric g as

$$I[f] = \int_{\mathcal{M}} \mathrm{d}V\, f = \int_{\mathcal{M}} \mathrm{d}^n x\, \sqrt{|g(x)|}\, f(x). \tag{62}$$

Stoke's theorem generalizes the fundamental theorem of calculus and the divergence theorem to manifolds. Let $\mathcal{M}$ be a differential manifold of dimension n, with the boundary $\partial\mathcal{M}$. Stoke's theorem says that an $n-1$-form ω and its exterior derivative $\mathrm{d}\omega$ are related by

$$\int_{\mathcal{M}} \mathrm{d}\omega = \int_{\partial\mathcal{M}} \omega. \tag{63}$$

This theorem implies a generalized divergence theorem. The boundary of $\mathcal{M}$ is a $n-1$ manifold dimensional, and a metric g on $\mathcal{M}$ will induce a new metric γ on $\partial\mathcal{M}$. This metric corresponds to the restriction of g to $\partial\mathcal{M}$. Furthermore, there will be a vector field n^μ of normalized vectors orthogonal to all elements of $T\partial\mathcal{M}$. This theorem states that for a vector field V^μ on $\mathcal{M}$,

$$\int_{\mathcal{M}} \mathrm{d}^n x\, \sqrt{|g|}\, \nabla_\mu V^\mu = \int_{\partial\mathcal{M}} \mathrm{d}^{n-1} y\, \sqrt{|\gamma|}\, n_\mu V^\mu. \tag{64}$$

2 Lie theory

This section is based on [**leeIntroductionSmoothManifolds2003d**, **peskinIntroductionQuantumField1995**, **schwartzQuantumFieldTheory2013**, **weinbergQuantumTheoryFields1995**, **weinbergQuantumTheoryFields**

One application of differential geometry is in the theory of Lie groups. The primary use case of these groups is to study the symmetries of manifolds, particularly those that represent physical systems. Symmetries are a vital part of modern physics and will be used throughout this text. In this section, we will develop the tools that we need for this study.

2.1 Lie groups

Inspired by Èvariste Galois' use of finite groups to study the finite solution set of algebraic equations, Sophus Lie introduced Lie groups, *topological groups*, to study the solutions of differential equations. Groups are natural structures to capture symmetries, as they can be defined as actions on an object which leave it unchanged. A group is a set, G, together with a map

$$(\cdot,\cdot) : G \times G \longmapsto G, \tag{65}$$

$$(g_1, g_2) \longmapsto g_3 = g_1 g_2, \tag{66}$$

called group multiplication. This map obeys the group axioms, which are the existence of an identity element $\mathbb{1}$, associativity, and the existence of an inverse element g^{-1} for all $g \in G$. These can be written as

$$\begin{array}{ll} \exists \mathbb{1} \in G, \text{ s.t.}, \forall g \in G, & g\mathbb{1} = g, \\ \forall g_1, g_2, g_3 \in G, & g_1(g_2 g_3) = (g_1 g_2) g_3, \\ \forall g \in G,\ \exists g^{-1} \in G, \text{ s.t.}, & g g^{-1} = \mathbb{1}. \end{array} \tag{67}$$

A Lie group is a manifold G with a group structure. Elements $g_1, g_2 \in G$ can thus be combined by group multiplication and mapped to their inverses. We additionally require these maps to be smooth, which is equivalent to $(g_1, g_2) \to g_1 g_2^{-1}$ being smooth.

As we will discuss in more detail in the next chapter, a symmetry transformation is a map between physical states which leaves the equations governing that system unchanged. Assume the field, or set of fields, φ is governed by the equation $f(\varphi) = 0$. A symmetry transformation $\varphi \mapsto g\varphi$, where $g\varphi$ represents the action g acting on φ, will then obey $f(g\varphi) = 0$. This is what makes groups the natural structures to describe symmetries. Assume G is the set of all symmetries of a system, or a subset of them closed under compositions

$$G = \{\, g \mid f(g\varphi) = 0 \,\}. \tag{68}$$

The group G might act on φ linearly, so $(g\varphi)_i = g_{ij}\varphi_j$, or in a more complicated manner. A linear realization of a Lie group is called a *representation*. In any case, the group multiplication is composition, i.e., performing transformations in succession. This map is closed, as the composite of two symmetry transformations is another symmetry transformation. The identity map is a symmetry transformation, and composition is associative. This means that invertible symmetry transformations form a group, and for continuous sets, this group is a Lie group.

We will focus on connected Lie groups, in which all elements $g \in G$ are in the same connected piece as the identity map $\mathbb{1}\varphi = \varphi$. This means that for each $g \in G$, one can find a continuous path $\gamma(t)$ in the manifold, such that $\gamma(0) = \mathbb{1}$ and $\gamma(1) = g$. Given such a path, we can study transformations close to the identity element. As the Lie group is a smooth manifold, we can write[8]

$$\gamma(\epsilon) = \mathbb{1} + i\epsilon V + \mathcal{O}(\epsilon^2). \tag{69}$$

V is a generator, and is defined as

$$iV = \frac{\mathrm{d}\gamma}{\mathrm{d}t}\Big|_{t=0}. \tag{70}$$

The generator is thus a member of the tangent space of the identity element, $T_{\mathbb{1}}G$. We denote the coordinates of G by $\eta_\alpha \in \mathbb{R}^n$. As before, we can denote a path γ in a manifold G by its path through $\mathbb{R}^n$, $\gamma(t) = g(\eta(t))$. We will assume, without loss of generality, that $\eta_\alpha(0) = 0$ and $g(0) = \mathbb{1}$. We can then write the generator as

$$V = \frac{\mathrm{d}\gamma}{\mathrm{d}t}\Big|_{t=0} = \frac{\mathrm{d}\eta_\alpha}{\mathrm{d}t}\Big|_{t=0} \frac{\partial g}{\partial \eta_\alpha}\Big|_{\eta=0} = v_\alpha T_\alpha, \quad T_\alpha = \frac{\partial g}{\partial \eta_\alpha}\Big|_{\eta=0}, \quad v_\alpha = \frac{\mathrm{d}\eta_\alpha}{\mathrm{d}t}\Big|_{t=0}. \tag{71}$$

[8] The factor i is a physics convention and differs from how mathematicians define generators of a Lie group.

Infinitesimal transformations can therefore be written as

$$\gamma(\epsilon) = \mathbb{1} + i\epsilon v_\alpha T_\alpha + \mathcal{O}(\epsilon^2). \tag{72}$$

The generators form a new algebraic structure, the Lie algebra. This is the linearization of the Lie group, but it encodes information about the whole group. We will, in fact, mostly focus on the generators rather than the full transformations, which makes the Lie algebra an important structure.

2.2 Lie algebras

An abstract Lie algebra $\mathfrak{g}$ is a vector space V, together with a binary operation,

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \longmapsto \mathfrak{g}, \tag{73}$$
$$(V_1, V_2) \longmapsto V_3 = [V_1, V_2]. \tag{74}$$

This map is linear in both arguments, antisymmetric, and obeys the *Jacobi identity*,

$$[V_1,[V_2,V_3]] + [V_2,[V_3,V_1]] + [V_3,[V_1,V_2]] = 0, \quad \forall V_1, V_2, V_3 \in \mathfrak{g}. \tag{75}$$

In our case, we will work with concrete Lie algebras $\mathfrak{g}$, corresponding to Lie groups G. As long as G is simply connected, i.e., connected to the identity and without holes, then this is a one-to-one correspondence. This was Sophus Lie's main result. The space V is then $T_{\mathbb{1}}G$ with the basis T_α, and we can define the Lie bracket by

$$[T_\alpha, T_\beta] = i f^\gamma_{\alpha\beta} T_\gamma, \tag{76}$$

where $f^\gamma_{\alpha\beta}$ are the structure constants of the algebra. These obey $f^\gamma_{\alpha\beta} = -f^\gamma_{\beta\alpha}$ and they follow their own version of the Jacobi identity,

$$f^\mu_{\alpha\beta} f^\nu_{\gamma\mu} + f^\mu_{\beta\gamma} f^\nu_{\alpha\mu} + f^\mu_{\gamma\alpha} f^\nu_{\beta\mu} = 0. \tag{77}$$

An algebra is called *abelian* if $f^\alpha_{\beta\gamma} = 0$. Any abelian algebra is just a direct sum of N of the Lie algebras of $\mathrm{U}(1)$, $\mathfrak{u}(1)$. A simple Lie algebra is a non-abelian algebra that does not contain any non-trivial *ideals*, also called invariant sub-algebras.[9] An ideal $\mathfrak{i} \subset \mathfrak{g}$ is a set such that $[\mathfrak{i}, \mathfrak{g}] \subset \mathfrak{i}$. A semi-simple Lie algebra can be written as a direct sum of simple algebras. The total classification of simple Lie algebras was done by Cartan and Killing and involved four infinite families, the classical algebras such as $\mathfrak{su}(N)$, and five exceptional algebras which do not fit into these families. There is a natural metric on Lie groups, called the Killing form, $B_{\alpha\beta} = -f^\gamma_{\alpha\eta} f^\eta_{\beta\gamma}$. This is non-degenerate if the corresponding Lie algebra is semi-simple. Additionally, if the Lie group is compact, it is positive definite. In that case, one can choose a basis of the algebras so that the structure constants are *totally antisymmetric*. This is trivially true for abelian groups, and as a consequence, the structure constants of a direct sum of compact, simple algebras and abelian algebras are totally antisymmetric, and we write $f^\alpha_{\beta\gamma} = f_{\alpha\beta\gamma}$ [**weinbergQuantumTheoryFields1996**].

As with Lie groups, Lie algebras have representations. A representation of a Lie algebra $\mathfrak{g}$ is a homomorphism, i.e., a map that preserves the Lie bracket, from $\mathfrak{g}$ to the Lie algebra of linear maps on a vector space V, i.e., matrices, called $\mathfrak{gl}(V)$. In $\mathfrak{gl}(V)$, the Lie bracket is the matrix commutator. A representation is faithful if the homomorphism is injective. The most important representations are the fundamental and the adjoint. The fundamental representation is the smallest faithful representation, and in the case of the familiar groups such as $\mathfrak{so}(N)$ and $\mathfrak{su}(N)$, the fundamental representation is the defining one. In the adjoint representation, the generators T^A_α are the structure constants, $(T^A_\alpha)^\beta_\gamma = -i f^\beta_{\alpha\gamma}$. For compact Lie algebras, i.e., the algebras of compact Lie groups, the representations are Hermitian.

A subset of the original Lie group, $H \subset G$, closed under the group action, is called a subgroup. H then has its own Lie algebra $\mathfrak{h}$, with a set of $m = \dim H$ generators, t_a, which is a subset of the original generators T_α. We denote the remaining set of generators x_i, such that t_a and x_i together span $\mathfrak{g}$. Assume G is compact and its Lie algebra semi-simple. The commutators of t_a must be closed, which means that we can write

$$[t_a, t_b] = i f_{abc} t_c, \tag{78}$$
$$[t_a, x_i] = i f_{aik} x_k, \tag{79}$$
$$[x_i, x_j] = i f_{ijk} x_k + i f_{ijc} t_c, \tag{80}$$

[9] Some authors do not have the non-abelian criterion and includes $\mathfrak{u}(1)$ as simple.

where abc runs over the generators of $\mathfrak{h}$, and ijk runs over the rest. The second formula can be derived using the total anti-symmetry of the structure constants, which implies that $f_{abk} = 0 = -f_{akb}$. This is called a Cartan decomposition. One parameter subgroups are one special case of Lie subgroups. If a curve $\gamma(t)$ through G obey

$$\gamma(t)\gamma(s) = \gamma(t+s), \quad \gamma(0) = \mathbb{1}, \tag{81}$$

then all the points on this curve from a one parameter subgroup of G. This path is associated with a generator,

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t}\Big|_{t=0} = i\eta_\alpha T_\alpha. \tag{82}$$

This association is one-to-one, and allows us to define the exponential map,

$$\exp\{i\eta_\alpha T_\alpha\} := \gamma(1). \tag{83}$$

For connected and compact Lie groups, all elements of the Lie group $g \in G$ can be written as an exponential of elements in the corresponding Lie algebra $\eta_\alpha T_\alpha \in \mathfrak{g}$. For matrix groups, the exponential equals the familiar series expansion

$$\exp\{X\} = \sum_n \frac{1}{n!} X^n. \tag{84}$$

3 Classical field theory

integrate onl and new notes

3.1 Old notes

A field in the sense of classical field theory is a function which takes in a point in space time, $x \in \mathcal{M}$, an returns a value in some space X, for example scalars ($X = \mathbb{R}$), spinors or vectors. Let the set of all fields $\psi_\alpha(x) \in X$ in consideration be $C(X)$. The dynamics of these fields are given by the Lagrangian density $\mathcal{L}$. The Lagrangian density might be tough of in two ways. Firstly as a functional which takes in a field and returns a scalar field

$$\begin{aligned} \mathcal{L} : C(X) &\longrightarrow C(\mathbb{R}) \\ \psi_\alpha(x) &\longrightarrow \mathcal{L}[\psi_\alpha](x) \end{aligned}$$

The action is then defined as the integral of the Lagrangian over some part of space time $\Omega \subseteq \mathcal{M}$:

$$S[\psi_\alpha] = \int_\Omega \mathrm{d}x\, \mathcal{L}[\psi_\alpha](x).$$

The second way is as a scalar function of the value of the field, its derivatives and possibly space time itself: $\mathcal{L}(\varphi_\alpha, \partial_\mu \varphi_\alpha, ..., x)$. This second form is how the Lagrangian of a specific theory is defined, for example

$$\mathcal{L}_{KG} = \frac{1}{2}(\partial\varphi)^2 - \frac{1}{2}m\varphi^2, \quad \mathcal{L}_D = \bar{\psi}(i\partial\!\!\!/ - m)\psi.$$

If we have a smooth one-parameter family of transformation of the field

$$\psi_\alpha(x) \longrightarrow \psi'_\alpha(x;t), \quad t \in \mathbb{R},\ \psi'_\alpha(x;0) = \psi_\alpha(x),$$

This becomes, for $\epsilon \ll 1$

$$\psi_\alpha(x) \longrightarrow \psi'_\alpha(x;\varepsilon) \sim \psi_\alpha(x) + \varepsilon f_\alpha[\psi](x) = \psi_\alpha(x) + \delta\psi_\alpha(x).$$

The dynamical equations of the field, the Euler-Lagrange equations, are obtained using an arbitrary transformation of the fields

$$\psi_\alpha(x) \longrightarrow \psi_\alpha(x) + \varepsilon\eta(x), \quad \eta(\partial\Omega) = 0,$$

and demand that the action is stationary:

$$\delta S[\psi] = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\left(S[\psi_\alpha] - S[\psi'_\alpha(\varepsilon)]\right) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\left(\int_\Omega \mathrm{d}x\, \left(\mathcal{L}[\psi_\alpha](x) - \mathcal{L}[\psi'_\alpha(\varepsilon)](x)\right)\right.$$

We then use the fact that $\mathcal{L}$ may be viewed as a smooth function of the field and its derivatives, $\mathcal{L}(\psi'_\alpha, \partial_\mu \psi'_\alpha) \sim \mathcal{L}(\psi_\alpha + \epsilon\eta_\alpha, \partial_\mu(\psi_\alpha + \varepsilon\eta_\alpha))$ to Taylor expand it. Furthermore, the fact that the variation η vanish at the boundaries may be used to do partial integration, meaning

$$\delta S[\psi] = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_\Omega \mathrm{d}x \left(\frac{\partial \mathcal{L}}{\partial \psi_\alpha} \epsilon\eta_\alpha(x) + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_\alpha)} \epsilon \partial_\mu \eta_\alpha(x) \right) = \int_\Omega \mathrm{d}x \left(\frac{\partial \mathcal{L}}{\partial \psi_\alpha} - \mathrm{d}_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_\alpha)} \right) \right) \eta_\alpha(x) = 0$$

here, d_μ is used to remind us that we must consider the derivative of the Lagrangian as a scalar field $\frac{\partial *}{\partial \mathcal{L}}(\partial_\mu \psi_\alpha)[\psi](x)$ in the same way as the Lagrangian, and thus take a total derivative w.r.t. x^μ. As η_α is arbitrary, this implies

$$\frac{\partial \mathcal{L}}{\partial \psi_\alpha} - \mathrm{d}_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_\alpha)} \right) = 0.$$

Nöthers theorem

If a transformation leaves the Lagrangian unchanged modulo a divergence, i.e. $\mathcal{L}[\psi'](x) = \mathcal{L}[\psi](x) + \mathrm{d}_\mu K^\mu[\psi](x)$, then the one may use the divergence theorem to show that

$$\Delta S = \int_\Omega \mathrm{d}x\, \partial_\mu K^\mu = \int_{\partial\Omega} \mathrm{d}x\, n_\mu K^\mu,$$

where n^μ is the surface normal of Ω. [10] If we repeat the procedure to obtain the EL-equations, this part of the action will not contribute, as the variation on the boundary is assumed to be zero, and the equations of motion is unchanged. This is therefore called a (continuous) symmetry of the Lagrangian. Assuming the transformation is of the form as before, $\psi'_\alpha \sim \psi_\alpha + \varepsilon f_\alpha[\psi]$, we may use the second form of the Lagrangian to write

$$\delta\mathcal{L} = \varepsilon \frac{\partial \mathcal{L}}{\partial \psi_\alpha} f_\alpha[\psi] + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_\alpha)} \varepsilon \mathrm{d}_\mu f_\alpha[\psi] = \varepsilon \mathrm{d}_\mu K^\mu[\psi]\varepsilon.$$

Using the EL-equations, this gives us a conserved current

$$J^\mu[\psi] = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_\alpha)} f_\alpha[\psi] - K^\mu[\psi], \quad \mathrm{d}_\mu J^\mu[\psi] = 0.$$

Using translation, $\psi'_\alpha(x) = \psi(x - \varepsilon\xi) \sim \psi_\alpha - \varepsilon\xi^\mu \partial_\mu \psi_\alpha(x)$, we get $f_\alpha[\psi] = \xi^\mu \partial_\mu \psi_\alpha(x)$. As the Lagrangian is a scalar, we also get $\mathcal{L}[\psi'](x) = \mathcal{L}[\psi](x - \varepsilon\xi) = \mathcal{L}[\psi] - \varepsilon\xi^\mu \partial_\mu \mathcal{L}[\psi](x)$. As the translation vector is arbitrary, we get the canonical stress-energy tensor

$$\Theta^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_a)} \partial_\nu \psi_\alpha - \delta^\mu_\nu \mathcal{L}.$$

Functional derivative

Functional differentiation generalizes the property

$$\mathrm{d}F(x) = \sum_i \frac{\partial F}{\partial x_i} \mathrm{d}x_i$$

We have $S[\phi]$, so we get, and define the variation of S by an arbitrary $\eta(x)$, as in the derivation of the EL-equations:

$$\delta S[\psi_\alpha] = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \left(S[\psi_\alpha] - S[\psi'_\alpha(\varepsilon)] \right) =: \int_\Omega \mathrm{d}x\, \frac{\delta S}{\delta \psi_\beta(x)} \delta\psi_\beta(x)$$

The EL-equations are thus given by $\frac{\delta S[\psi]}{\delta \psi_\alpha(x)} = 0$. We may, as in classical mechanics, introduce a canonical momentum and Hamiltonian density. Let $\partial_0 f = \dot{f}$, and $\bar{\partial}_\mu = (0, \partial_1, ...)$. Then

$$\pi_\alpha = \frac{\partial \mathcal{L}}{\partial(\dot{\psi}_\alpha)}, \quad \mathcal{H}[\psi, \pi] = \mathcal{H}(\psi, \hat{\partial}\psi, \pi) = \pi_\alpha \dot{\psi}_\alpha - \mathcal{L}[\psi].$$

[10] Uppercase delta is used here to distiguish it from the variation we used to obtain the EL equations.

The equations of motion are given by Hamiltons equation,

$$\dot{\psi_\alpha} = \{\pi_\alpha, \mathcal{H}\}, \quad \dot{\pi}_\alpha = \{\psi_\alpha, \mathcal{H}\}.$$

We have here introduced the Poisson brackets,

$$\{A, B\} = \int \mathrm{d}y \int \mathrm{d}x \left(\frac{\delta A}{\delta \psi_\alpha(x)} \frac{\delta B}{\delta \pi_\alpha(y)} - \frac{\delta B}{\delta \psi_\alpha(y)} \frac{\delta A}{\delta \pi_\alpha(x)} \right).$$

Field theory on curved manifolds

A action on curved manifold may defined as with the metric $g_{\mu\nu}$, for a scalar $\mathcal{L}$

$$S[\psi_\alpha] = \int_\Omega \mathrm{d}^n x \, \sqrt{g} \mathcal{L}[\psi_\alpha]$$

When the metric is the free field, i.e. $\psi_\alpha = g_{\mu\nu}$, one needs to vary with respect to $\sqrt{g} = \sqrt{\det(g)}$. We then get, for arbitrary $\eta(x)_{\mu\nu}$

$$\begin{aligned}
\delta\sqrt{g} &= \lim_{\varepsilon\to 0} \sqrt{\det(g^{\mu\nu} + \varepsilon\eta^{\mu\nu})} - \sqrt{g} = \frac{1}{2}\sqrt{g}\delta \ln(g) \\
\delta \ln(g) &= \lim_{\varepsilon\to 0} \ln(\det(g_{\mu\rho}[g^\rho{}_\nu - \varepsilon g^{\rho\sigma}\eta_{\sigma\nu}])) - \ln(g) = \lim_{\varepsilon\to 0} \ln(\det(g^\rho{}_\nu - \varepsilon g^{\rho\sigma}\eta_{\sigma\nu})) \\
\det(g + \varepsilon M) &= \frac{1}{n!}\varepsilon^{\mu_1\cdots}\varepsilon_{\nu_1\ldots}(g + \varepsilon M)^{\nu_1}{}_{\mu_1}\ldots = \frac{1}{n!}\varepsilon^{\mu_1\cdots}\varepsilon_{\nu_1\ldots}(g^{\nu_1}{}_{\mu_1}\ldots + \varepsilon M^{\nu_1}{}_{\mu_1} g^{\nu_2}{}_{\mu_2}\ldots + \ldots) \\
&= 1 + \frac{1}{n!}\varepsilon\epsilon^{\mu\mu_2\cdots}\epsilon_{\nu\mu_2\ldots} M^\nu{}_\mu + \mathcal{O}(\varepsilon^2) = 1 + \varepsilon M^\mu{}_\mu + \mathcal{O}(\varepsilon^2)
\end{aligned}$$

Using $\ln(1 + \varepsilon x) \sim x$, and $\delta(g^{\mu\nu} g_{\mu\nu}) = 0$,

$$\delta\sqrt{g} = -\frac{1}{2}\sqrt{g} g_{\mu\nu}\delta g^{\mu\nu}.$$

Furthermore, as $\delta\Gamma = \Gamma - \Gamma'$, it is a tensor, and we may write

$$\begin{aligned}
\delta R^\rho{}_{\sigma\mu\nu} &= \delta(\partial_{[\mu}\Gamma^\rho_{\nu]\sigma} + \Gamma^\rho_{\lambda[\mu}\Gamma^\lambda_{\nu]\sigma}) = \partial_{[\mu}\delta\Gamma^\rho_{\nu]\sigma} + (\delta\Gamma^\rho_{\lambda[\mu})\Gamma^\lambda_{\nu]\sigma} + \Gamma^\rho_{\lambda[\mu}(\delta\Gamma^\lambda_{\nu]\sigma}) \\
&= \partial_\mu\delta\Gamma^\rho_{\nu\sigma} + \Gamma^\rho_{\lambda\mu}(\delta\Gamma^\lambda_{\nu\sigma}) - \Gamma^\lambda_{\mu\sigma}(\delta\Gamma^\rho_{\lambda\nu}) - \left(\partial_\nu\delta\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\lambda\nu}(\delta\Gamma^\lambda_{\mu\sigma}) - \Gamma^\lambda_{\nu\sigma}(\delta\Gamma^\rho_{\lambda\mu})\right) + (\Gamma^\lambda_{\mu\nu}\delta\Gamma^\rho_{\lambda\sigma} - \Gamma^\lambda_{\mu\nu}\delta\Gamma^\rho_{\lambda\sigma}) \\
&= \nabla_\mu\delta\Gamma^\rho_{\nu\sigma} - \nabla_\nu\delta\Gamma^\rho_{\mu\sigma} = \nabla_\eta\left(g^\eta{}_\mu\delta\Gamma^\rho_{\nu\sigma} - g^\eta{}_\nu\delta\Gamma^\rho_{\mu\sigma}\right) = \nabla_\eta(K^\rho{}_{\sigma\mu\nu})^\eta,
\end{aligned}$$

where K is a tensorial quantity, which vanish at the boundary of our spacetime.

3.2 Functionals — from master

The principle of stationary action and the path integral method relies on functional calculus, where ordinary, n-dimensional calculus is generalized to an infinite-dimensional calculus on a space of functions. A functional, S, takes in a function $\varphi(x)$, and returns a real number $S[\varphi]$. We will be often be dealing with functionals of the form

$$S[\varphi] = \int_{\mathcal{M}} \mathrm{d}^n x \, \mathcal{L}[\varphi](x), \tag{85}$$

Here, $\mathcal{L}[\varphi](x)$, the Lagrangian density, is a functional which takes in a function φ, and returns a real number $\mathcal{L}[\varphi](x)$ *for each point* $x \in \mathcal{M}$. $\mathcal{L}$ thus returns a real-valued function, not just a number. $\mathcal{M}$ is the manifold, in our case space-time, of which both $\varphi(x)$ and $\mathcal{L}[\varphi](x)$ are functions. The function φ can, in general, take on the value of a scalar, complex number, spinor, vector, etc…, while $\mathcal{L}[\varphi](x)$ must be a scalar-valued function. This strongly constraints the form of any Lagrangian and is an essential tool in constructing quantum field theories. Although this section is written with a single scalar-valued function φ, this can easily be generalized by adding an index, $\varphi \to \varphi_\alpha$, enumerating all the degrees of freedom, then summing over this index when restating the arguments [**carrollSpacetimeGeometryIntroduction2019**, **schwartzQuantumFieldTheory2013**, **peskinIntroductionQuantumField1995**].

3.3 Functional derivative

The functional derivative is base on an arbitrary *variation* η of the function φ. The variation η, often written $\delta\varphi$, is an arbitrary function only constrained to vanish *quickly enough* at the boundary $\partial\mathcal{M}$.[11] The variation of the functional S is defined as

$$\delta_\eta S[\varphi] = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left(S[\varphi+\epsilon\eta] - S[\varphi]\right) = \frac{\mathrm{d}}{\mathrm{d}\epsilon}S[\varphi+\epsilon\eta]|_{\epsilon=0}. \tag{86}$$

We can regard the variation of a functional as the generalization of the differential of a function, Equation 18, as the best linear approximation around a point. In regular differential geometry, a function f can be approximated around a point x by

$$f(x+\epsilon v) = f(x) + \epsilon \mathrm{d}f(v), \tag{87}$$

where v is a vector in the tangent space at x. In functional calculus, the functional S is analogous to f, φ to x, and η to v. We can more clearly see the resemblance by writing

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}f(x+\epsilon v) = \mathrm{d}f(v) = \frac{\partial f}{\partial x^\mu}v^\mu. \tag{88}$$

In the last line we expanded the differential using the basis-representation, $v = v^\mu\partial_\mu$. To generalize this to functional, we define the *functional derivative*, by

$$\delta_\eta S[\varphi] = \int_{\mathcal{M}} \mathrm{d}^n x\, \frac{\delta S[\varphi]}{\delta\eta(x)}\eta(x). \tag{89}$$

If we let $S[\varphi] = \varphi(x)$, for some fixed x, the variation becomes

$$\delta_\eta S[\varphi] = \eta(x) = \int \mathrm{d}^n y\, \delta(x-y)\eta(y), \tag{90}$$

which leads to the identity

$$\frac{\delta\varphi(x)}{\delta\varphi(y)} = \delta(x-y). \tag{91}$$

There is also a generalized chain rule for functional derivatives. If ψ is some new functional variable, then

$$\frac{\delta S[\varphi]}{\delta\varphi(x)} = \int_{\mathcal{M}} \mathrm{d}^n y\, \frac{\delta S[\varphi]}{\delta\psi(y)}\frac{\delta\psi(y)}{\delta\varphi(x)}. \tag{92}$$

Higher functional derivatives are defined in terms of higher-order variations,

$$\delta_\eta^m S[\varphi] = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\delta_\eta^{m-1}S[\varphi+\epsilon\eta]|_{\epsilon=0} = \int_{\mathcal{M}}\left(\prod_{i=1}^m \mathrm{d}^n x_i\,\eta(x_i)\right)\frac{\delta^m S[\varphi]}{\delta\varphi(x_1)...\delta\varphi(x_m)}. \tag{93}$$

With this, we can write the functional Taylor expansion,

$$S[\varphi_0+\varphi] = S[\varphi_0] + \int_{\mathcal{M}}\mathrm{d}^n x\,\varphi(x)\frac{\delta S[\varphi_0]}{\delta\varphi(x)} + \frac{1}{2}\int_{\mathcal{M}}\mathrm{d}^n x\mathrm{d}^n y\,\varphi(x)\varphi(y)\frac{\delta^2 S[\varphi_0]}{\delta\varphi(x)\,\delta\varphi(y)} + \ldots. \tag{94}$$

Here, the notation $\frac{\delta S[\varphi_0]}{\delta\varphi}$ indicate that $S[\varphi]$ is first differentiated with respect to φ, then evaluated at $\varphi = \varphi_0$ [**peskinIntroductionQuantumField1995**, **schwartzQuantumFieldTheory2013**].

3.4 Gaussian integrals and the stationary phase approximation

An integral that we will use a lot is the Gaussian integral,

$$\int_{\mathbb{R}}\mathrm{d}x\,\exp -\frac{1}{2}ax^2 = \sqrt{\frac{2\pi}{a}}, \tag{95}$$

for $a\in\mathbb{R}$. The imaginary version,

$$\int_R \mathrm{d}x\,\exp i\frac{1}{2}ax^2 \tag{96}$$

[11] The condition of "quickly enough" is to ensure that we can integrate by parts and ignore the boundary condition, which we will do without remorse.

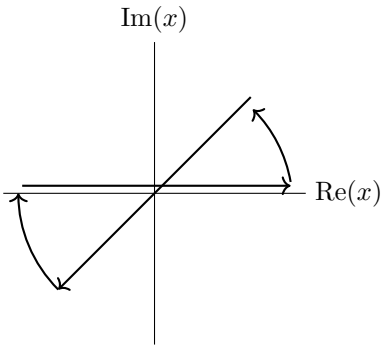


Figure 5: Wick rotation

does not converge. However, if we change $a \to a + i\epsilon$, the integrand is exponentially suppressed.

$$f(x) = \exp i\frac{1}{2}ax^2 \to \exp i\frac{1}{2}ax^2 - \frac{1}{2}\epsilon x^2. \tag{97}$$

As the integrand falls exponentially for $x \to \infty$ and contains no poles in the upper right nor lower left quarter of the complex plane, we may perform a wick rotation by closing the contour as shown in Figure 5. This gives the result

$$\begin{aligned}\int_{\mathbb{R}} \mathrm{d}x\, \exp i\frac{1}{2}(a+i\epsilon)x^2 &= \int_{\sqrt{i}\mathbb{R}} \mathrm{d}x\, \exp i\frac{1}{2}ax^2 \\ &= \sqrt{i}\int_{\mathbb{R}} \mathrm{d}y\, \exp -\frac{1}{2}(-a)y^2 \\ &= \sqrt{\frac{2\pi i}{(-a)}},\end{aligned} \tag{98}$$

where we have made the change of variable $y = (1+i)/\sqrt{2}x = \sqrt{i}x$. In n dimensions, the Gaussian integral formula generalizes to

$$\int_{\mathbb{R}^n} \mathrm{d}^n x\, \exp -\frac{1}{2}x_n A_{nm} x_m = \sqrt{\frac{(2\pi)^n}{\det(A)}}, \tag{99}$$

where A is a matrix with n real, positive eigenvalues. We may also generalize Equation 98,

$$\int_{\mathbb{R}^n} \mathrm{d}^n x\, \exp i\frac{1}{2}x_n(A_{nm} + i\epsilon\delta_{nm})x_m = \sqrt{\frac{(2\pi i)^n}{\det(-A)}}. \tag{100}$$

The final generalization is to functional integrals. If $A(x,x')$ is a linear operator, then

$$\begin{aligned}\int \mathcal{D}\varphi\, \exp -\frac{1}{2}\int \mathrm{d}x\mathrm{d}x'\, \varphi(x)A(x,x')\varphi(x') &= C\det(A)^{-1/2}, \\ \int \mathcal{D}\varphi\, \exp i\frac{1}{2}\int \mathrm{d}x\mathrm{d}x'\, \varphi(x)A(x,x')\varphi(x') &= C'\det(-A)^{-1/2}.\end{aligned} \tag{101}$$

C and C' are divergent constants but will either fall away as we are only looking at the logarithm of I_∞ and can throw away additive constants, or ratios between quantities that are both multiplied by C or $C*$.

The Gaussian integral can be used for the stationary phase approximation. By Taylor expanding around a stationary point x_0, we may approximate

$$\int \mathrm{d}x\, \exp i\alpha f(x) \sim \sqrt{\frac{2\pi}{\alpha f''(x_0)}} \exp \alpha f(x_0), \quad \alpha \to \infty, \tag{102}$$

where the point x_0 is defined by $f'(x_0) = 0$. The functional generalization of this is

$$\int \mathcal{D}\varphi\, \exp i\alpha S[\varphi] \sim C\det\left(-\frac{1}{\alpha}\frac{\delta^2 S[\varphi_0]}{\delta\varphi^2}\right)^{-1/2} \exp i\alpha S[\varphi_0], \tag{103}$$

Here, $S[\varphi]$ is a general functional of φ, we have used the Taylor expansion, and φ_0 obeys

$$\frac{\delta S[\varphi_0]}{\delta\varphi(x)} = 0. \tag{104}$$

3.5 The Euler-Lagrange equation

The Lagrangian may also be written as a scalar function of the field-values at x, $\varphi(x)$, as well as its derivatives, $\partial_\mu\varphi(x)$, for example

$$\mathcal{L}(\varphi, \partial_\mu\varphi) = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 - \frac{1}{4!}\lambda\varphi^4 + \dots. \tag{105}$$

We have omitted the evaluation at x for the brevity of notation. We use this to evaluate the variation of a functional in the of Equation 85,

$$\delta_\eta S[\varphi] = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\int_{\mathcal{M}} \mathrm{d}^n x\, \mathcal{L}[\varphi + \epsilon\eta](x), \tag{106}$$

by Taylor expanding the Lagrangian density as a function of φ and its derivatives,

$$\mathcal{L}[\varphi + \epsilon\eta] = \mathcal{L}\left(\varphi + \epsilon\eta, \partial_\mu\{\varphi + \epsilon\eta\}\right) = \mathcal{L}[\varphi] + \epsilon\left(\frac{\partial\mathcal{L}}{\partial\varphi}\eta + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\partial_\mu\eta\right) + \mathcal{O}(\epsilon^2). \tag{107}$$

Inserting this into Equation 106 and partially integrating the last term allows us to write the variation on the form Equation 89. The functional derivative is

$$\frac{\delta S}{\delta\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}. \tag{108}$$

The principle of stationary action says that the equation of motion of a field obeys $\delta_\eta S = 0$. As η is arbitrary, this is equivalent to setting the functional derivative of S equal to zero. The result is the Euler-Lagrange equations of motion [**schwartzQuantumFieldTheory2013**],

$$\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} = 0. \tag{109}$$

3.6 Functional calculus on a curved manifold

As discussed in subsection 1.5, when integrating a scalar on a curved manifold, we must include the $\sqrt{|g|}$-factor to get a coordinate-independent result. The action in curved spacetime is therefore

$$S[g, \varphi] = \int_{\mathcal{M}} \mathrm{d}^n x\, \sqrt{|g|}\mathcal{L}[g, \varphi], \tag{110}$$

where the action and the Lagrangian now is a functional of both the matter-field φ and the metric $g_{\mu\nu}$. Our example Lagrangian from last section now takes the form

$$\mathcal{L}(g_{\mu\nu}, \varphi, \nabla_\mu\varphi) = \frac{1}{2}g^{\mu\nu}\nabla_\mu\varphi\nabla_\nu\varphi - \frac{1}{2}m^2\varphi^2 - \frac{1}{4!}\lambda\varphi^4\dots, \tag{111}$$

where partial derivatives are substituted with covariant derivatives following the minimal coupling rule. We define the functional derivative as

$$\delta_\eta S = \int_{\mathcal{M}} \mathrm{d}^n x\sqrt{|g|}\frac{\delta S}{\delta\eta(x)}\eta(x). \tag{112}$$

If this is a variation in φ only, this gives the same result as before. However, in general relativity, the metric itself is a dynamic field, and we may therefore vary it. Consider $g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}$. The variation of the action is then assuming $\mathcal{L}$ only depends on g and not its derivatives, we get

$$\delta_g S = \int_{\mathcal{M}} \mathrm{d}^n x\, \left[\left(\delta\sqrt{|g|}\right)\mathcal{L}[g] + \sqrt{|g|}\delta\mathcal{L}[g]\right] \tag{113}$$

The variation of the $\sqrt{|g|}$-factor can be evaluated using the Levi-Civita symbol $\varepsilon_{\mu_1\dots\mu_n}$. The determinant of a $n \times n$-matrix may be written as

$$\det(A) = \frac{1}{n!}\varepsilon_{\mu_1\dots\mu_n}\varepsilon^{\nu_1\dots\nu_n}A^{\mu_1}{}_{\nu_1}\dots A^{\mu_n}{}_{\nu_n}. \tag{114}$$

For a matrix M, then, we can write

$$\begin{aligned}\det(\mathbb{1}+\varepsilon M) &= \frac{1}{n!}\varepsilon_{\mu_1\dots}\varepsilon^{\nu_1\cdots}(\mathbb{1}+\epsilon M)^{\mu_1}{}_{\nu_1}(\mathbb{1}+\epsilon M)^{\mu_2}{}_{\nu_2}\dots \\ &= \frac{1}{n!}\varepsilon_{\mu_1\dots}\varepsilon^{\nu_1\cdots}[\delta^{\mu_1}_{\nu_1}\delta^{\mu_2}_{\nu_2}\cdots+\epsilon(M^{\mu_1}{}_{\nu_1}\delta^{\mu_2}_{\nu_2}\cdots+M^{\mu_2}{}_{\nu_2}\delta^{\mu_1}_{\nu_1}\cdots+\dots)+\mathcal{O}(\epsilon^2)] \\ &= 1+M^{\mu}{}_{\mu}+\mathcal{O}(\epsilon^2). \end{aligned} \tag{115}$$

Using

$$\det(g^{\mu}{}_{\nu}+\epsilon\eta^{\mu}{}_{\nu}) = \det(g^{\mu}{}_{\rho}[\delta^{\rho}_{\nu}+\epsilon g^{\rho\lambda}\eta_{\lambda\nu}]) = \det(g^{\mu}{}_{\rho})(1+\epsilon g^{\nu\lambda}\eta_{\lambda\nu}), \tag{116}$$

we get

$$\delta\sqrt{|g|} = \frac{1}{2}\frac{1}{\sqrt{|g|}}\frac{g}{|g|}\delta\det(g^{\mu}{}_{\nu}) = -\frac{1}{2}\sqrt{|g|}g^{\mu\nu}\delta g_{\mu\nu} \tag{117}$$

The minus sign is due to the determinant of a Lorentzian metric being negative. Assuming the Lagrangian only depends on the metric directly, and not its derivatives, the variation of the action is

$$\delta_g S = \int_{\mathcal{M}} \mathrm{d}^n x\,\sqrt{|g|}\left(\frac{\partial\mathcal{L}}{\partial g^{\mu\nu}}-\frac{1}{2}g_{\mu\nu}\mathcal{L}\right)\delta g^{\mu\nu}. \tag{118}$$

With the Lagrangian in Equation 111, we get

$$\frac{\delta S}{\delta g^{\mu\nu}} = \frac{\partial\mathcal{L}}{\partial g^{\mu\nu}}-\frac{1}{2}g_{\mu\nu}\mathcal{L} = -\frac{1}{2}\left(\frac{1}{2}\nabla_\mu\varphi\nabla_\nu\varphi+\frac{1}{2}m^2\varphi^2 g_{\mu\nu}+\dots\right). \tag{119}$$

We recognize the $(\mu,\nu)=(0,0)$-component as negative half the Hamiltonian density, which supports the definition of the definition of the stress-energy tensor Equation 144 [**carrollSpacetimeGeometryIntroduction2019**].

3.7 Functional derivative of the Einstein-Hilbert action

In the Einstein-Hilbert action, Equation 141, the Lagrangian density is $\mathcal{L}=kR=kg^{\mu\nu}R_{\mu\nu}$, where k is a constant and $R_{\mu\nu}$ the Ricci tensor, Equation 43. As the Ricci tensor is dependent on both the derivative and second derivative of the metric, the variation is

$$\delta S_{\mathrm{EH}} = k\int_{\mathcal{M}}\mathrm{d}^n x\,\sqrt{|g|}\left(\delta R-\frac{1}{2}g_{\mu\nu}\delta g^{\mu\nu}R\right). \tag{120}$$

The variation of the Ricci scalar is

$$\delta R = R_{\mu\nu}\delta g^{\mu\nu}+g^{\mu\nu}\delta R_{\mu\nu}, \tag{121}$$

We can write the variation of the Ricci scalar, and thus the Riemann curvature tensor, in terms of variations in Christoffel symbols, $\delta\Gamma^{\rho}_{\mu\nu}$ using the explicit formula for a symmetric, metric-compatible covariant derivative, Equation 39. As $\delta\Gamma=\Gamma-\Gamma'$, it is a tensor, and we may write

$$\begin{aligned}\delta R^{\rho}{}_{\sigma\mu\nu} &= \delta(\partial_\mu\Gamma^{\rho}_{\nu\sigma}-\partial_\nu\Gamma^{\rho}_{\mu\sigma}+\Gamma^{\rho}_{\lambda\mu}\Gamma^{\lambda}_{\nu\sigma}-\Gamma^{\rho}_{\lambda\nu}\Gamma^{\lambda}_{\mu\sigma}) \\ &= \partial_\mu\delta\Gamma^{\rho}_{\nu\sigma}+\Gamma^{\rho}_{\lambda\mu}\delta\Gamma^{\lambda}_{\nu\sigma}-\Gamma^{\lambda}_{\mu\sigma}\delta\Gamma^{\rho}_{\lambda\nu}-\left(\partial_\nu\delta\Gamma^{\rho}_{\mu\sigma}+\Gamma^{\rho}_{\lambda\nu}\delta\Gamma^{\lambda}_{\mu\sigma}-\Gamma^{\lambda}_{\nu\sigma}\delta\Gamma^{\rho}_{\lambda\mu}\right) \\ &\quad+(\Gamma^{\lambda}_{\mu\nu}\delta\Gamma^{\rho}_{\lambda\sigma}-\Gamma^{\lambda}_{\mu\nu}\delta\Gamma^{\rho}_{\lambda\sigma}) \\ &= \nabla_\mu\delta\Gamma^{\rho}_{\nu\sigma}-\nabla_\nu\delta\Gamma^{\rho}_{\mu\sigma} = \nabla_\eta\left(g^{\eta}{}_{\mu}\delta\Gamma^{\rho}_{\nu\sigma}-g^{\eta}{}_{\nu}\delta\Gamma^{\rho}_{\mu\sigma}\right) := \nabla_\eta K^{\rho\eta}{}_{\mu\rho\nu}, \end{aligned}$$

where K is a tensorial quantity, which vanishes at the boundary of our spacetime. Using the generalized divergence theorem, Equation 64, we see the contribution to the action from this quantity vanish. The contribution comes from an integral over $g^{\mu\nu}\delta R_{\mu\nu}=g^{\mu\nu}\delta R^{\rho}{}_{\mu\rho\nu}=g^{\mu\nu}\nabla_\eta K^{\rho\eta}{}_{\mu\rho\nu}$ Using metric compatibility, we can exchange the covariant derivative and the metric, and we have $g^{\mu\nu}\delta R_{\mu\nu}=\nabla_\eta[g^{\mu\nu}K^{\rho\eta}{}_{\mu\rho\nu}]$. The contribution to the action therefore becomes

$$\int_{\mathcal{M}}\mathrm{d}^4x\,\sqrt{|g|}g^{\mu\nu}\delta R_{\mu\nu} = \int_{\mathcal{M}}\mathrm{d}^4x\,\sqrt{|g|}\nabla_\eta K^{\eta\rho\mu}{}_{\rho\mu} = \int_{\partial\mathcal{M}}\mathrm{d}^3y\,\sqrt{|\gamma|}n_\eta K^{\eta\rho\mu}{}_{\rho\mu}=0, \tag{122}$$

where we used the fact that $\delta g_{\mu\nu}$, and thus K, vanish at $\partial\mathcal{M}$, and the generalized form of the divergence theorem, Equation 64. The variation of the action is therefore

$$\delta S_{\mathrm{EH}} = k\int_{\mathcal{M}}\mathrm{d}^n x\sqrt{|g|}\left[R_{\mu\nu}-\frac{1}{2}Rg_{\mu\nu}\right]\delta g^{\mu\nu}, \tag{123}$$

and by the definition of the functional derivative,

$$\frac{\delta S_{\mathrm{EH}}}{\delta g^{\mu\nu}} = k\left(R_{\mu\nu}-\frac{1}{2}Rg_{\mu\nu}\right). \tag{124}$$

3.8 Nöther's theorem

One of the most profound consequences of symmetry in physics is the appearance of conserved quantities. Assume we have a set of fields φ_i. Nöther's theorem tells us that if the Lagrangian $\mathcal{L}[\varphi_i]$ has a continuous symmetry, then there is a corresponding conserved current [**carrollSpacetimeGeometryIntroduction2019**, **peskinIntroductionQuantumField1995**]. Consider an infinitesimal transformation,

$$\varphi_i(x) \longrightarrow \varphi_i'(x) = \varphi_i(x) + \delta\varphi_i(x). \tag{125}$$

Applying this transformation to the Lagrangian will in general change its form,

$$\mathcal{L}[\varphi] \to \mathcal{L}[\varphi'] = \mathcal{L}[\varphi] + \delta\mathcal{L}. \tag{126}$$

We assume this transformation is a symmetry, i.e.,

$$\delta\mathcal{L} = \partial_\mu K^\mu. \tag{127}$$

By considering the Lagrangian as a function of the field and its derivatives, $\mathcal{L} = \mathcal{L}(\varphi_i, \partial_\mu\varphi_i)$, we can write the difference term as a Taylor expansion around $(\varphi_i, \partial_\mu\varphi_i)$,

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi_i}\delta\varphi_i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_i)}\delta(\partial_\mu\varphi_i), \tag{128}$$

where $\delta(\partial_\mu\varphi_i) = \partial_\mu\varphi_i' - \partial_\mu\varphi_i$. By the linearity of the derivative,

$$\delta(\partial_\mu\varphi_i) = \partial_\mu\varphi_i' - \partial_\mu\varphi_i = \partial_\mu(\varphi_i' - \varphi_i) = \partial_\mu\delta\varphi_i. \tag{129}$$

With this, and the Euler-Lagrange equations

$$\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_i)} - \frac{\partial\mathcal{L}}{\partial\varphi_i} = 0, \tag{130}$$

we can rewrite

$$\delta\mathcal{L} = \left(\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_i)}\right)\delta\varphi_i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_i)}(\partial_\mu\delta\varphi_i) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_i)}\delta\varphi_i\right). \tag{131}$$

If we define the current

$$J^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_i)}\delta\varphi_i(x) - K^\mu, \tag{132}$$

then combining Equation 127 and Equation 131 gives

$$\partial_\mu J^\mu = 0. \tag{133}$$

This is Nöther's theorem; a continuous symmetry implies the existence of a conserved current.

The current flux through some spacelike surface V defines a conserved charge. The surface of constant time in some reference frame has the normal vector $n_\mu = (1, 0, 0, 0)$, so the charge is

$$Q = \int_V \mathrm{d}^4x\, n_\mu J^\mu = \int_V \mathrm{d}^3x\, J^0. \tag{134}$$

We can then use the divergence theorem. Assume ∂V is the boundary of V, which has the space-like normal vector k_i, and that the current falls quickly towards infinity. Then, using $\partial_\mu J^\mu = 0$,

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \int_V \mathrm{d}^3\,\partial_0 J^0 = -\int_V \mathrm{d}^3x\,\partial_i J^i = -\int_{\partial V} \mathrm{d}^2x\, k_i J^i = 0, \tag{135}$$

proving that the charge is conserved over time. These results, however, rely on the equations of motion. This is generally valid in classical mechanics. However, in quantum mechanics, the path integral is over *all* paths, not just those that are on-shell, and we can therefore not assume the equations of motion vanish. We will later derive the quantum version of the conservation law Equation 133.

4 General Relativity

General relativity describes how matter and energy curve the fabric of space and time. Einstein first wrote down the theory more than a century ago, and it is still our most accurate theory of gravitational effects.

integrate old notes with new

4.1 General Relatvity — Old notes

Einsteins field equations are given by the action ($g = \text{diag}(1,-1,-1,-1)$)$S[g_{\mu\nu}] = S_H + S_M$. Here, S_M is the action of matter and energy, which gives the (non-cannonical) energy-momentum tensor

$$T_{\mu\nu} = \frac{-2}{\sqrt{g}}\frac{\delta S_M}{\delta g^{\mu\nu}} \implies \delta S_M[g_{\mu\nu}] = \int_\Omega \mathrm{d}^4x(-2T_{\mu\nu})\delta g^{\mu\nu}$$

The Hilbert Lagrangian is

$$\mathcal{L}_h[g_{\mu\nu}] = \frac{1}{2\kappa}(R - 2\Lambda),$$

thus vary Hamilton pricinple yields

$$\begin{aligned} 2\kappa\,\delta S_H &= \delta\int_\Omega \mathrm{d}^4x\,\sqrt{g}\frac{1}{2\kappa}\,(R-2\Lambda) = \int_\Omega \mathrm{d}^4x\,\left(\sqrt{g}\delta(g^{\mu\nu}R_{\mu\nu}) - \frac{1}{2}\sqrt{g}g_{\mu\nu}\delta g^{\mu\nu}\,[R+2\Lambda]\right) \\ &= \int_\Omega \mathrm{d}^4x\,\sqrt{g}\left(R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu}\right) + \int_\Omega \mathrm{d}^4x\,\sqrt{g}\,\nabla_\eta g^\mu{}_\rho\left(g^\eta{}_\mu\delta\Gamma^\rho_{\nu\sigma} - g^\eta{}_\nu\delta\Gamma^\rho_{\mu\sigma}\right) = -\delta S_M \end{aligned}$$

This gives us the Einstein field equation

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

The Einstein equation encodes the information of the trace of the Riemann tensor. Taking the trace-reverse of the equation, and Weyl tensor and the Bianchi identity implies $\nabla_\lambda R_{\rho\sigma\mu\nu} = -\nabla_\rho R_{\sigma\lambda\mu\nu} - \nabla_\sigma R_{\lambda\rho\mu\nu}$ and $\nabla_\mu R^\mu{}_\nu = \frac{1}{2}\nabla_\nu R$, we find using the symmetries of the Reimann tensor that $g^{\rho\lambda}\nabla_\lambda R_{\rho\sigma\mu\nu} = 2\nabla_{[\mu}R_{\nu]\sigma}$, and thus that

$$\begin{aligned} \nabla^\rho C_{\rho\sigma\mu\nu} &= g^{\lambda\rho}\nabla_\lambda R_{\mu\nu\rho\sigma} - \frac{2}{n-2}\left(\nabla_{[\mu}R_{\nu]\sigma} - \nabla^\rho g_{\sigma[\mu}R_{\nu]\rho}\right) + \frac{2}{(n-1)(n-2)}\nabla_{[\mu}g_{\nu]\sigma}R. \\ &= 2\frac{n-3}{n-2}\nabla_{[\mu}\left(R_{\nu]\sigma} + \frac{1}{2(n-1)}g_{\nu]\sigma}R\right) := 2\frac{n-3}{n-2}C_{[\mu\nu]\sigma} \end{aligned}$$

This is the Cotton tensor. Using the fact that $g^{\mu\nu}T_{\mu\nu} := T = -1/\kappa(R - n\Lambda)$, and trace reversing the Einstein equation, $R_{\mu\nu} = \kappa(T_{\mu\nu} - 1/2Tg_{\mu\nu} - \Lambda g_{\mu\nu})$, we get a differential relationship between the Weyl tensor and the energy momentum tensor,

$$\nabla^\rho C_{\rho\sigma\mu\nu} = 2\kappa\frac{n-3}{n-2}\nabla_{[\mu}\left(T_{\nu]\sigma} - \frac{n}{2(n-1)}Tg_{\nu]\sigma}\right)\,(?????)$$

4.2 Newtonian Gravity

This section is based on [**carrollSpacetimeGeometryIntroduction2019**].

Newton's famous law of gravity states that the force of gravity from an object of mass M acting on another object of mass m is proportional to both objects' masses and is inversely proportional to the distance between them squared. This force is directed radially inwards. Formulating this as an equation, with $\vec{r}$ as the vector from the object with mass M to that with mass m with length $|\vec{r}| = r$, gives

$$\vec{F}_g = -\frac{GMm}{r^2}\hat{r}. \tag{136}$$

Here, G is Newton's gravitation constant, and $\hat{r} = \vec{r}/r$. The vector $\vec{r}$ is a purely spatial three-vector, as space is separated from time in the Newtonian picture. The law of gravitation, together with Newton's second law of motion

$$\sum_i \vec{F}_i = m\vec{a}, \tag{137}$$

where $\vec{a}$ is the acceleration of an object and $\vec{F}_i$ the forces acting upon it, can account for the motion of stellar objects. These laws work well in all but the most extreme circumstances, involving very massive objects or very short distances. As we will see, such extreme circumstances can be quantified by $2GM \approx r$. Newtonian gravity works well as an approximation because G is small in everyday units. Highly precise measurements of the orbit of Mercury were needed before any deviation from it was noticed.

We restate Newtonian gravity in a field-theoretic language to better compare Newtonian gravity to its successor theory. The gravitational potential from a mass M, which gives rise to Newton's force law, obeys the equation

$$\nabla^2 \Phi_g = -4\pi G\rho. \tag{138}$$

Here, ρ is the mass density. For a single point mass, $\rho(\vec{r}) = M\delta(\vec{r})$, this has the solution $\Phi_g = -GM/r$. The acceleration due to gravity is then

$$\vec{a} = -\vec{\nabla}\Phi_g. \tag{139}$$

We see that mass acts as a gravitational charge. Due to the success of Newtonian gravity, we expect it to be a limit of any theory that succeeds it. This gives us the ability to theoretically test any new theory of gravity, as well as to connect parameters in the new theory to old, known quantities.

4.3 General relativity

This section is based on [**carrollSpacetimeGeometryIntroduction2019**, **glendenningCompactStarsNuclear2012**]

The derivation of the spherically symmetric metric is done using computer code, as described in **??**.

4.4 Einstein's field equations

General relativity describes spacetime as a smooth manifold $\mathcal{M}$, with a (pseudo-Riemannian) metric, $g_{\mu\nu}$. This metric is treated as a dynamical field, which is affected by the presence of matter and energy. The matter and energy contents of spacetime are encoded in the stress-energy tensor $T_{\mu\nu}$, while the behavior of $g^{\mu\nu}$ is encoded in a scalar Lagrangian density. We employ the minimal-coupling rule to reformulate laws from special relativity to curved spacetime. This rule states that laws written in an inertial frame in a coordinate-independent way remain true in curved spacetime. In an inertial frame, the Crisoffel-symbols vanish, so $\nabla_\mu = \partial_\mu$. We can thus write any laws containing partial derivatives in a coordinate impenitent way by exchanging them for covariant derivatives. To generalize Newton's second law Equation 137, we must first make it relativistic by introducing a 4-force, $F^\mu = \frac{\mathrm{d}}{\mathrm{d}\tau}p^\mu$, where p^μ is the 4-momentum. When applying the minimal coupling rule, Newton's second law then becomes Equation 37, [**hartleGravityIntroductionEinstein2021**]

$$\sum_i F_i^\mu = m\left(\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu_{\nu\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}\right). \tag{140}$$

In the absence of any external forces, objects will follow geodesics in spacetime. With this, we must now find the law governing $g^{\mu\nu}$. As this is a field, we will do this by assigning it a Lagrangian density. The most obvious—and correct—choice as the Lagrangian is the Ricci scalar, which results in the Einstein-Hilbert action,

$$S_{\mathrm{EH}} = \frac{1}{2\kappa}\int_{\mathcal{M}} \mathrm{d}^n x\,\sqrt{|g|}\,R. \tag{141}$$

The $\sqrt{|g|}$-factor is included for the integral to be coordinate-independent, as discussed in subsection 1.5.[12] The κ is a constant and decides how strong the coupling of gravity to matter and energy is. This constant can then be related to Newton's constant of gravitation G by $\kappa = 8\pi G$. When including the contributions from other fields with an action S_{m}, the total action becomes

$$S = S_{\mathrm{EH}} + S_{\mathrm{m}}. \tag{142}$$

The equations of motion of the dynamical field, which in this case is the metric, are given by Hamilton's principle of stationary action. Using functional derivatives, as defined in subsection 3.6, this is stated as

$$\frac{\delta S}{\delta g^{\mu\nu}} = 0. \tag{143}$$

We define the stress-energy tensor as

$$T_{\mu\nu} = -2\frac{\delta S_{\mathrm{m}}}{\delta g^{\mu\nu}}. \tag{144}$$

[12]The gravitational action can also include a cosmological constant, modifying the Lagrangian to $R + 2\Lambda$. This constant does not affect the subject of this thesis and is therefore not included here.

The functional derivative of the Einstein-Hilbert action is evaluated in subsection 3.7, and with the result, Equation 124, we get the Einstein field equations

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \kappa T_{\mu\nu}. \tag{145}$$

The left-hand side of the Einstein field equations is called the Einstein tensor, $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$. This tensor obeys the identity

$$\nabla^\mu G_{\mu\nu} = 0, \tag{146}$$

as a consequence of the more general Bianchi identity, Equation 42.

4.5 Spherically symmetric spacetime

To model stars, we will assume that the metric is spherically symmetric and time-independent. In this case, the most general metric can be written, at least locally, as [**carrollSpacetimeGeometryIntroduction2019**]

$$\mathrm{d}s^2 = e^{2\alpha(r)}\mathrm{d}t^2 - e^{2\beta(r)}\mathrm{d}r^2 - r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2), \tag{147}$$

where α and β are general functions of the radial coordinate r. In matrix form, this corresponds to

$$g_{\mu\nu} = \begin{pmatrix} e^{2\alpha(r)} & 0 & 0 & 0 \\ 0 & -e^{2\beta(r)} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2(\theta) \end{pmatrix}. \tag{148}$$

Using Equation 35, we can now compute the Christoffel symbols in terms of the unknown functions. These computations in this subsection are done using computer code, which is shown in **??**. The results are

$$\Gamma^t_{\mu\nu} = \begin{pmatrix} 0 & \frac{\mathrm{d}}{\mathrm{d}r}\alpha(r) & 0 & 0 \\ \frac{\mathrm{d}}{\mathrm{d}r}\alpha(r) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{149}$$

$$\Gamma^r_{\mu\nu} = \begin{pmatrix} e^{2\alpha(r)}e^{-2\beta(r)}\frac{\mathrm{d}}{\mathrm{d}r}\alpha(r) & 0 & 0 & 0 \\ 0 & \frac{\mathrm{d}}{\mathrm{d}r}\beta(r) & 0 & 0 \\ 0 & 0 & -re^{-2\beta(r)} & 0 \\ 0 & 0 & 0 & -re^{-2\beta(r)}\sin^2(\theta) \end{pmatrix}, \tag{150}$$

$$\Gamma^\theta_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{r} & 0 \\ 0 & \frac{1}{r} & 0 & 0 \\ 0 & 0 & 0 & -\sin(\theta)\cos(\theta) \end{pmatrix}, \tag{151}$$

$$\Gamma^\phi_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{r} \\ 0 & 0 & 0 & \frac{\cos(\theta)}{\sin(\theta)} \\ 0 & \frac{1}{r} & \frac{\cos(\theta)}{\sin(\theta)} & 0 \end{pmatrix}. \tag{152}$$

The symbols not included are zero. Substituting these results into Equation 39 gives the Riemann curvature tensor. We can then obtain the Ricci tensor by taking the trace, as shown in Equation 43. The results are

$$R_{tt} = \left[r\left(\frac{\mathrm{d}}{\mathrm{d}r}\alpha(r)\right)^2 - r\frac{\mathrm{d}}{\mathrm{d}r}\alpha(r)\frac{\mathrm{d}}{\mathrm{d}r}\beta(r) + r\frac{\mathrm{d}^2}{\mathrm{d}r^2}\alpha(r) + 2\frac{\mathrm{d}}{\mathrm{d}r}\alpha(r)\right]\frac{e^{2\alpha(r)}e^{-2\beta(r)}}{r}, \tag{153}$$

$$R_{rr} = -\frac{1}{r}\left[r\left(\frac{\mathrm{d}}{\mathrm{d}r}\alpha(r)\right)^2 - r\frac{\mathrm{d}}{\mathrm{d}r}\alpha(r)\frac{\mathrm{d}}{\mathrm{d}r}\beta(r) + r\frac{\mathrm{d}^2}{\mathrm{d}r^2}\alpha(r) - 2\frac{\mathrm{d}}{\mathrm{d}r}\beta(r)\right], \tag{154}$$

$$R_{\theta\theta} = -\left[r\frac{\mathrm{d}}{\mathrm{d}r}\alpha(r) - r\frac{\mathrm{d}}{\mathrm{d}r}\beta(r) - e^{2\beta(r)} + 1\right]e^{-2\beta(r)}, \tag{155}$$

$$R_{\varphi\varphi} = R_{\theta\theta}\sin^2(\theta). \tag{156}$$

All other components are zero. Finally, the trace of the Ricci tensor gives the Ricci scalar,

$$R = \frac{2e^{-2\beta(r)}}{r^2}\Bigg[r^2\left(\frac{\mathrm{d}}{\mathrm{d}r}\alpha(r)\right)^2 - r^2\frac{\mathrm{d}}{\mathrm{d}r}\alpha(r)\frac{\mathrm{d}}{\mathrm{d}r}\beta(r) + r^2\frac{d^2}{dr^2}\alpha(r) + 2r\frac{\mathrm{d}}{\mathrm{d}r}\alpha(r) - 2r\frac{\mathrm{d}}{\mathrm{d}r}\beta(r) - e^{2\beta(r)} + 1\Bigg]. \tag{157}$$

The unknown functions α and β are now determined by the matter and energy content of the universe, which is encoded in $T_{\mu\nu}$, through Einstein's field equation, Equation 145.

4.6 The Schwarzschild metric

The simples case for a matter distribution in spacetime is $T_{\mu\nu} = 0$. Although this might only seem to be useful to model a non-empty universe, it can be combined with a central point particle and empty space elsewhere. In this case, the Einstein equations are simply $R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 0$. We can show that the trace of the Ricci tensor is zero by taking the trace of this equation, simplifying it to $R_{\mu\nu} = 0$. By combining Equation 153 and Equation 154, we find

$$R_{tt} + e^{2(\alpha-\beta)}R_{rr} = 2\frac{\mathrm{d}}{\mathrm{d}r}(\alpha+\beta) = 0, \tag{158}$$

which implies $\alpha = -\beta + \text{const}$. The constant corresponds to rescaling the coordinates, which allows us to set it to zero. From Equation 155, we get

$$e^{2\beta}R_{\theta\theta} = -2r\frac{\mathrm{d}}{\mathrm{d}r}\alpha - e^{-2\alpha} + 1 = 0, \tag{159}$$

which may be restated as

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(re^{2\alpha}\right) = 1. \tag{160}$$

This equation has the solution

$$e^{2\alpha(r)} = e^{-2\beta(r)} = \left(1 - \frac{R_s}{r}\right), \tag{161}$$

where R_s, the Schwarzschild radius, is a constant. With this, we should obtain Newton's law of gravity with a small perturbation, $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, and at slow speeds, $\frac{\mathrm{d}*}{\mathrm{d}x^i}\tau \ll \frac{\mathrm{d}*}{\mathrm{d}t}\tau$. Inserting this into the geodesic equation Equation 140 with $F_i = 0$, using $\partial_0 g_{\mu\nu} = 0$ and expanding to leading order, we get

$$\frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} = \frac{1}{2}\eta^{ij}\partial_j h_{00}. \tag{162}$$

We now obtain Newtonian gravity, as formulated in Equation 139, if we identify $h_{00} = 1 - e^{2\alpha} = -2\Phi_g$. The Schwarzschild radius is thus $R_s = 2GM$, and we see that this solution corresponds to a point-mass M located at $\vec{r} = 0$. We will extend our discussion to finite densities in the next section. Inserting our results into the metric, we get the Schwarzschild metric

$$\mathrm{d}s^2 = \left(1 - \frac{2GM}{r}\right)\mathrm{d}t^2 - \left(1 - \frac{2GM}{r}\right)^{-1}\mathrm{d}r^2 - r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right). \tag{163}$$

5 QFT

5.1 QFT via path integrals

This section is based on [**peskinIntroductionQuantumField1995**, **weinbergQuantumTheoryFields1995**, **weinbergQuantumTheoryFields1996**, **schwartzQuantumFieldTheory2013**]

In the path integral formalism, one evaluates quantum observable by integrating the contributions of all possible configurations. If the system has specified initial and final states, this amounts to all possible paths the system might evolve between these, hence the name. We assume the reader has some familiarity with this formalism. However, if a refresher is needed, **??** contains a derivation of the closely related

imaginary-time formalism and compares it with the path integral approach. A summary of functional calculus is given in subsection 3.2.

In the path integral formalism, the vacuum-to-vacuum transition amplitude, i.e., overlap between the vacuum at $t = -\infty$ and the vacuum at time $t = \infty$, is given by

$$\begin{aligned} Z &= \lim_{T\to\infty} \langle \Omega, T/2 | -T/2, \Omega\rangle \\ &= \lim_{T\to\infty} \left\langle \Omega \left| e^{-iHT} \right| \Omega \right\rangle \\ &= \int \mathcal{D}\pi\mathcal{D}\varphi \, \exp i \int \mathrm{d}^4x \, (\pi\dot{\varphi} - \mathcal{H}[\pi,\varphi]), \end{aligned}$$

where $|\Omega\rangle$ is the vacuum state. The φ are the fields of the theory, and π their canonical momenta. We will work as if φ are a bosonic field. However, this can be readily generalized to fermions. By introducing a source term into the Hamiltonian density, $\mathcal{H} \to \mathcal{H} - J(x)\varphi(x)$, we get the generating functional

$$Z[J] = \int \mathcal{D}\pi\mathcal{D}\varphi \, \exp i \int \mathrm{d}^4x \, (\pi\dot{\varphi} - \mathcal{H}[\pi,\varphi] + J\varphi). \tag{164}$$

If $\mathcal{H}$ is quadratic in π, we can complete the square and integrate out π to obtain

$$Z[J] = C \int \mathcal{D}\varphi \, \exp i \int \mathrm{d}^4x \, (\mathcal{L}[\varphi] + J\varphi). \tag{165}$$

C is infinite, but constant, and will drop out of physical quantities. In scattering theory, the main objects of study are correlation functions $\langle\varphi(x_1)\varphi(x_2)...\rangle = \langle\Omega|T\{\varphi(x_1)\varphi(x_2)\dots\}|\Omega\rangle$, where T is the time ordering operator. These are given by functional derivatives of $Z[J]$,

$$\langle\varphi(x_1)\varphi(x_2)...\rangle = \frac{\int \mathcal{D}\varphi(x)\,[\varphi(x_1)\varphi(x_2)...]e^{iS[\varphi]}}{\int \mathcal{D}\varphi(x)\,e^{iS[\varphi]}} = \frac{1}{Z[0]}\prod_i \left(-i\frac{\delta}{\delta J(x_i)}\right) Z[J]\Big|_{J=0}, \tag{166}$$

where $S[\varphi] = \int \mathrm{d}^4x\,\mathcal{L}[\varphi]$ is the action of the theory. The functional derivative is described in subsection 3.2. In a free theory, we are able to write

$$Z_0[J] = Z_0[0]\exp iW_0[J], \quad iW_0[J] = -\frac{1}{2}\int \mathrm{d}^4x\mathrm{d}^4y\, J(x)D_0(x-y)J(y), \tag{167}$$

where D_0 is the propagator of the free theory. Using this form of the generating functional, Equation 166 becomes

$$\begin{aligned} &\frac{1}{Z[0]}(-i)^n\frac{\delta}{\delta J(x_1)}\cdots\frac{\delta}{\delta J(x_n)}Z_0[J]\Big|_{J=0} = (-i)^n\frac{\delta}{\delta J(x_1)}\cdots\frac{\delta}{\delta J(x_n)}e^{iW_0[J]}\Big|_{J=0} \\ &= (-i)^n\frac{\delta}{\delta J(x_1)}\cdots\frac{\delta}{\delta J(x_{n-1})}\left(i\frac{\delta W_0[J]}{\delta J(x_n)}\right)e^{iW_0[J]}\Big|_{J=0} \\ &= (-i)^n\frac{\delta}{\delta J(x_1)}\cdots\frac{\delta}{\delta J(x_{n-2})}\left(i\frac{\delta^2 W_0[J]}{\delta J(x_{n-1})\,\delta J(x_n)} + i^2\frac{\delta W_0[J]}{\delta J(x_{n-1})}\frac{\delta W_0[J]}{\delta J(x_n)}\right)e^{iW_0[J]}\Big|_{J=0} \\ &= \dots \\ &= (-i)^{\lfloor n/2\rfloor}\sum_{(a,b)}\prod_{i=1}^{\lfloor n/2\rfloor}\frac{\delta^2 W_0[J]}{\delta J(x_{a(i)})\,\delta J(x_{b(i)})}\Big|_{J=0}. \end{aligned}$$

In the last line, we have introduced the functions a, b, which define a way to pair n elements. $\lfloor\cdot\rfloor$ is the floor function. The domain of these functions are the integers between 1 and $\lfloor n/2\rfloor$, the image a subset of the integers between 1 and n of size $\lfloor n/2\rfloor$. A valid pairing is a set $\{(a(1),b(1)),\dots(a(\lfloor n/2\rfloor),b(\lfloor n/2\rfloor))\}$, where all elements $a(i)$ and $b(j)$ are different, such that all integers up to and including n are featured. A pair is not directed, so $(a(i),b(i))$ is the same pair as $(b(i),a(i))$. The sum is over the set $\{(a,b)\}$ of all possible, unique pairings. If n is odd, the expression is equal to 0. This is Wick's theorem, and it can more simply be stated as *a correlation function is the sum of all possible pairings of 2-point functions*,

$$\left\langle\prod_{i=1}^n\varphi(x_i)\right\rangle_0 = \sum_{\{(a,b)\}}\prod_{i=1}^{\lfloor n/2\rfloor}\left\langle\varphi(x_{a(i)})\varphi(x_{b(i)})\right\rangle_0. \tag{168}$$

The subscript on the expectation value indicates that it is evaluated in the free theory.

If we have an interacting theory, that is, a theory with an action $S = S_0 + S_I$, where S_0 is a free theory, the generating functional can be written

$$Z[J] = Z_0[0] \left\langle \exp iS_I + i \int \mathrm{d}^4x \, J(x)\varphi(x) \right\rangle_0 . \tag{169}$$

We can expand the exponential in power series, which means the expectation value in Equation 169 becomes

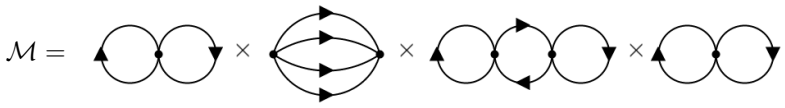
$$\sum_{n,m} \frac{1}{n!m!} \left\langle (iS_I)^n \left(i \int \mathrm{d}^4x \, J(x)\varphi(x) \right)^m \right\rangle_0 . \tag{170}$$

The terms in this series are represented by Feynman diagrams, constructed using the Feynman rules, and can be read from the action. We will not further detail how the Feynman rules are derived. The Feynman rules for a free scalar field in thermal field theory are derived in **??**, and the general procedure is found in any of the main sources for this section [**peskinIntroductionQuantumField1995**, **schwartzQuantumFieldTheory2013**, **weinbergQuantumTheoryFields1995**, **weinbergQuantumTheoryFields**
The source terms give rise to an additional vertex

(171)

The generating functional $Z[J]$ thus equals $Z_0[0]$ times *the sum of all diagrams with external sources* $J(x)$.

Consider a general diagram without external legs, built up of N different connected subdiagrams, where subdiagram i appears n_i times. As an illustration, a generic vacuum diagram in φ^4-theory has the form



(172)

If sub-diagram i as a stand-alone diagram equals $\mathcal{M}_i$, each copy of that subdiagram contributes a factor $\mathcal{M}_i$ to the total diagram. However, due to the symmetry of permuting identical subdiagrams, one must divide by the extra symmetry factor $s = n_i!$, the total number of permutations of all the copies of diagram i. The full diagram therefore equals

$$\mathcal{M} = \prod_{i=1}^{N} \frac{1}{n_i!} \mathcal{M}_i^{n_i}. \tag{173}$$

$\mathcal{M}$ is uniquely defined by a finite sequence of integers, $(n_1, n_2, \ldots n_N, 0, 0, \ldots)$, so the sum of all diagrams is the sum over the set S of all finite sequences of integers. This allows us to write the sum of all diagrams as

$$\sum_{(n_1,\ldots)\in S} \prod_i \frac{1}{n_i!} \mathcal{M}_i^{n_i} = \prod_{i=1}^{\infty} \sum_{n_i=1}^{\infty} \frac{1}{n_i!} \mathcal{M}_i^{n_i} = \exp \sum\nolimits_i \mathcal{M}_i. \tag{174}$$

We showed that the generating functional $Z[J]$ were the $Z_0[0]$ times the sum of all diagrams due to external sources. From Equation 174, if we define

$$Z[J] = Z_0[0] \exp iW[J], \tag{175}$$

then $W[J]$ is the sum of all connected diagrams. This is trivially true for the free theory, where the only connected diagram is

(176)

The two-point function in the full, interacting theory can thus be written

$$-i \frac{\delta^2 W[J]}{\delta J(x)\, \delta J(y)} = D(x - y). \tag{177}$$

5.2 The 1PI effective action

This section is based on [**peskinIntroductionQuantumField1995**, **schwartzQuantumFieldTheory2013**, **weinbergQuantumTheoryFields1995**, **weinbergQuantumTheoryFields1996**].

The generating functional for connected diagrams, $W[J]$, is dependent on the external source current J. We can define a new quantity with a different independent variable, using the Legendre transformation analogously to what is done in thermodynamics and Lagrangian mechanics. The new independent variable is

$$\varphi_J(x) := \frac{\delta W[J]}{\delta J(x)} = \langle \varphi(x) \rangle_J \,. \tag{178}$$

The subscript J on the expectation value indicate that it is evaluated in the presence of a source. The Legendre transformation of W is then

$$\Gamma[\varphi_J] := W[J] - \int \mathrm{d}^4x \, J(x)\varphi_J(x). \tag{179}$$

Using the definition of φ_J, we have that

$$\frac{\delta\Gamma[\varphi_J]}{\delta\varphi_J(x)} = \int \mathrm{d}^4y \, \frac{\delta J(y)}{\delta\varphi_J(x)} \frac{\delta W[J]}{\delta J(y)} - \int \mathrm{d}^4y \, \frac{\delta J(y)}{\delta\varphi_J(x)}\varphi_J(y) - J(x) = -J(x). \tag{180}$$

If we compare this to the classical equations of motion of a field φ with the action S,

$$\frac{\delta S[\varphi]}{\delta\varphi(x)} = -J(x), \tag{181}$$

we see that Γ is an action that gives the equation of motion for the expectation value of the field, given a source current $J(x)$.

To interpret Γ further, we observe what happens if we treat $\Gamma[\varphi]$ as a classical action with a coupling g. The generating functional in this new theory is

$$Z[J,g] = \int \mathcal{D}\varphi \, \exp ig^{-1} \left(\Gamma[\varphi] + \int \mathrm{d}^4x \, \varphi(x)J(x) \right). \tag{182}$$

The free propagator in this theory will be proportional to g, as it is given by the inverse of the equation of motion for the free theory. All vertices in this theory, on the other hand, will be proportional to g^{-1}, as they are given by the higher-order terms in the action $g^{-1}\Gamma$. This means that a diagram with V vertices and I internal lines is proportional to g^{I-V}. Regardless of what the Feynman-diagrams in this theory are, the number of loops of a connected diagram is[13]

$$L = I - V + 1. \tag{183}$$

To see this, we first observe that diagrams with one single loop must have equally many internal lines as vertices, so the formula holds for $L = 1$. The formula still holds if we add a new loop to a diagram with n loops by joining two vertices. If we attach a new vertex with one line, the formula still holds, and as the diagram is connected, any more lines connecting the new vertex to the diagram will create additional loops. This ensures that the formula holds by induction. As a consequence of this, any diagram is proportional to g^{L-1}. This means that in the limit $g \to 0$, the theory is fully described at the tree-level, i.e., by only considering diagrams without loops. In this limit, we may use the stationary phase approximation, as described in subsection 3.2, which gives

$$Z[J, g \to 0] \sim C \det \left(-g \frac{\delta^2\Gamma[\varphi_J]}{\delta\varphi(x)\,\delta\varphi(y)} \right)^{-1/2} \exp ig^{-1} \left(\Gamma[\varphi_J] + \int \mathrm{d}^4x \, J(x)\varphi_J(x) \right). \tag{184}$$

This means that

$$-ig \ln(Z[J,g]) = gW[J,g] = \Gamma[\varphi_J] + \int \mathrm{d}^4x \, J(x)\varphi_J(x) + \mathcal{O}(g), \tag{185}$$

which is exactly the Legendre transformation we started out with, modulo the factor g. Γ is, therefore, the action that describes the full theory at the tree level. For a free theory, the classical action S equals the effective action.

As we found in the last section, the propagator $D(x,y) = \langle \varphi(x)\varphi(y) \rangle_J$ is given by $-i$ times the second functional derivative of $W[J]$. Using the chain rule, together with Equation 180, we get

$$\begin{aligned} &(-i) \int \mathrm{d}^4z \frac{\delta^2 W[J]}{\delta J(x)\delta J(z)} \frac{\delta^2\Gamma[\varphi_J]}{\delta\varphi_J(z)\,\delta\varphi_J(y)} \\ &\quad = (-i) \int \mathrm{d}^4z \frac{\delta\varphi_J[z]}{\delta J(x)} \frac{\delta^2\Gamma[\varphi_J]}{\delta\varphi_J(z)\,\delta\varphi_J(y)} = (-i) \frac{\delta}{\delta J(x)} \frac{\delta\Gamma[\varphi_J]}{\delta\varphi_J(y)} = i\delta(x-y). \end{aligned}$$

[13] This is a consequence of the Euler characteristic $\chi = V - E + F$, a topological invariant of graphs.

This is exactly the definition of the inverse propagator, which means

$$\frac{\delta^2\Gamma[\varphi_J]}{\delta\varphi_J(x)\,\delta\varphi_J(y)} = D^{-1}(x,y). \tag{186}$$

The inverse propagator is the sum of all one-particle-irreducible (1PI) diagrams, with two external vertices. More generally, Γ is the generating functional for 1PI diagrams, which is why it is called the 1PI effective action.

We define η as the fluctuations around the expectation value of the field, $\varphi(x) = \varphi_J(x) + \eta(x)$, and use this to change variables of integration in the path integral. The expectation value φ_J is constant with respect to the integral, so

$$\int \mathcal{D}\varphi\, \exp iS[\varphi] = \int \mathcal{D}\eta\, \exp iS[\varphi_J + \eta]. \tag{187}$$

By assumption, $\langle\eta\rangle_J = 0$, which means this path integral is described by only 1PI diagrams, connected or not. We can therefore write

$$\exp i\Gamma[\varphi_J] = \int \mathcal{D}\eta\, \exp iS[\varphi_J + \eta]. \tag{188}$$

As we will discuss later, we interpret this form as *integrating out* the η degrees of freedom, leaving us with an effective description of the physics dependent only on the ground state solution φ_J. The disadvantage of this is that there is no bound on how complicated Γ might be—it is not restricted by the usual assumptions of the form of the action, such as locality [**schwartzQuantumFieldTheory2013**]. With some simplifying assumptions, though, we can still make use of the 1PI effective action, as we will see in the next subsection.

5.3 Effective potential

For a constant field configuration $\varphi(x) = \varphi_0$, the effective action, a functional, becomes a regular function. We define the effective potential $\mathcal{V}_{\text{eff}}$ by

$$\Gamma[\varphi_0] = -VT\,\mathcal{V}_{\text{eff}}(\varphi_0), \tag{189}$$

where VT is the volume of space-time. For a constant ground state, the effective potential will equal the energy of this state. To calculate the effective potential, we can expand the action around this state to calculate the effective action, by changing variables to $\varphi(x) = \varphi_0 + \eta(x)$. $\eta(x)$ now parametrizes fluctuations around the ground state and has by assumption a vanishing expectation value. The generating functional becomes

$$Z[J] = \int \mathcal{D}(\varphi_0 + \eta)\, \exp iS[\varphi_0 + \eta] + i\int \mathrm{d}^4x\, [\varphi_0 + \eta(x)]J(x). \tag{190}$$

The functional version of a Taylor expansion, as described in subsection 3.2, is

$$S[\varphi_0 + \eta] = S[\varphi_0] + \int \mathrm{d}x\, \eta(x)\, \frac{\delta S[\varphi_0]}{\delta\varphi(x)} + \frac{1}{2}\int \mathrm{d}x\mathrm{d}y\, \eta(x)\eta(y)\, \frac{\delta^2 S[\varphi_0]}{\delta\varphi(x)\delta\varphi(y)} + \dots. \tag{191}$$

The notation

$$\frac{\delta S[\varphi_0]}{\delta\varphi(x)} \tag{192}$$

indicates that the functional $S[\varphi]$ is differentiated with respect to $\varphi(x)$, then evaluated at $\varphi(x) = \varphi_0$. We define

$$S_0[\eta] := \frac{1}{2}\int \mathrm{d}^4x\mathrm{d}^4y\, \eta(x)\eta(y)\, \frac{\delta^2 S[\varphi_0]}{\delta\varphi(x)\,\delta\varphi(y)}, \tag{193}$$

$$S_I[\eta] := \frac{1}{6}\int \mathrm{d}^4x\mathrm{d}^4y\mathrm{d}^4z\, \eta(x)\eta(y)\eta(z)\, \frac{\delta^3 S[\varphi_0]}{\delta\varphi(x)\,\delta\varphi(y)\,\delta\varphi(z)} + \dots, \tag{194}$$

where the dots indicate higher derivatives. When we insert this expansion into the generating functional $Z[J]$ we get

$$Z[J] = \int \mathcal{D}\eta \exp\left\{ i\int \mathrm{d}^4x\, (\mathcal{L}[\varphi_0] + J\varphi_0) + i\int \mathrm{d}^4x\, \eta(x)\, \left(\frac{\delta S[\varphi_0]}{\delta\varphi(x)} + J(x)\right) + iS_0[\eta] + iS_I[\eta] \right\}. \tag{195}$$

The first term is constant with respect to η and may be taken outside the path integral. The second term gives rise to tadpole diagrams, which alter the expectation value of $\eta(x)$. For $J = 0$, this expectation value should vanish, and this term can be ignored. Furthermore, this means that the ground state must minimize the classical potential,

$$\frac{\partial \mathcal{V}(\varphi_0)}{\partial \varphi} = 0. \tag{196}$$

This leaves us with

$$-i \ln Z[J] = W[J] = \int \mathrm{d}^4x \left(\mathcal{L}[\varphi_0] + J\varphi_0\right) - i \ln \left(\int \mathcal{D}\eta \, \exp iS_0[\eta] + iS_I[\eta] \right). \tag{197}$$

We can now use the definition of the 1PI effective action to obtain a formula for the effective potential,

$$\mathcal{V}_{\text{eff}}(\varphi_0) = -\frac{1}{VT}\left(W[J] - \int \mathrm{d}^4x \, J(x)\varphi_0 \right) = \mathcal{V}(\varphi_0) - i \ln \left(\int \mathcal{D}\eta \, \exp iS_0[\eta] + iS_I[\eta] \right). \tag{198}$$

We showed that the 1PI effective action describes the whole quantum theory of the original action at the tree level. This was done by inspecting a theory with an action proportional to g^{-1}. In this theory, Feynman diagrams with L loops are proportional to g^{L-1}. We can use the same argument to expand the effective potential in loops. This is done by modifying the action $S[\varphi] \to g^{-1}S[\varphi]$, and then expand in power of g. The first term in the effective potential is modified by $\mathcal{V} \to g^{-1}\mathcal{V}$, which means that it is made up of tree-level terms. This is as expected since the tree-level result corresponds to the classical result without any quantum corrections. The second term becomes

$$\begin{aligned} \ln \left(\int \mathcal{D}\eta \, e^{iS_0 + iS_I} \right) &\longrightarrow \ln \left(\int \mathcal{D}\eta \, e^{ig^{-1}S_0 + ig^{-1}S_I} \right) \\ &= \ln \left(\int \mathcal{D}\eta \, e^{ig^{-1}S_0} \right) + \ln \left(\frac{\int \mathcal{D}\eta \, e^{ig^{-1}S_I} \, e^{ig^{-1}S_0}}{\int \mathcal{D}\eta \, e^{ig^{-1}S_0}} \right) \end{aligned}$$

The first term is quadratic in η, and can therefore be evaluated as a generalized Gaussian integral, as described in subsection 3.2,

$$\begin{aligned} &\ln \left[\int \mathcal{D}\eta \, \exp ig^{-1} \frac{1}{2} \int \mathrm{d}^4x \mathrm{d}^4y \, \eta(x)\eta(y) \, \frac{\delta^2 S[\varphi_0]}{\delta\varphi(x)\,\delta\varphi(y)} \right] \\ &= \ln \left[\det \left(-g^{-1} \frac{\delta^2 S[\varphi_0]}{\delta\varphi(x)\,\delta\varphi(y)} \right)^{-1/2} \right] = -\frac{1}{2} \mathrm{Trln} \left[-\frac{\delta^2 S[\varphi_0]}{\delta\varphi(x)\,\delta\varphi(y)} \right] + \text{const.} \end{aligned}$$

We used the identity $\ln\{\det(M)\} = \mathrm{Trln}(M)$. After we remove the constant, this term is proportional to g^0, i.e., it is made up of one-loop terms.

The last term can be evaluated by first expanding the exponential containing the S_I term, then using $\ln(1+x) = \sum_n \frac{1}{n} x^n$. Using

$$\langle A \rangle_0 = \frac{\int \mathcal{D}\varphi \, A \, e^{ig^{-1}S_0}}{\int \mathcal{D}\varphi \, e^{ig^{-1}S_0}}, \tag{199}$$

we can write

$$\frac{\int \mathcal{D}\eta \, e^{ig^{-1}S_I} e^{ig^{-1}S_0}}{\int \mathcal{D}\eta \, e^{ig^{-1}S_0}} = \sum_{n=0}^{\infty} \frac{1}{n!} \left\langle (ig^{-1}S_I)^n \right\rangle_0. \tag{200}$$

We recognize this as the sum of all Feynman diagrams, with Feynman rules from the interaction term S_I. The logarithm of this is then the sum of all connected Feynman diagrams. We know that S_I is made up of terms that are third power or higher in the fields. Each internal line is connected to two vertices, and each vertex is connected to at least three internal lines, i.e., $I \geq 3/2V$. The number of loops is therefore $L = I - V + 1 \geq (3/2 - 1)V + 1$. There is at leas one vertex, i.e. $L \geq 3/2$. This shows that the first logarithm contains *all* one-loop contributions. The effective potential at one-loop order is therefore

$$\mathcal{V}_{\text{eff}}(\varphi_0) = \mathcal{V}(\varphi_0) - \frac{i}{VT} \frac{1}{2} \mathrm{Trln} \left(-\frac{\delta^2 S[\varphi_0]}{\delta\varphi(x)\,\delta\varphi(y)} \right). \tag{201}$$

5.4 Symmetry and Goldstone's theorem

This section is based on [**leeIntroductionSmoothManifolds2003d**, **peskinIntroductionQuantumField1995**, **schwartzQuantumFieldTheory2013**, **weinbergQuantumTheoryFields1995**, **weinbergQuantumTheoryFields**

Symmetry plays a prominent role in modern physics. If we can transform a physical state in such a way that the governing equations of this system are unchanged, we call that transformation a *symmetry transformation*. All such transformations are known as the symmetries of that theory. The symmetries of a theory encode a lot of physics, such as the presence of conserved quantities and the system's low energy behavior. We distinguish between internal and external symmetries. An external symmetry is an active coordinate transformation, such as rotations or translations. They relate degrees of freedom at different space-time points, while internal symmetries transform degrees of freedom at each space-time point independently. A further distinction is between global and local symmetry transformations. Global transformations have one rule for transforming degrees of freedom at each point, which is applied everywhere, while local transformations are functions of space-time.

In classical field theory, symmetries are encoded in the behavior of the Lagrangian when the fields are transformed. We will consider continuous transformations, which can in general be written as

$$\varphi(x) \longrightarrow \varphi'(x) = f_t[\varphi](x), \quad t \in [0,1]. \tag{202}$$

Here, $f_t[\varphi]$ is a functional in φ, and a smooth function of t, with the constraint that $f_0[\varphi] = \varphi$. This allows us to look at "infinitesimal" transformations,

$$\varphi'(x) = f_\epsilon[\varphi] = \varphi(x) + \epsilon \frac{\mathrm{d} f_t[\varphi]}{\mathrm{d}t}\bigg|_{t=0} + \mathcal{O}(\epsilon). \tag{203}$$

When considering infinitesimal transformations, we will not always write $+\mathcal{O}(\epsilon)$, but rather consider it implicit. We will consider internal, global transformations which act linearly on φ. For N fields, φ_i, this can be written

$$\varphi_i'(x) = \varphi_i(x) + i\epsilon\, V_{ij}\varphi_j(x). \tag{204}$$

V_{ij} is called the generator of the transformation. A symmetry transformation of the system is then a transformation in which the Lagrangian left is unchanged, or at most differs by a 4-divergence term. That is, a transformation $\varphi \to \varphi'$ is a symmetry if

$$\mathcal{L}[\varphi'] = \mathcal{L}[\varphi] + \partial_\mu K^\mu[\varphi], \tag{205}$$

where $K^\mu[\varphi]$ is a functional of φ.[14] This is a requirement for symmetry in quantum field theory too. However, as physical quantities in quantum field theory are given not just by the action of a single state but the path integral, the integration measure $\mathcal{D}\varphi$ has to be invariant as well. If a classical symmetry fails due to the non-invariance of the integration measure, it is called an *anomaly*.

To investigate the symmetry properties of a quantum theory, we explore what constraints a symmetry imposes on the effective action. To that end, assume

$$\mathcal{D}\varphi'(x) = \mathcal{D}\varphi(x), \quad S[\varphi'] = S[\varphi]. \tag{206}$$

In the generating functional, such a transformation corresponds to a change of integration variable. Using the infinitesimal version of the transformation, we may write

$$\begin{aligned} Z[J] &= \int \mathcal{D}\varphi\, \exp\left\{ iS[\varphi] + i\int \mathrm{d}^4x\, J_i(x)\varphi_i(x) \right\} \\ &= \int \mathcal{D}\varphi'\, \exp\left\{ iS[\varphi'] + i\int \mathrm{d}^4x\, J_i(x)\varphi_i'(x) \right\} \\ &= Z[J] + i\epsilon \int \mathrm{d}^4x\, J_i(x) \int \mathcal{D}\varphi\, [V_{ij}\varphi_j(x)]\, e^{iS[\varphi]}. \end{aligned} \tag{207}$$

Using Equation 180, we can write this as

$$\int \mathrm{d}^4x\, \frac{\delta\Gamma[\varphi_J]}{\delta\varphi_i(x)}\, V_{ij}\, \langle\varphi_j(x)\rangle_J = 0. \tag{208}$$

[14] Terms of the form $\partial_\mu K^\mu$ do not affect the physics, as the variational principle $\delta S = 0$ does not vary the fields at infinity. Together with the divergence theorem, this means that such terms do not influence the equations of motion.

This constraint will allow us to deduce the properties of a theory close to the ground state, only using information about the symmetries of the theory.

The archetypical example of an internal, global, and continuous symmetry is the linear sigma model, which we will use as an example throughout this section. The linear sigma model is made up of N real scalar fields φ_i, whose Lagrangian is

$$\mathcal{L}[\varphi] = \frac{1}{2}\partial_\mu\varphi_i(x)\partial^\mu\varphi_i(x) - \mathcal{V}(\varphi), \quad \mathcal{V}(\varphi) = -\frac{1}{2}\mu^2\varphi_i(x)\varphi_i(x) + \frac{1}{4}\lambda[\varphi_i(x)\varphi_i(x)]^2. \tag{209}$$

This system is invariant under the rotation of the N fields into each other,

$$\varphi_i \longrightarrow \varphi_i' = M_{ij}\varphi_j, \quad M^{-1} = M^T. \tag{210}$$

The set of all such transformations forms the Lie group $\mathrm{O}(N)$. For $N = 2$, and $\mu^2, \lambda > 0$ we get the ubiquitous Mexican hat potential, as illustrated in Figure 6.

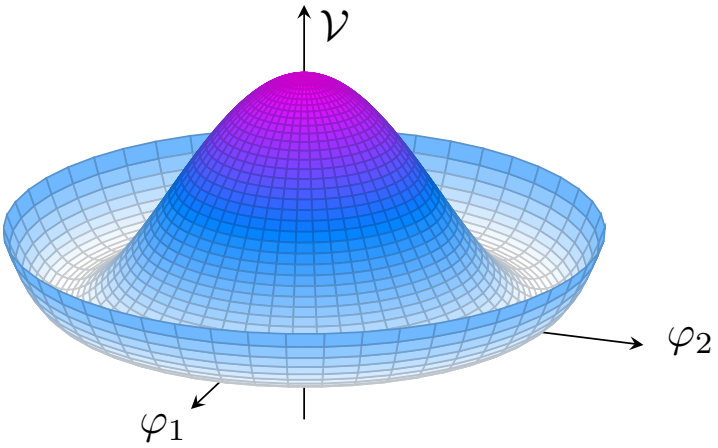


Figure 6: The Mexican hat potential is the classical potential $\mathcal{V}$ for the $N = 2$ linear sigma model.

5.5 Goldstone's theorem

A symmetry transformation will leave the governing equation of a theory unchanged. This, however, does not imply that physical states, such as the ground state, are invariant under this transformation. The $N = 2$ linear sigma model illustrates this. If we assume the ground state φ_0 is translationally invariant, then it is given by minimizing the effective potential, of which the classical potential, $\mathcal{V}$, is the leading order approximation. This potential is illustrated in Figure 6. The ground state is therefore given by any of the values along the brim of the potential. If we, without loss of generality, choose $\varphi_0 = (0, v)$ as the ground state, then any rotation will change this state. We say that the symmetry has been *spontaneously broken*.

To explore this in a general context, assume a theory of N real scalar fields φ_i are invariant under the actions of some Lie group, G. A symmetry $g \in G$ is broken if the vacuum expectation value of the original fields and the transformed fields differ. That is, if

$$\langle\varphi\rangle_0 \neq \langle\varphi'\rangle_0 = \langle g\varphi\rangle_0 \,. \tag{211}$$

We can now exploit what we learned about Lie groups to write the infinitesimal transformation as

$$\langle\varphi'\rangle_0 = \langle\varphi\rangle_0 + i\epsilon\eta_\alpha T_\alpha \langle\varphi\rangle_0 \,. \tag{212}$$

Let x_ℓ be the set of generators corresponding to broken symmetries, i.e.,

$$x_\ell \langle\varphi\rangle_0 \neq 0. \tag{213}$$

These are called the *broken generators*. The remaining set of generators t_a, which obey

$$t_a \langle\varphi\rangle_0 = 0, \tag{214}$$

are called unbroken, and generate a subgroup $H \subset G$ as the set of symmetry transformations of the vacuum is a group.

In Equation 208 we found that, if V is the generator of some symmetry, then the effective action obeys

$$\int \mathrm{d}^4x \, \frac{\delta\Gamma[\varphi_J]}{\delta\varphi_i} V_{ij} \left\langle \varphi_j \right\rangle_0 = 0. \tag{215}$$

We now differentiate this expression with respect to $\varphi_k(y)$, which gives

$$\int \mathrm{d}^4x \, \frac{\delta^2\Gamma[\varphi_0]}{\delta\varphi_k(y)\,\delta\varphi_i(x)} V_{ij} \left\langle \varphi_j \right\rangle_0 = 0. \tag{216}$$

With the assumption that the ground state is constant, we get

$$\frac{\partial^2 \mathcal{V}_{\text{eff}}}{\partial\varphi_k \partial\varphi_i} V_{ij} \left\langle \varphi_j \right\rangle_0 = 0. \tag{217}$$

This is trivial for unbroken symmetries, as $t^a_{ij} \left\langle \varphi_j \right\rangle_0 = 0$ by definition. However, in the case of broken symmetries, the second derivative of the effective potential must have an eigenvector $x^\ell_{ij} \left\langle \varphi_j \right\rangle_0$ with a zero eigenvalue for each broken generator. Here, ℓ labels the set of generators, while (ij) are the indices corresponding to field-components φ_i. We found that the second derivative of the effective action is the inverse propagator,

$$D^{-1}_{ij}(x,y) = \frac{\delta^2\Gamma[\varphi_0]}{\delta\varphi_i(y)\,\delta\varphi_j(x)} = \int \frac{\mathrm{d}^4p}{(2\pi)^4} e^{-ip(x-y)} \, \tilde{D}^{-1}_{ij}(p). \tag{218}$$

Using this, we can write

$$\tilde{D}^{-1}_{ij}(p=0)\, x^\ell_{jk} \left\langle \varphi_k \right\rangle_0 = 0. \tag{219}$$

Zeros of the inverse propagator correspond to the physical mass of particles. In Lorentz-invariant systems, each zero-eigenvalue vector corresponds to a massless particle, called a Goldstone boson.[15] This means there are $n_G = \dim G - \dim H$ zero-mass modes. In general, the counting of massless modes is complicated and depends on the dispersion relation of the particles at low momenta. Systems with Goldstone bosons with quadratic dispersion relation, that is $E \propto |\vec{p}|^2$ when $\vec{p} \to 0$, often exhibit a lower number of massless modes. An example is ferromagnetic, where the SU(2) rotational symmetry is broken down to U(1) when they align along one axis. This corresponds to two broken generators, yet the system exhibits only one massless mode [**braunerSpontaneousSymmetryBreaking2010**].

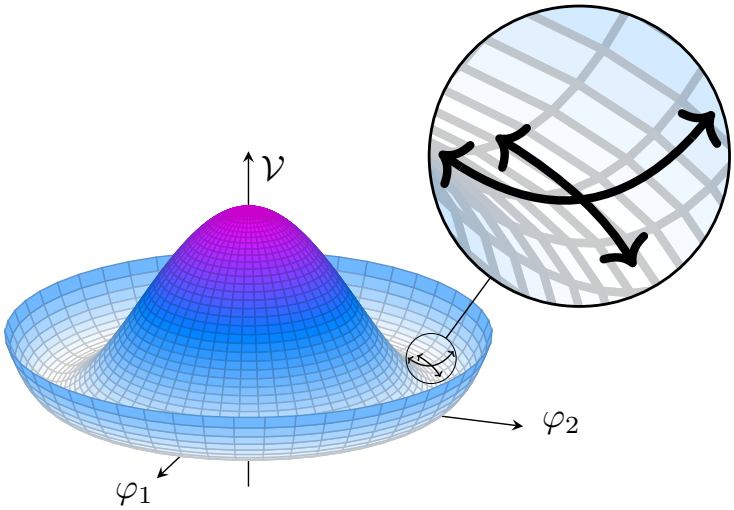


Figure 7: Excitations along the brim does not cost any energy, as the potential is flat, unlike excitations in the radial direction.

The linear sigma model gives an intuition for the Goldstone mode. In the case of $N = 2$, the symmetry of the Lagrangian are rotations in the plane. As the ground state is a point along the "brim" of the hat, this rotational symmetry is broken. However, any excitations in the angular direction do not cost any energy,

[15] The particles are bosons due to the bosonic nature of the transformations, g. If the generators are Grassmann numbers, the resulting particles, called goldstinos, are fermions.

which is indicative of a massless mode. This is illustrated in Figure 7. In this example, the original symmetry group is one-dimensional, so there are no unbroken symmetries. Consider instead the $N = 3$ linear sigma model, which has the three-dimensional symmetry group SO(3), rotations of the sphere. We see that the ground state is left invariant under a subgroup of the original symmetry transformations. The ground state manifold of this system, the set of all states that minimizes the effective potential, is then a sphere. When the system chooses one single ground state, this symmetry is broken, but only for two of the generators. The generator for rotations around the ground state leaves that point unchanged and is thus an unbroken symmetry. Any excitations in the direction of the broken symmetries do not cost energy, as it is in the ground state manifold. On the other hand, the unbroken symmetry does not correspond to an excitation. This is illustrated in Figure 8.

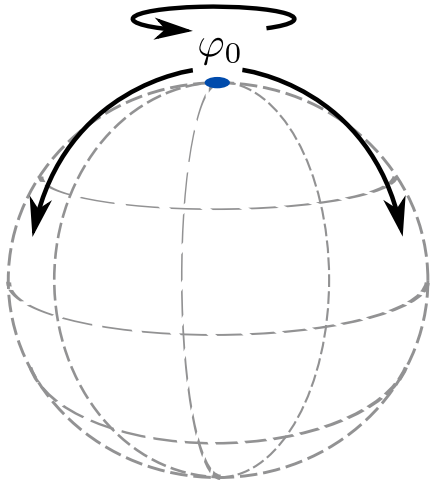


Figure 8: Excitations for the $N = 3$ sigma model. Two of the symmetries are broken, while rotations around the ground state leaves the system unchanged.

5.6 Schwinger-Dyson equations and Ward identities

Given a system of fields φ_a governed by some action $S[\varphi]$, the expectation value of a functional of the fields, $F[\varphi]$, is given by

$$\langle F[\varphi] \rangle = \int \mathcal{D}\varphi \, e^{iS[\varphi]} F[\varphi]. \tag{220}$$

If we perform a *local* transformation of the field on the form $\varphi(x) \to \varphi(x) + \epsilon\eta(x)$, the integral measure will remain unchanged. The expectation value, to first order in ϵ, then changes as

$$\langle F \rangle \to \int \mathcal{D}\varphi \, e^{i(S+\epsilon\delta S)}(F + \epsilon\delta F) = \langle F \rangle + \epsilon \langle i(\delta_\eta S)F \rangle + \epsilon \langle \delta_\eta F \rangle . \tag{221}$$

The variation δ_η is related to the functional dervative through $\delta_\eta S = \int \mathrm{d}^n x \, \frac{\delta S[\varphi]}{\delta\varphi(x)}\eta(x)$, as defined in subsection 3.2. As this amounts to a change of integration variable, the expectation value should remain unchanged. This gives us the important identity

$$\langle i(\delta_\eta S[\varphi])F[\varphi] \rangle + \langle \delta_\eta F[\varphi] \rangle = 0. \tag{222}$$

Inserting the integral form of the variation, and using the fact that η is independent of φ, we may write this identity as

$$\left\langle \frac{\delta S[\varphi]}{\delta\varphi(x)} F[\varphi] \right\rangle = i \left\langle \frac{\delta F[\varphi]}{\delta\varphi(x)} \right\rangle . \tag{223}$$

The Schwinger-Dyson equations are important special cases of this identity. They are the equations of motion of correlation functions. They thus incorporate the dynamics of a theory. We derive them by setting $F[\varphi] = \varphi(x_1)...\varphi(x_n)$. If we have a Lagrangian on the form $\mathcal{L} = -\frac{1}{2}\varphi(\partial^2 + m^2)\varphi - V[\varphi]$, then Equation 223 becomes

$$\begin{aligned}(\partial_x^2 + m^2) \langle \varphi(x)\varphi(x_1)\dots\varphi(x_n) \rangle = &- \langle \mathcal{V}'[\varphi](x)\varphi(x_1)\dots\varphi(x_n) \rangle \\ &- i\sum_i \delta(x - x_i) \langle \varphi(x_1)\dots\varphi(x_{i-1})\varphi(x_{i+1})\dots\varphi(x_n) \rangle .\end{aligned}$$

If $n = 1$ and $\mathcal{V} = 0$, we get the defining relation for the free Greens function,

$$(\partial_x^2 + m^2) \langle \varphi(x)\varphi(y) \rangle = -i\delta(x - y). \tag{224}$$

We may also consider slightly more general transformations of $\varphi(x)$, such as local phase-transformations $\varphi(x) \to e^{i\epsilon(x)}\varphi(x)$, as long as they do not affect the measure of the path integral. We will use this to derive identities related to the Schwinger-Dyson equations that incorporate the symmetries of a given theory. If $\varphi(x) \to \varphi(x) + \delta\varphi(x)$ is a global symmetry transformation, so that $\delta\mathcal{L} = \partial_\mu K^\mu$ and the integration measure is unchanged, then $\varphi(x) \to \varphi(x)+\eta(x)\delta\varphi(x)$ is a corresponding local transformation. We recover the global transformation for $\eta = 1$. The variation of the action from this transformation will be

$$\begin{aligned}
\delta S &= \int \mathrm{d}^4x \left(\frac{\partial \mathcal{L}}{\partial \varphi} \eta \delta\varphi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \partial_\mu(\eta\delta\varphi) \right) \\
&= \int \mathrm{d}^4x \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \delta\varphi \right) \partial_\mu \eta + \int \mathrm{d}^4x\, \eta(x) \left(\frac{\partial \mathcal{L}}{\partial \varphi} \delta\varphi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \partial_\mu \delta\varphi \right) \\
&= -\int \mathrm{d}^4x\, \eta(x) \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \delta\varphi - K^\mu \right).
\end{aligned}$$

In the last line, we integrated by parts, and used $\delta\mathcal{L} = \partial_\mu K^\mu$. From subsection 3.8, we recognize the term within the parenthesis as precisely the Nöther current J^μ, so

$$\delta S = -\int \mathrm{d}^4x\, \eta(x) \partial_\mu J^\mu. \tag{225}$$

As φ is an integration variable in the path integral, it is not necessarily on-shell. We can therefore not use Nöther's theorem, $\partial_\mu J^\mu = 0$, as this relies on the equation of motion. However, for $F = 1$ and thus $\delta F = 0$, we can insert Equation 225 into Equation 220 to obtain the quantum version of the current conservation equation of Nöther's theorem, Equation 133,

$$\partial_\mu \langle J^\mu \rangle = 0. \tag{226}$$

With $F = \varphi(x_1)\varphi(x_2)$, we get [**schwartzQuantumFieldTheory2013**, **peskinIntroductionQuantumField1995**]

$$\partial_{x,\mu} \langle J^\mu(x)\varphi(x_1)\varphi(x_2) \rangle = -i\delta(x - x_1) \langle \delta\varphi(x_1)\varphi(x_2) \rangle - i\delta(x - x_2) \langle \varphi(x_1)\delta\varphi(x_2) \rangle. \tag{227}$$

Identities of this form are called Ward-Takashi identities, often just Ward-identities, and encode the symmetries of a theory. In case symmetry is only approximate, so $\delta\mathcal{L} = \partial_\mu K^\mu + \Delta$, where Δ is some small symmetry breaking operator, or it is subject to an anomaly, so $\mathcal{D}\varphi \to \mathcal{D}\varphi(1 + \Delta)$, then the conservation equation is modified to

$$\partial_\mu \langle J^\mu \rangle = \langle \Delta \rangle. \tag{228}$$

To create the generating functional, we must add external currents j. However, these new terms in the Lagrangian might break the invariance under a symmetry transformation $\varphi \to \varphi'$. If we transform the external currents as $j \to j'$ to counteract the transformation of the fields, then the theory should remain invariant. As before, we make both these transformations local, making sure that they leave the measure invariant. We can then perform a change of variable in the path integral, which relates generating functionals with different external currents, $Z[j] = Z[j']$. This relation must not only be obeyed by the underlying theory but also by any effective theory, which significantly constrains the form of the effective Lagrangian. As an illustration, we consider an example of spinor fields adapted from [**schererIntroductionChiralPerturbation2002**], as this is closely related to the construction of chiral perturbation theory. Spinors and gauge theory, which are relevant for this example, are discussed in more depth in **??**. Consider a massless spinor field with the Lagrangian,

$$\mathcal{L} = i\bar{\psi}\partial\!\!\!/\psi - \mathcal{V}[\psi, \bar{\psi}]. \tag{229}$$

Assume this theory has a global SU (N) symmetry, so the $\mathcal{V}$ remains unchanged under the transformation $\psi \to U\psi$, $\bar{\psi} \to \bar{\psi}U^\dagger$. The system then has a corresponding conserved current,

$$J^\mu_\alpha = \bar{\psi} T_\alpha \gamma^\mu \psi, \tag{230}$$

where T_α are the generators of SU (N). We then include spinor sources $\eta = \eta_\alpha T_\alpha$ and vector sources $v_\mu = v^\alpha_\mu T_\alpha$ by adding the terms $\bar{\eta}\psi$, $\bar{\psi}\eta$, and $v^\alpha_\mu J^\mu_\alpha$ to the Lagrangian. Under a local SU (N) transformation, $\psi \to e^{i\theta_\alpha(x)T_\alpha}\psi$, the action changes as

$$S \to \int \mathrm{d}^4x \left[i\bar{\psi}\partial\!\!\!/\psi - \mathcal{V}[\psi, \bar{\psi}] + \bar{\eta}U\psi + \bar{\psi}U^\dagger\eta + \bar{\psi}\gamma^\mu(U^\dagger v_\mu U + iU^\dagger\partial_\mu U)\psi \right]. \tag{231}$$

The last term corresponds to the change in action without sources, which we found earlier Equation 225. We then define transformations of the external fields to counteract the transformation of ψ. As these transformations are local, the sources now act as gauge fields. The gauge transformation of the external sources are

$$\eta \to U\eta, \quad \bar{\eta} \to \bar{\eta}U^\dagger, \quad v_\mu \to U(v_\mu + i\partial_\mu)U^\dagger. \tag{232}$$

This gives the relation $S[\psi', \bar{\psi}', \eta', \bar{\eta}', v'] = S[\psi, \bar{\psi}, \eta, \bar{\eta}, v]$, where the prime indicates gauge transformed fields. As we argued in the subsection on the Dyson-Schwinger equations, we can change the integration variables inside the path integral without changing the result. Considering an infinitesimal transformation, and expanding to first order in θ, we get

$$\begin{aligned} 0 &= Z[\eta', \bar{\eta}', v'] - Z[\eta, \bar{\eta}, v] \\ &= i \int \mathrm{d}^4x \left\langle i\theta_\alpha(x)\bar{\psi}T_\alpha\eta - i\theta_\alpha(x)\bar{\eta}T_\alpha\psi + i\bar{\psi}\gamma^\mu(i\theta_\alpha(x)[T_\alpha, v_\mu] - i\partial_\mu\theta_\alpha(x)T_\alpha)\psi \right\rangle \end{aligned} \tag{233}$$

As the transformation, and thus θ_α, is arbitrary, the integrand must vanish. After partial integration, we are left with

$$\left\langle \bar{\psi}T_\alpha\eta - \bar{\eta}T_\alpha\psi + D_\mu^{\alpha\beta}J_\beta^\mu \right\rangle = 0. \tag{234}$$

Here, $D_\mu^{\alpha\beta}$ is the covariant derivative in the adjoint representation, $D_\mu^{\alpha\beta}J_\nu^\beta = (\delta_{\alpha\beta}\partial_\mu + iv_\mu^\gamma f^{\alpha\gamma\beta})J_\beta^\nu$, and $f^{\alpha\beta\gamma}$ are the structure constants of $\mathfrak{su}(N)$. We can get a more general expression by writing this using the generating functional W,

$$\left(\frac{\delta}{\delta\eta_\alpha(x)}\eta - \bar{\eta}\frac{\delta}{\delta\bar{\eta}_\alpha(x)} + D_\mu^{\alpha\beta}\frac{\delta}{\delta v_\mu^\beta(x)} \right) W[\eta, \bar{\eta}, v] = 0. \tag{235}$$

If we evaluate this at $\eta = \bar{\eta} = v_\mu = 0$, we get the quantum conservation equation $\partial_\mu \langle J_\alpha^\mu \rangle = 0$. From Equation 235, we can also get more general Ward identities by taking functional derivatives with respect to the external sources.

5.7 CCWZ construction

This section is based on [**callanStructurePhenomenologicalLagrangians1969**, **colemanStructurePhenomenologi**
morrisonColemanCallanWessZuminoConstruction2017, **panicoCompositeNambuGoldstoneHiggs2016**, **weinbergQuantumTheoryFields1996**, **pichEffectiveFieldTheory2020**].

As Goldstone bosons are massless, they play a crucial role in low-energy dynamics. To best describe this limit, we seek a parametrization of the theory in which they are the degrees of freedom. This can be done using the CCWZ construction, named after Callan, Coleman, Wess, and Zumino.

5.8 Parametrizing the Goldstone manifold

We saw that the Goldstone bosons correspond to excitations within the vacuum manifold. The vacuum manifold corresponds to points in field space φ that can be reached from the vacuum φ_0 with a transformation $g \in G$. Assume this group acts linearly on the fields. This means that we can write such excitations as

$$\varphi_i = (\tilde{\Sigma}\varphi_0)_i = \tilde{\Sigma}_{ij}(\varphi_0)_j, \quad \tilde{\Sigma} = \tilde{\Sigma}(\eta) = \exp i\eta_\alpha T_\alpha. \tag{236}$$

We will drop the indices (i, j) for the sake of compact notation. $\tilde{\Sigma}$ is thus a function from the parameter space, $\eta_\alpha \in \mathbb{R}^n$, to G,

$$\tilde{\Sigma} : \mathbb{R}^n \longmapsto G. \tag{237}$$

We then get field configurations by making the parameters dependent on space-time. We will for now assume η_α is constant. This parametrization is highly redundant. Two elements $\tilde{\Sigma}$ and $\tilde{\Sigma}'$, related by

$$\tilde{\Sigma}' = \tilde{\Sigma}e^{i\theta_a t_a} \tag{238}$$

results in the same φ. This is because $e^{i\theta_a t_a} = h \in H$, and $h\varphi_0 = \varphi_0$, by assumption. The set of all equivalent $\tilde{\Sigma}$'s is exactly the left coset, $gH = \{\, gh \mid h \in H \,\}$. The set of cosets forms a new manifold, G/H, called the Goldstone manifold. This is a manifold of dimension $\dim(G/H) = \dim(G) - \dim(H)$,

which is the number of broken generators and thus also the number of Goldstone modes. Membership of a certain coset is an equivalence relation, $g \sim g'$ if $g' = gh$. This means that the cosets gH form a partition of G and that each element $g \in G$ belongs to one, and only one, coset. To remove the redundancy in the parametrization, we need to choose one representative element from each coset.

By the inverse function theorem, any mapping between manifolds $f : \mathcal{M} \mapsto \mathcal{N}$ that has a non-degenerate differential, that is an invertible Jacobian, at a point $p \in \mathcal{M}$, is invertible in a neighborhood of p. If we write

$$\tilde{\Sigma}(\xi, \theta) = \exp i\xi_i x_i \exp i\theta_a t_a, \tag{239}$$

then the map is invertible at $p = (\xi_i = 0, \theta_a = 0)$, as the Jacobian is the identity matrix. This point is mapped to the identity element of G. This means that, in a neighborhood $U \subset G$ of the identity, each element g has a unique representation $g = \tilde{\Sigma}$ [**leeIntroductionSmoothManifolds2003d**]. Furthermore, two elements $\tilde{\Sigma}'$ and $\tilde{\Sigma}$ related by $\tilde{\Sigma}' = \tilde{\Sigma}h$, $h \in H$ have the same ξ-arguments. We see that ξ_i parametrize G/H, in the neighborhood of the identity. We therefore demand that $\tilde{\Sigma}$ always appear in the standard form

$$\Sigma(\xi) = \tilde{\Sigma}(\xi, 0) = \exp i\xi_i x_i. \tag{240}$$

The field $\varphi(x)$ can therefore be written as

$$\varphi(x) = \Sigma(x)\varphi_0 = \exp i\xi_i(x) x_i \varphi_0, \tag{241}$$

and $\xi_i(x)$ can be associated with the Goldstone bosons.

In the linear sigma model, the original $\mathrm{O}(N)$ symmetry is broken down to $\mathrm{O}(N-1)$, which transforms the remaining $N-1$ fields with vanishing expectation values into each other. However, $\mathrm{O}(N)$ consists of two disconnected subsets, those matrices with determinant 1 and those with determinant -1. There is no continuous path that takes an element of $\mathrm{O}(N)$ with determinant -1 to an element with determinant 1.[16] The set of symmetries that are connected to the identity is

$$G = SO(N) = \{\, M \in \mathrm{O}(N) \mid \det M = 1 \,\}. \tag{242}$$

If we choose $\varphi_0 = (0, 0, ..., v)$, then it is apparent that the ground state is invariant under the rotations of the $N-1$ first fields, so the unbroken symmetry is $H = \mathrm{SO}(N-1)$. The Goldstone manifold is $G/H = \mathrm{SO}(N)/\mathrm{SO}(N-1)$.

Consider the case of $N = 3$, which is illustrated in Figure 8. G is the rotations of the sphere, while H is rotations around φ_0, SO(2). The Goldstone manifold consists of the rotations of φ_0 to other points of the sphere, i.e. $G/H = \mathrm{SO}(3)/\mathrm{SO}(2) = S^2$, the 2-sphere. This is not a Lie group, as translating φ in a closed path around the sphere may result in a rotation around the z-axis. This is illustrated in Figure 9

To check that ξ_i are the Goldstone modes, we study the way they appear in the Lagrangian. As they are massless, no mass term of the form $M_{ij}\xi_i\xi_j$ should appear. The original Lagrangian $\mathcal{L}[\varphi]$ was invariant under global transformations $\varphi(x) \mapsto g\varphi(x)$. However, any terms that only depend on $\varphi(x)$, and not its derivatives, will also be invariant under a *local* transformation, $\varphi(x) \mapsto g(x)\varphi(x)$. Our parametrization of the fields, $\varphi(x) = \Sigma(x)\varphi_0$ is exactly such a transformation, which means that any such terms are independent of the Goldstone bosons. We can therefore write

$$\mathcal{L}[\varphi] = \mathcal{L}_{\mathrm{kin}}[\varphi] + \mathcal{V}(\varphi_0), \tag{243}$$

where all terms in $\mathcal{L}_{\mathrm{kin}}$ are proportional to at least one derivative term, $\partial_\mu\varphi(x)$, while $\mathcal{V}$ is independent of $\Sigma(x)$. Inserting the parametrization into this derivative term, we get

$$\partial_\mu\varphi(x) = \partial_\mu[\Sigma(x)\varphi_0] = \Sigma(x)[\Sigma(x)^{-1}\partial_\mu\Sigma(x)]\varphi_0. \tag{244}$$

The Lagrangian will therefore depend on ξ_i via terms of the form $\Sigma(x)^{-1}\partial_\mu\Sigma(x)$, which is called the Mauer-Cartan form. This is a $\mathfrak{g}$-valued function, which means that it can be written as

$$i\Sigma(x)^{-1}\partial_\mu\Sigma(x) = d_\mu(x) + e_\mu(x), \tag{245}$$

$$d_\mu = ix_i d_{ij}(\xi)\partial_\mu\xi_j, \tag{246}$$

$$e_\mu = it_a e_{ai}(\xi)\partial_\mu\xi_i, \tag{247}$$

where d_{ij} and e_{ai} are as-of-yet unknown real valued functions of ξ [**watanabeEffectiveLagrangianNonrelativistic2014**, **weinbergQuantumTheoryFields1996**].

[16] A simple proof of this is the fact that the determinant is a continuous function, while any path $\det M(t)$ such that $\det M(1) = -1$, $\det M(0) = 1$ must make a discontinuous jump.

5.9 Transformation properties of Goldstone bosons

We can deduce how the Goldstone bosons transform under G from how φ transforms. In general,

$$\varphi' = g\varphi = (g\Sigma(\xi))\varphi_0 = \Sigma(\xi')\varphi_0 \quad g \in G. \tag{248}$$

While $\Sigma(\xi')$ has the standard form by assumption,

$$\Sigma(\xi') = \exp i\xi'_i x_i, \tag{249}$$

$g\Sigma(\xi)$ does not, in general.

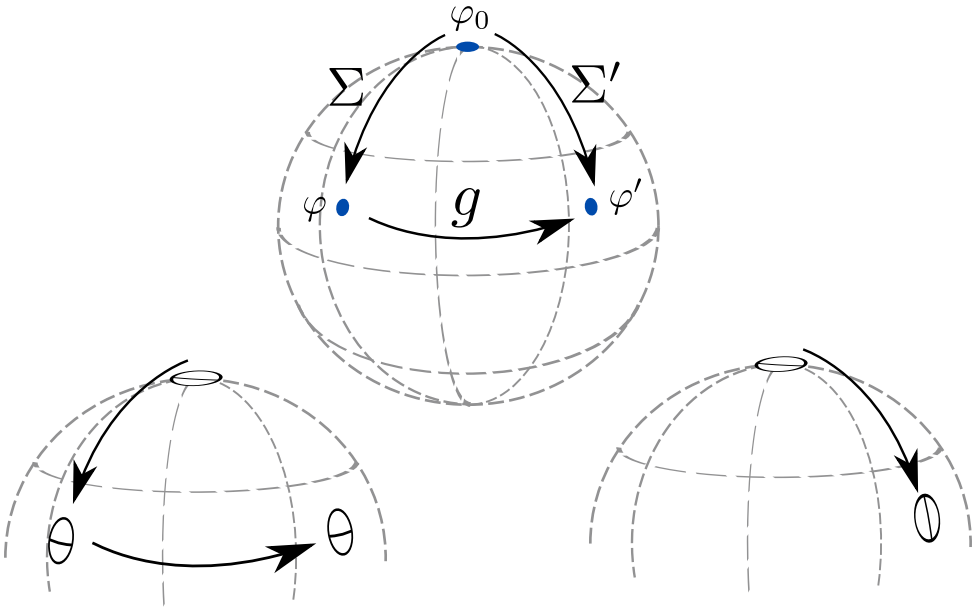


Figure 9: The top figure illustrates the transformation of φ_0 to φ and then φ', and the alternative, direct transformation $\varphi_0 \to \varphi'$. The bottom figure illustrates how this can rotate a neighborhood of φ_0 differently.

Figure 9 illustrates this in the case of $G = \mathrm{SO}(3)$. $\Sigma(\xi)$ transforms φ_0 to φ, then g transforms φ to $\varphi' = \Sigma(\xi')\varphi_0$. Assuming φ and φ' are close enough to φ_0, we can write $\Sigma(\xi)$ and $\Sigma(\xi')$ on the standard form. However, if we follow a small neighborhood around φ_0 as it is acted on by $\Sigma(\xi)$, then g, it will be rotated by the time it arrives at φ' when compared to the same neighborhood if it was acted on by $\Sigma(\xi')$.

$g\Sigma(\xi)$ and $\Sigma(\xi')$ are in the same coset, as they by assumption corresponds to the same physical state. This means that we can write $g\Sigma(\xi) = \Sigma(\xi')h(g,\xi)$, where $h(g,\xi) \in H$. The transformation rule of ξ under G is therefore implicitly defined by

$$\Sigma(\xi') = g\Sigma(\xi)[h(g,\xi)]^{-1}. \tag{250}$$

This is, in general, not a linear representation, which is why this construction also is called a *non-linear realization*. Using the transformation rule, we can obtain the transformation rule of the Maurer-Cartan form. We use the shorthand $\Sigma = \Sigma(\xi)$, $\Sigma' = \Sigma(\xi')$, and $h = h(g,\xi)$. This gives

$$\begin{aligned}
\Sigma^{-1}\partial_\mu\Sigma \to{}& \Sigma'^{-1}\partial_\mu\Sigma' \\
&= (g\Sigma h^{-1})^{-1}\partial_\mu(g\Sigma h^{-1}) \\
&= (h\Sigma^{-1}g^{-1})g[(\partial_\mu\Sigma)h^{-1} + \Sigma\partial_\mu h^{-1}] \\
&= h\Sigma^{-1}(\partial_\mu\Sigma)h^{-1} + h\partial_\mu h^{-1} \\
&= h(\Sigma^{-1}\partial_\mu\Sigma + \partial_\mu)h^{-1}.
\end{aligned}$$

In terms of d_μ and e_μ,

$$d_\mu \to h d_\mu h^{-1}, \tag{251}$$

$$e_\mu \to h(e_\mu + i\partial_\mu)h^{-1}. \tag{252}$$

The transformation rule of e_μ is that of a gauge field, with the gauge group H. We will discuss gauge fields in more detail in **??**. If we include massive degrees of freedom and not only the Goldstone modes, e_μ is used to create a covariant derivative of the massive modes. We are only interested in the Goldstone modes and will therefore be satisfied with d_μ. We have now greatly constrained the way the Goldstone modes may appear in the Lagrangian. However, this does not yet solve the problem of the strong force being non-perturbative. To do this, we need to create an effective field theory in which the strong force has been integrated out.

6 Quantum Mechanics

Finding solutions to problems in quantum mechanics ultimatley comes down to solving the Schrødinger equation

$$\hat{H}\left|\Psi, t\right\rangle = i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\left|\Psi, t\right\rangle .$$

The most straight forward way to do this is to project it down to a basis, like position q or momentum p. The Schrødinger equation then might become a partial differential equation, which as we all know can be nasty things. However, there is an more abstract algebraic way of working with the equation in a basis independent form, which relies on a few basic assumption about the operators. The main only assumption going in is the commutation relation between the operators, like

$$[q, p] = i\hbar,\ [L_i, L_j] = i\hbar\epsilon_{ijk}L_k, \quad \text{where } [A, B] = AB - BA$$

These operators might be represented as variables, matrices or differential operators, but here, this is the *only* information about them we use. This means that as soon as you find something that behaves like this, everything derived here comes for free. We can then extract a lot of information about the system, without having to solve differential equations, or do a single integral. To remove clutter, this text sets $\hbar = c = 1$.

6.1 Harmonic oscillator

The hamiltonian for a harmonic oscillator is

$$\hat{H} = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 q^2 = \frac{1}{2}\omega\left(m\omega q^2 + \frac{p^2}{m\omega}\right)$$

This leads to very natural versions of the position and momentum operators, and their fundamental commutation relation

$$q' := \sqrt{m\omega}\, q, \quad p' := \frac{p}{\sqrt{\omega m}}, \implies [q', p'] = i \tag{253}$$

Creation operators

We can then take this new hamiltonian and factor it, utilizing the commutation-relation,

$$\hat{H}' = \frac{1}{2}\omega(q'^2 + p'^2) = \frac{1}{2}\omega(q' - ip')(q' + ip') - \frac{i}{2}\omega[q', p']$$

This leads us to the introduction of the anhelation and creation operators, resp.

$$a := \frac{1}{\sqrt{2}}(q' + ip'),\ a^\dagger := \frac{1}{\sqrt{2}}(q' - ip') \implies H' = aa^\dagger + \frac{1}{2}. \tag{254}$$

The commutation relation is

$$\begin{aligned}[a, a^\dagger] = \frac{1}{2}(q' + ip')(q' - ip') - \frac{1}{2}(q' - ip')(q' + ip') = -\frac{i}{2}[q', p'] - \frac{i}{2}[q', p'] \\ \implies [a, a^\dagger] = 1\end{aligned} \tag{255}$$

and we can find the reverse relation

$$q' = \frac{1}{\sqrt{2}}(a + a^\dagger), \quad p' = \frac{-i}{\sqrt{2}}(a - a^\dagger).$$

The last tool we need is the number operator, $\hat{N} := a^\dagger a$. This means $\hat{H} = \hat{N} + 1/2$ and these operators share commutation relations,

$$[\hat{H}', a] = [\hat{N}, a] = a^\dagger aa - aa^\dagger a = [a^\dagger, a]a = -a, \tag{256}$$
$$[\hat{H}', a^\dagger] = [\hat{N}, a^\dagger] = a^\dagger aa^\dagger - a^\dagger a^\dagger a = a^\dagger [a, a^\dagger] = a^\dagger.$$

Eigenvalues

Assume $|\psi\rangle$ is an eigenvector of $\hat{H}'$, so $\hat{H}' |\psi\rangle = \lambda |\psi\rangle$. We can exploit the commutators of the hamiltonian and the a operators to create new eigenvectors.

$$\hat{H}'(a |\psi\rangle) = a(\hat{H}' + [\hat{H}', a]) |\psi\rangle = (\lambda - 1)(a |\psi\rangle)$$
$$\hat{H}'(a^\dagger |\psi\rangle) = a^\dagger(\hat{H}' + [\hat{H}', a^\dagger]) |\psi\rangle = (\lambda + 1)(a^\dagger |\psi\rangle)$$

This seem to create an infinite chain of eigenvalues and vectors in both directions. However, there must be a bottom, as

$$0 \leq \| a |\psi\rangle \|^2 = \langle \psi | a^\dagger a |\psi\rangle = \lambda - 1/2 \implies \lambda \geq 1/2.$$

This holds true for all eigenvalues of $\hat{H}'$. For both of these statements to be true, there must be a vector $|0\rangle$ such that $a |0\rangle = 0$, which can be reach from *any* eigenvector of $\hat{H}'$ through repeated application of a, and thus all eigenvectors can be reached from application of $a^\dagger$. [17] The eigenvalue for the ground state, E_0', is

$$\hat{H}' |0\rangle = (a^\dagger a + \frac{1}{2}) |0\rangle = \frac{1}{2} |0\rangle \implies E_0' = \frac{1}{2}.$$

The next eigenvalues up the chain is then given by:

$$\hat{H}'[(a^\dagger)^n |0\rangle] = (E_0 + n)[(a^\dagger)^n |0\rangle] := E_n[(a^\dagger)^n |0\rangle], \quad E_n = n + \frac{1}{2}. \tag{257}$$

Eigenvectors

If we apply the $\hat{N} = \hat{H}' - 1/2$ operator to one of these new eigenvectors, we get the number of times $a^\dagger$ has been applied to get to it,

$$\hat{N}(a^\dagger)^n |0\rangle = n(a^\dagger)^n |0\rangle$$

We assume $|0\rangle$ is normalized, and let $|n\rangle$ be the n'th eigenvector in the chain. Therefore $a^\dagger |n\rangle = c_n |n+1\rangle$, $a |n\rangle = c_n' |n-1\rangle$ for some constants c_n, c_n', and

$$\langle n | aa^\dagger |n\rangle = \langle n | \hat{N} + 1 |n\rangle = \langle n+1 | c_n^* c_n |n+1\rangle = |c_n|^2 = n+1, \quad |c_n'|^2 = n$$

Choosing real and positive values for the constants, we get

$$a^\dagger |n\rangle = \sqrt{n+1} |n\rangle , \, a |n\rangle = \sqrt{n} |n-1\rangle , \quad |n\rangle = \frac{1}{\sqrt{n!}} |0\rangle . \tag{258}$$

Expectationvalues

The expectation value of any odd power of q', p' will be zero, as it can be written as a sum of odd combinations $a, a^\dagger$, and we will then end up with a inner product of the form $\langle m | n\rangle , m \neq n$. Thus,

$$\langle q^{2n+1} \rangle = \langle p^{2n+1} \rangle = 0. \tag{259}$$

Even numbers, however, can give nonzero answers. For example

$$\langle q'^2 \rangle = \frac{1}{\sqrt{2}} \langle n | (a + a^\dagger)^2 |n\rangle = \frac{1}{\sqrt{2}} \langle n | a^\dagger a + aa^\dagger |n\rangle = \frac{\sqrt{n} + \sqrt{n-1}}{\sqrt{2}}, \quad \langle p'^2 \rangle = \frac{-\sqrt{n} + \sqrt{n-1}}{\sqrt{2}} \tag{260}$$

[17] This is at least true for eigenvalues, but can there be parallel chains of eigenvectors?

Coherent states

Assume $|\alpha\rangle$ is the eigenvector of a, i.e. $a|\alpha\rangle = \alpha|\alpha\rangle$. Then we can expand it in terms of the energy eigenvectors,

$$|\alpha\rangle = \sum_n c_n |n\rangle .$$

Using what we have found for the lowering operator,

$$a|\alpha\rangle = a\sum_n c_n|n\rangle = \sum_n c_n\sqrt{n}|n-1\rangle = \alpha\sum_n c_n|n\rangle$$

This means we can get a recursion relation for c_n, $c_{n+1} = \alpha c_n/\sqrt{n+1}$, which explicitly evaluates to $c_n = \alpha^n c_0/\sqrt{n!}$. If we enforce $\langle\alpha\rangle = 1$, then we get

$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_n \frac{\alpha^n}{\sqrt{n!}}|n\rangle . \tag{261}$$

In the Heisenberg picture, the time dependence of the operators is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}a(t) = i[\hat{H}, a(t)] = -ia(t) \implies a(t) = ae^{-it}, \quad a^\dagger(t) = a^\dagger e^{it}$$

The time dependent expectation value of the coherent state is thus

$$\langle\alpha|\, a(t)\,|\alpha\rangle = \alpha e^{-it}, \quad \langle\alpha|\, a^\dagger(t)\,|\alpha\rangle = \alpha^* e^{it},$$

and the position/momentum expectationvalues are

$$\langle q(t)\rangle_\alpha = \sqrt{2}|\alpha|\cos(t - \arg(\alpha)), \quad \langle p(t)\rangle_\alpha = -\sqrt{2}|\alpha|\sin(t-\arg(\alpha))$$

Position representation

The rule for finding the position-representation of a operator $\hat{F}(\hat{x},\hat{p})$ is

$$\langle x|\,\hat{F}(\hat{x},\hat{p})\,|\psi\rangle = \hat{F}\left(x, -i\frac{\partial}{\partial x}\right)\langle x|\psi\rangle ,$$

where $\langle x|\psi\rangle := \psi(x)$ is the wave function of the system. This means we can get a differential equation for the wavefunction, if we have an operator equation, which we do: $a|0\rangle = 0$. The ground state wave function is therefore given by

$$a|0\rangle = 0 \implies \left(x + \frac{\partial}{\partial x}\right)\psi_0 = 0 \implies \psi_0(x) = C\exp(-x^2/2).$$

The exited states are given by

$$\psi_n(x) = \left(x - \frac{\partial}{\partial x}\right)^n \psi_0(n).$$

6.2 Field of harmonic oscillators

Consider the quantized photon field, or EM vectorpotential, in a square box with side lengths L, volume $V = L^3$ and periodic boundary conditions. The field can be split up into modes with different wave vectors k with polarization vectors $\epsilon_{k\lambda}$. They are given by

$$k = \frac{2\pi n_i}{L}\hat{e}_i, \quad \epsilon_{k\lambda}\cdot k = 0, \quad i \in \{1,2,3\},\ \lambda \in \{1,2\}.$$

Let

$$S = \{(k_1\lambda_1), (k_2\lambda_2), ...\}$$

be the set of all unique modes. The Hamiltonian for a free EM field is then

$$H = \sum_{j\in S}\omega_j\left(a_j^\dagger a_j + \frac{1}{2}\right),$$

where $\omega_{k\lambda} = k$ ($c = 1$, remember) is the frequency of the mode $j = (k\lambda)$. The sum is over the wave vectors k and polarizations λ corresponding to a polarization vector $\epsilon_{k\lambda}$. The a's here are operators, just as in the case of the single oscillator case, however it is one for each mode/polarization. In reality, there is a lot of suppressed tensor products here. Writing the operators out, they are

$$a_{k\lambda} = 1_{k_1\lambda_1} \otimes 1_{k_2\lambda_2} \otimes ... \otimes a_{k,\lambda} \otimes ... \tag{262}$$

These operators obey the commutation relation

$$[a_i, a_j^\dagger] = \delta_{ij},$$

which means we get everything from the single harmonic oscillator for free. Even though we started with the creation operators, we can still define the canonical operators

$$q_j = \sqrt{\frac{1}{2\omega_j}} \left(a_j + a_j^\dagger\right), \; p_j = \sqrt{\frac{\omega_j}{2}} \left(a_j + a_j^\dagger\right), \quad [q_i, p_j] = i\delta_{ij}$$

Notice that we can't get away with removing the ω's this time, as there are different frequencies for each oscillator. We can again define a number operator $N_j = a_j^\dagger a_j$. Then, the energy eigenvalues are of the form

$$|F\rangle = \bigotimes_{j\in S} |n_j\rangle = |n_{k_1\lambda_1}, n_{k_1\lambda_1}, ...\rangle, \quad H|\psi\rangle = (E_0 + \sum_{j\in S} \omega_j n_j)|\psi\rangle,$$

where $E_0 = \sum_{j\in A} \omega_j/2$ is the (very infinite) ground state energy. This, however, is not a worry, as it is a constant. Let $|0\rangle = |0, 0, ...\rangle$ be the ground state of the system. The different states are then built up by applying the raising operators. If we assume there is a highest exited state N (anything else would require an infinite amount of energy), then the state is characterized by the number of photons in each state, $F = (n_1, n_2, ...n_N)$. The state is then

$$|F\rangle = |n_1, n_2, ...n_N, 0, 0, ...\rangle = \bigotimes_{n_j\in F} \frac{(a_j^\dagger)^{n_j}}{\sqrt{n_j!}} |0\rangle.$$

Position representation

We now have a wave function for each mode. To find the position representation of the system, we need to know the position eigenvector. However, now we have one position operator for each mode, so the different eigenvectors are

$$\hat{q}_j |q_j\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes ... (\hat{q}_j |q_j\rangle) \otimes ... = |\psi_1\rangle \otimes |\psi_2\rangle ... \otimes (q_j |q\rangle) \otimes ... = q_j |q_j\rangle$$

as the operator applies only to he corresponding mode. We can then construct the position operator for *all* modes/photons, $\hat{x} = \bigotimes_{j\in S} q_j$, with the corresponding eigenvectors $|x\rangle$:

$$\hat{x}|x\rangle = \bigotimes_{j\in S} \hat{q}_j |q_j\rangle = \left(\prod_{j\in S} q_j\right) |x\rangle.$$

What does this mean? I have no idea...

Field operators and coherent states

The field strength at a position x in the box is given by the field operator

$$\hat{A}(x) = \sum_{j\in S} \frac{\epsilon_j}{\sqrt{2k_j V}} \left(a_j e^{ik_j\cdot x} + a_j^\dagger e^{-ik_j\cdot x}\right). \tag{263}$$

We can then make coherent states, as in the one oscillator case:

$$|\alpha_j\rangle = e^{-|\alpha_j|^2/2} \sum_{n_j} \frac{\alpha_j}{\sqrt{n_j!}} |0, 0, ...n_j, 0, ...\rangle.$$

As we have time dependence again given by

$$\frac{\mathrm{d}t}{\mathrm{d}a_j}(t) = i[H, a_j(t)] = -i\omega_j a_j(t) \implies a_j(t) = a_j e^{-i\omega_j t},\ a_j^\dagger(t) = a_j^\dagger e^{i\omega_j t},$$

we get

$$A(x,t) := \langle \alpha_j | \hat{A}(x,t) | \alpha_j \rangle = \epsilon_j \frac{\sqrt{2}|\alpha_j|}{\sqrt{k_j V}} \cos(k_j \cdot x - \omega_j t - \arg(\alpha)).$$

The classical electromagnetic linearly polarized plane waves are not photon-states, but rather coherent states, superpositions of different photon states. In fact, as $|F\rangle$ is not an eigenvector of $\hat{A}(x)$, the photon state does not have a definite, well define field-strength anywhere, in the same way a electron with a well defined momentum does not have a well defined position.

6.3 Angular momentum

Angular momentum operators, in units of $\hbar$, are 3-dimensional vectors

$$J = \sum_{i=1}^{3} J_i e_i, \quad J^2 = J_1^2 + J_2^2 + J_3^2$$

defined by the commutation relation

$$[J_i, J_j] = \sum_{k=1}^{3} i\epsilon_{ijk} J_k, \quad i,j,k \in \{1,2,3\} \tag{264}$$

where ϵ_{ijk} is the Levi-Cevita symbol. None of the components commutate with each other, but they all commutate with the length squared, as

$$\begin{aligned}[J^2, J_i] &= [J_j^2, J_i] + [J_k^2, J_i] = J_j[J_j, J_i] + [J_j, J_i]J_j + J_k[J_k, J_i] + [J_k, J_i]J_k \\ &= i(J_j J_k + J_k J_j) - i(J_j J_k + J_k J_j) = 0,\end{aligned}$$

where the indices are cyclic $\ldots k \to i \to j \to k \to i \ldots$. We then make a specific choice, picking out component 3 as 'special', but this is completely arbitrary, and define the raising and lowering operators, analogous to the creation and anhelation operators (254), and find their commutation relation with J_3

$$J_\pm = J_1 \pm iJ_2, \quad [J_3, J_\pm] = \pm J_\pm. \tag{265}$$

Notice the similarities with the commutation relation (256). As the operators J^2, J_z commute, so they have a common set of eigenvectors. [18] Assume $|\lambda, \mu\rangle$ is an eigenvector of J_3 so that $J_3 |\lambda,\mu\rangle = \mu |\lambda,\mu\rangle$, $J^2 |\lambda,\mu\rangle = \lambda(\lambda+1)|\lambda,\mu\rangle$, we can get new eigenvectors by applying $J_\pm$, and get

$$J_3(J_\pm |\lambda,\mu\rangle) = (J_\pm J_3 + [J_3, J_\pm])|\lambda,\mu\rangle = (\mu \pm 1)(J_\pm |\lambda,\mu\rangle), \implies J_\pm |\lambda,\mu\rangle = |\lambda, \mu \pm 1\rangle$$

To find relations for the length square operator, we should express it in terms of these other operators.

$$J_\pm J_\mp = J_1^2 + J_2^2 \pm [J_1, J_2] = J^2 - J_z^2 \pm J_3, \implies J^2 = J_\pm J_\mp + J_z^2 \mp J_z.$$

Assuming $\langle \lambda,\mu\rangle\, \lambda, \mu = 1$, we know

$$\begin{aligned}|J_+ |\lambda,\mu\rangle|^2 &= \langle \lambda,\mu | J_- J_+ |\lambda,\mu\rangle = \langle \lambda,\mu | J^2 - J_z^2 - J_3 |\lambda,\mu\rangle = (\lambda(\lambda+1) - \mu(\mu+1)) > 0 \\ |J_- |\lambda,\mu\rangle|^2 &= \langle \lambda,\mu | J_+ J_- |\lambda,\mu\rangle = \langle \lambda,\mu | J^2 - J_z^2 + J_3 |\lambda,\mu\rangle = (\lambda(\lambda+1) - \mu(\mu-1)) > 0.\end{aligned}$$

Again, we see that the only way to both have these kind of raising, lowering operators, and for the positivity condition to hold true, we must have $|\mu| \leq \lambda$, so there must be two values μ_-, μ_+ s.t. $J_\pm |\lambda, \mu_p m\rangle = 0$. Thus, as μ, take integer steps, λ must be integer, or half integer, and we get the conditions for the eigenvalues

$$J^2 |l,m\rangle = l(l+1),\ J_3 |l,m\rangle = m, \quad 2l \in Z,\ m \in \{-l, -l+1, ..., l\}. \tag{266}$$

[18] This should probably be proved here somewhere...

Combining angular momenta

Given two hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$, we can create a composite system $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, which have vectors of the form $|v\rangle = |v_1\rangle \otimes |v_2\rangle$. We can also define operators on the new space. given operators on the old space A_1, A_2, we get a new operator which acts on vectors in the new space $A = A_1 \otimes A_2$, $A|v\rangle = (A_1|v_1\rangle) \otimes (A_2|v_2\rangle)$. Thus, if we have angular momentum operators acting on the old spaces, J_1, J_2, which obey

$$\begin{aligned}
&[J_{1,i}, J_{1,j}] = i\epsilon_{ijk}J_{1,k}, \quad [J_{2,i}, J_{2,j}] = i\epsilon_{ijk}J_{1,k}, \\
&J_1^2|j_1, m_1\rangle = j_1(j_1+1)|j_1, m_1\rangle\,,\ J_{1,3}|j_1, m_1\rangle = m_1|j_1, m_1\rangle\,, \\
&J_2^2|j_2, m_2\rangle = j_2(j_2+1)|j_2, m_2\rangle\,,\ J_{2,3}|j_2, m_2\rangle = m_2|j_2, m_2\rangle\,,
\end{aligned}$$

we can create the total angular momentum operator

$$J = J_1 \otimes 1 + 1 \otimes J_2 = J_1 + J_2.$$

Notice that the tensor product is suppressed when the operator is tensored with the identity operator. The two original operators commute, as

$$[J_1, J_2] = [J_1 \otimes 1, 1 \otimes J_2] = [J_1, 1] \otimes [1, J_2] = 0.$$

This also implies

$$[J_1^2, J_2^2] = 0.$$

Thus, the new operator is a angular momentum operator

$$\begin{aligned}
&[J_i, J_j] = [J_{i,1} + J_{i,2}, J_{j,1} + J_{j,2}] \\
=&[J_{i,1}, J_{j,1}] + [J_{i,1}, J_{j,2}] + [J_{i,2}, J_{j,1}] + [J_{i,2}, J_{j,2}] = i\epsilon_{ijk}J_{1,k} + i\epsilon_{ijk}J_{2,k} \\
\implies &[J_i, J_j] = i\epsilon_{ijk}J_k. \qquad (267)
\end{aligned}$$

It thus also have creation operators as (265). The square of the new operator can be expressed in terms of the old, as

$$J^2 = J \cdot J = (J_1 \otimes 1 + 1 \otimes J_2) \cdot (J_1 \otimes 1 + 1 \otimes J_2) = J_1^2 \otimes 1 + 1 \otimes J_2^2 + 2(J_1 \otimes 1 \cdot 1 \otimes J_2).$$

To find the last terms, we must remember that the angular momentum operators are vectors. Using that $(A_1 \otimes A_2)(B_1 \otimes B_2) = A_1B_1 \otimes A_2B_2$ (this comes just from applying the operators in succession) we get

$$\begin{aligned}
&J_1 \otimes 1 = (J_{1,i} \otimes 1)e_i, \quad 1 \otimes J_{2,i} = (J_{1,i} \otimes 1)e_i \\
&\implies J_1 \cdot J_2 = (J_1 \otimes 1) \cdot (1 \otimes J_2) = \sum_i (J_{1,i} \otimes 1)(1 \otimes J_{2,i}) = \sum_i (J_{1,i} \otimes J_{2,i}).
\end{aligned}$$

Using (265), we can write

$$\begin{aligned}
J_1 \cdot J_2 &= \left[\frac{1}{2}[(J_{1,+} + J_{1,-})e_1 - i(J_{1,+} - J_{1,-})e_2]J_{1,3} + J_{1,3}e_3\right] \cdot \left[\frac{1}{2}[(J_{2,+} + J_{2,-})e_1 - i(J_{2,+} - J_{2,-})e_2]J_{2,3} + J_{2,3}e_3\right] \\
&= J_{1,3}J_{2,3} + \frac{1}{4}\Big(J_{1,+}J_{2,+} + J_{1,-}J_{2,-} + J_{1,+}J_{2,-} + J_{2,-}J_{2,+} - [(J_{1,+}J_{2,+} + J_{1,-}J_{2,-} - J_{1,+}J_{2,-} - J_{2,-}J_{2,+}]\Big) \\
&= J_{1,3}J_{2,3} + \frac{1}{2}(J_{1,+}J_{2,-} + J_{1,-}J_{2,+})
\end{aligned}$$

This means we can write

$$J^2 = J_1^2 + J_2^2 + 2J_{1,3}J_{2,3} + J_{1,+}J_{2,-} + J_{1,-}J_{2,+}.$$

We have two basis which span $\mathcal{H}$, the eigenvectors of the original operators,

$$|j_1, m_1\rangle \otimes |j_2, m\rangle = |j_1, m_1; j_2, m_2\rangle\,,$$

and the eigenvectors of the new operator

$$J^2|j, m\rangle = j(j+1)|j, m\rangle\,, \quad J_3|j, m\rangle = m|j, m\rangle$$

What is the correspondence? The original basis is eigenvalues of J_3, as

$J_3 |j_1, m_1; j_2, m_2\rangle = (J_{1,3} \otimes 1 + 1 \otimes J_{2,3}) |j_1, m_1\rangle \otimes |j_2, m_2\rangle = (J_{1,3} |j_1, m_1\rangle \otimes |j_2, m_2\rangle + |j_1, m_1\rangle \otimes J_{2,3} |j_2, m_2\rangle)$
$= m_1 |j_1, m_1\rangle \otimes |j_2, m_2\rangle + m_2 |j_1, m_1\rangle \otimes |j_2, m_2\rangle = (m_1 + m_2) |j_1, m_1; j_2, m_2\rangle$

So, $m = m_1 + m_2$. However, this is a very degenerate basis. If we take the max value for second quantum number $m_{1+} = j_1, m_{2+} = j_2$, then the original vector is a new basis vector, as

$$J^2 |j_1, m_{1+}; j_2, m_{2+}\rangle = (J_1^2 + J_2^2 + 2J_{1,3}J_{2,3} + J_{1,+}J_{2,-} + J_{2,-}J_{2,+}) |j_1, m_{1+}; j_2, m_{2+}\rangle .$$

We have

$J_{2,-}J_{2,+} |j_1, m_{1+}; j_2, m_{2+}\rangle = J_{1,+}J_{2,-} |j_1, m_{1+}; j_2, m_{2+}\rangle = 0,$
$(J_1^2 + J_2^2 + 2J_{1,3}J_{2,3}) |j_1, m_{1+}; j_2, m_{2+}\rangle = (j_1(j_1+1) + j_2(j_2+1) + 2m_{1+}m_{2+}) |j_1, m_{1+}; j_2, m_{2+}\rangle ,$
$j_1(j_1+1) + j_2(j_2+1) + 2m_{1+}m_{2+} = j_1(j_1+1) + j_2(j_2+1) + 2j_1j_2 = (j_1+j_2)(j_1+j_2+1),$

so

$$J^2 |j_1, m_{1+}; j_2, m_{2+}\rangle = (j_1 + j_2)(j_1 + j_2 + 1) |j_1, m_{1+}; j_2, m_{2+}\rangle$$

Thus, the max value of j is $j_1 + j_2$.