Coordinate Transformations and Orthogonal Coordinates

In Euclidean space it is common to consider non-Cartesian coordinate systems, often to take advantage of a certain symmetry such as rotational invariance. Examples are polar coordinates and spherical coordinates.

In an orthogonal coordinate system the metric is diagonal:

\[ g_{ij} = \hat{e}_i \cdot \hat{e}_j = h_i^2 \delta_{ij} \] (not summed over \( i \))

\( h_i \) are functions of coordinates.

The basis vectors in this case are orthogonal, but are not necessarily unit vectors.

Example: Consider 2D Euclidean space, described as the \( xy \) plane embedded in 3D Euclidean space.

Points on the plane are described by

\[ \vec{x}(x,y) = x \hat{e}_x + y \hat{e}_y + 0 \hat{e}_z \]

Unit vector in 3D Euclidean space.

Basis vectors in Cartesian coordinates:

\[ \begin{align*}
\hat{e}_x &= \frac{\partial \vec{x}}{\partial x} = \hat{e}_x \\
\hat{e}_y &= \frac{\partial \vec{x}}{\partial y} = \hat{e}_y
\end{align*} \]

\[ g_{ij} = \hat{e}_i \cdot \hat{e}_j = \delta_{ij} \]

\[ ds^2 = dx^2 + dy^2 \]

Now consider polar coordinates:

\[ x = r \cos \theta \quad \leftrightarrow \quad r = \sqrt{x^2 + y^2} \]

\[ y = r \sin \theta \quad \cos \theta = \frac{x}{\sqrt{x^2 + y^2}} \]
Polar coordinates: \( x(r, \theta) = r \cos \theta \hat{e}_x + r \sin \theta \hat{e}_y + 0 \hat{e}_z \)

Basis vectors:
\[
\begin{align*}
\hat{e}_r &= \frac{\partial \hat{x}}{\partial r} = \cos \theta \hat{e}_x + \sin \theta \hat{e}_y + 0 \hat{e}_z \\
\hat{e}_\theta &= \frac{\partial \hat{x}}{\partial \theta} = -r \sin \theta \hat{e}_x + r \cos \theta \hat{e}_y + 0 \hat{e}_z
\end{align*}
\]

\[ g_{ij} = \hat{e}_i \cdot \hat{e}_j = h_i^2 \delta_{ij} \]
where
\[
\begin{align*}
h_r &= \hat{e}_r \cdot \hat{e}_r = 1 \\
h_\theta &= \hat{e}_\theta \cdot \hat{e}_\theta = r^2
\end{align*}
\]

An orthonormal basis of vectors is obtained by dividing by the appropriate factor:
\[
\hat{e}_i = \frac{\hat{e}_i}{\sqrt{g_{ii}}}
\]

In polar coordinates:
\[
\begin{align*}
\hat{e}_r &= \hat{e}_r = \cos \theta \hat{e}_x + \sin \theta \hat{e}_y \\
\hat{e}_\theta &= \frac{\hat{e}_\theta}{r} = -\sin \theta \hat{e}_x + \cos \theta \hat{e}_y
\end{align*}
\]

Under a coordinate transformation, the basis vectors are covariant vectors:
\[
\hat{e}_i' = \frac{\partial x^i}{\partial x'^j} \hat{e}_j = \frac{\partial x^j}{\partial x'^i} \hat{e}_i'
\]

If \( V^i \) is a contravariant vector, then \( V^i \hat{e}_i \) is a scalar, i.e. coordinate-invariant.
The vector $\mathbf{v} = V^x \hat{e}_x + V^y \hat{e}_y$ becomes in $(r, \theta)$-coordinates

$\mathbf{v} = V^r \hat{e}_r + V^\theta \hat{e}_\theta$, where

$V^r = \frac{\partial V^x}{\partial r} + \frac{\partial V^y}{\partial \theta} \sin \theta V^r = \cos \theta V^x + \sin \theta V^y$

$V^\theta = \frac{\partial V^x}{\partial \theta} \sin \theta - \frac{\partial V^y}{\partial \theta} \cos \theta V^\theta = -\frac{1}{r} \sin \theta V^x + \cos \theta V^y$  

(Exercise)

with $\hat{e}_r = \cos \theta \hat{e}_x + \sin \theta \hat{e}_y$ and

$\hat{e}_\theta = -\sin \theta \hat{e}_x + \cos \theta \hat{e}_y$ as calculated earlier.

It is straightforward to check that $V^r \hat{e}_r + V^\theta \hat{e}_\theta = V^x \hat{e}_x + V^y \hat{e}_y$.

In terms of unit basis vectors we would write

$$\begin{pmatrix} \hat{e}_r \\ \hat{e}_\theta \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \hat{e}_x \\ \hat{e}_y \end{pmatrix}$$

$$\begin{pmatrix} V^r \\ V^\theta \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} V^x \\ V^y \end{pmatrix}$$

So that $V^r \hat{e}_r + V^\theta \hat{e}_\theta = V^x \hat{e}_x + V^y \hat{e}_y$

$$= \frac{V^r \hat{e}_r + V^\theta \hat{e}_\theta}{\sqrt{V^r + V^\theta}}$$

Hence, the components of $\mathbf{v}$ in orthonormal polar coordinates are related to the components of $\mathbf{v}$ in $(r, \theta)$-coordinates obtained by a coordinate transformation from $(x, y)$-coordinates by

$$V^r = \sqrt{g_{rr}} V^r, \quad V^\theta = \sqrt{g_{\theta\theta}} V^\theta.$$
Covariant Derivative

Vector fields on a manifold live in the tangent space, which in the case that the manifold is embedded in a higher-dimensional Euclidean space we can think of as the tangent plane at each point in the manifold.

\[ \vec{W}(x) = \Omega^m(x) \vec{e}_m(x) \]

We use Greek indices, but the discussion is essentially the same in space or spacetime.

Ordinary derivatives of \( \vec{W}(x) \) do not generally remain in the tangent space:

\[ \partial_v \vec{W}(x) = \partial_v (\Omega^m \vec{e}_m) = (\partial_v \Omega^m) \vec{e}_m + \Omega^m \partial_v \vec{e}_m \]

or, using \( \partial_v \vec{e}_m = \Gamma^k_{mn} \vec{e}_k + \kappa_{mn} \vec{e}_n \),

\[ \partial_v \vec{W} = (\partial_v \Omega^m) \vec{e}_m + \Omega^m \partial_v \vec{e}_m + \Omega^m \Gamma^k_{mn} \vec{e}_k + \Omega^m \kappa_{mn} \vec{e}_n \]

\[ = (\partial_v \Omega^m + \Gamma^m_{nk} \Omega^k) \vec{e}_m + \Omega^m \kappa_{mn} \vec{e}_n \]

The projection of \( \partial_v \vec{W} \) onto the tangent plane defines the covariant derivative of \( \vec{W}(x) \):

\[ D_v \vec{W} = (\partial_v \Omega^m + \Gamma^m_{nk} \Omega^k) \vec{e}_m = (\partial_v \Omega^m) \vec{e}_m \]

Note that \( D_v \Omega^m \) depends in general on all components of \( \Omega^m \), not just the component \( \partial x^m \).
The covariant derivative $\nabla W$ lives in the tangent space, and $\nabla W^m$ transforms as a tensor. This is another way to introduce the covariant derivative: we ask for a derivative that transforms covariantly under coordinate transformations.

The basic point is that neither $\nabla W^a$ nor $T_{ij}^m$ is a tensor, but the combination $\nabla W^m$ is.
The Affine Connection is not a tensor under general coordinate transformations.

\[ \Gamma^i_{\mu \nu} = \frac{\partial x^i}{\partial x^\mu} \frac{\partial x^j}{\partial x^\nu} \partial_j f, \quad f(x) \text{ is a locally standard coordinate system}. \]

In the coordinate system \( x' \),

\[ \Gamma'^{i'}_{\mu' \nu'} = \frac{\partial x'^{i'}}{\partial x'^{\mu'}} \frac{\partial x'^{j'}}{\partial x'^{\nu'}} \partial'_j f \]

chain rule \( \partial \)

\[ = \frac{\partial x^i}{\partial x'^{\mu'}} \frac{\partial x^j}{\partial x'^{\nu'}} \partial_j f \left( \frac{\partial x'^{i'}}{\partial x^i} \frac{\partial x^j}{\partial x^i} \right) \]

\[ = \frac{\partial x^i}{\partial x'^{\mu'}} \frac{\partial x^j}{\partial x'^{\nu'}} \left[ \frac{\partial x'^{i'}}{\partial x^i} \frac{\partial x^j}{\partial x^i} + \frac{\partial x^i}{\partial x'^{\mu'}} \frac{\partial x^j}{\partial x'^{\nu'}} \right] \]

\[ \frac{\partial x^i}{\partial x'^{\mu'}} \frac{\partial x^j}{\partial x'^{\nu'}} \]

\[ \partial x^i = \frac{\partial x^i}{\partial x'^{\mu'}} \frac{\partial x^j}{\partial x'^{\nu'}} \partial_j f \]

Differentiability of a tensor does not generally yield another tensor.

Under the transformation \( x \rightarrow x' \), \( V'_{\mu'} = \frac{\partial x'^{\mu'}}{\partial x^\mu} V^\mu \) for vector \( V^\mu \).

\[ \frac{\partial V_{\mu'}}{\partial x^\nu} = \frac{\partial x'^{\mu'}}{\partial x^\nu} \frac{\partial x^\rho}{\partial x^\mu} \frac{\partial V^\rho}{\partial x^\nu} + \frac{\partial^2 x'^{\mu'}}{\partial x^\rho \partial x^\mu} \frac{2}{\partial x^\rho} \frac{V^\rho}{\partial x^\mu} \]

\( \text{Tensor-like transformation, non-tensorial}. \)
However, the combination $D_i V^m = V^m_{;i} = \frac{\partial V^m}{\partial x^i} + \Gamma^m_{ik} V^k$, the covariant derivative, is a tensor:

$V^m_{;i} = \frac{\partial x^m}{\partial x^i} \frac{\partial}{\partial x^j} V^j_{;i}$

To show this we rewrite the transformation of $\Gamma^2_{\mu
u}$ in a different way.

Use $\frac{\partial x^i}{\partial x^\mu} \frac{\partial x^\rho}{\partial x^\nu} = \delta^i_{\rho}$

\[
\frac{\partial}{\partial x^\mu} : \frac{\partial x^i}{\partial x^\mu} \frac{2 \partial x^\rho}{\partial x^\mu} + \frac{\partial x^\rho}{\partial x^\mu} \frac{\partial x^i}{\partial x^\rho} \frac{\partial^2 x^\mu}{\partial x^i \partial x^\sigma} = 0
\]

\[
\Rightarrow \Gamma^2_{\mu
u} = \frac{\partial x^i}{\partial x^\mu} \frac{\partial x^\rho}{\partial x^\nu} \frac{\partial}{\partial x^i} \frac{\partial x^\mu}{\partial x^\sigma} \frac{\partial x^\rho}{\partial x^\sigma} - \frac{\partial x^\rho}{\partial x^\mu} \frac{\partial x^i}{\partial x^\nu} \frac{\partial^2 x^\mu}{\partial x^i \partial x^\sigma} \frac{\partial x^\rho}{\partial x^\sigma}
\]

It is now straightforward to show that the non-tensorlike term in the transformation of $\Gamma^2_{\mu
u}$ cancels the non-tensorlike term in the transformation of $D_i V^m$ in the combination $D_i V^m + \Gamma^m_{ik} V^k$. (Exercise)

Similarly, the covariant derivative of a covariant vector
\[
D V_m = V_{m;\nu} = \frac{\partial V_m}{\partial x^\nu} - \Gamma^i_{\nu \omega} V^\omega
\]

Under a coordinate transformation, $D V_m$ transforms as a tensor:

$V'_{m;\nu} = \frac{\partial x^\nu}{\partial x^\mu} \frac{\partial x^\sigma}{\partial x^\nu} V_{m;\sigma}$ (Exercise)
In general, covariant derivatives of tensors involve a sum of terms, each involving one factor of $\Gamma^\mu_{\nu\lambda}$, one for each index on the tensor.

Example: $T^\mu_\lambda \phi = \frac{\partial}{\partial x^\rho} T^\mu_\sigma \phi + \Gamma^\mu_\nu \Gamma^\nu_\lambda \phi + \Gamma^\mu_\mu \Gamma^\nu_\lambda \phi$ $- \Gamma^\mu_\lambda \Gamma^\nu_\nu \phi$

Exercise: Check that $T^\mu_\lambda \phi$ is a tensor.

Properties of Covariant Derivatives

1) $(\alpha A^\mu + \beta B^\mu)_\phi = \alpha A^\mu_\phi + \beta B^\mu_\phi$ Linearity

2) $(A^\mu B^\nu)_\phi = A^\mu_\phi B^\nu + A^\mu_\nu B^\phi$ Leibniz rule

3) $T^\mu_\lambda \phi = \frac{\partial}{\partial x^\rho} T^\mu_\rho \phi + \Gamma^\mu_\nu \Gamma^\nu_\lambda \phi$ Derivative of covariantly contravariant works as if contracted indicies were covariant.

Covariant Differentiation of the Metric

$\Gamma^\nu_\mu_\lambda \phi = \frac{\partial g_{\mu\lambda}}{\partial x^\nu} - \Gamma^\rho_\mu g_{\nu\lambda} - \Gamma^\rho_\lambda g_{\nu\mu}$

$= 0$ (very definition of $\Gamma^\mu_\nu$ in terms of $g_{\mu\nu}$)

We can also show this by considering a locally inertial coordinate system, in which $\frac{\partial g_{\mu\nu}}{\partial x^\rho} = 0$, $\Gamma^\rho_\mu = 0$ at some point $P$. But $\Gamma^\nu_\mu_\lambda$ is a tensor, so in a general coordinate system, $\Gamma^\nu_\mu_\lambda$ remains zero.
Similarly, $g^{\mu \nu} \partial_\nu = 0$

$\delta V^\mu = \mathbf{0}$

* Covariant Differentiation Commutes of Raising Lowering Indexes

\[(g^{\mu \nu} V^\lambda)_\mu = g^{\mu \nu} (\partial_\nu V^\lambda) + g^{\mu \nu} V^\lambda_{\nu} = g^{\mu \nu} V^\lambda_{\nu}\]

Special Cases of Covariant Differentiation

Covariant Derivative of a Scalar $S$:

\[S_{\mu \nu} = \frac{\partial S}{\partial x^\mu} \]

Covariant Curl: Recall $V_{\mu \nu} = \frac{\partial V^\lambda}{\partial x^\mu} - \Gamma^\lambda_\mu \nu V^\lambda$

\[\text{curl: } V_{\mu \nu} - V_{\nu \mu} = \frac{\partial V^\lambda}{\partial x^\mu} - \frac{\partial V^\lambda}{\partial x^\nu} = \text{ordinary curl.}\]

The covariant divergence of a covariant vector can be written in terms of $g = \det(g_{\mu \nu})$ using the following identity:

\[\text{Tr}\left\{M^{-1}(x) \frac{\partial M(x)}{\partial x^\mu} \right\} = \frac{\partial}{\partial x^\mu} \ln \det M(x)\]

Proof: If $x^\lambda \rightarrow x^\lambda + \delta x^\lambda$, then

\[\delta \ln \det M = \ln \det (M + \delta M) - \ln \det M\]
\[ \Delta \ln \det M = \ln \left( \frac{\det (M + \Delta M)}{\det M} \right) \]
\[ = \ln \det (M^{-1}(M + \Delta M)) \]
\[ = \ln \det (I + M^{-1}\Delta M) \approx \ln (1 + \text{Tr} M^{-1}\Delta M) \]
\[ \approx \text{Tr} M^{-1}\Delta M \]

\[ \lim_{\Delta x \to 0} \frac{\Delta \ln \det M}{\Delta x^2} = \frac{\partial}{\partial x^a} \ln \det M \]
\[ = \text{Tr} \left( M^{-1} \frac{\partial M}{\partial x^a} \right) \]

with \( M = g_{\mu \nu} \),
\[ g^{\mu \nu} \frac{\partial}{\partial x^\lambda} g_{\mu \nu} = \frac{\partial}{\partial x^\lambda} \ln g = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\lambda} \sqrt{g} \]

Then \( \Gamma^\mu_{\lambda \nu} = \frac{1}{2} g^{\mu \rho} \left\{ \partial_\nu g_{\rho \lambda} + \partial_\lambda g_{\rho \nu} - \partial_\rho g_{\lambda \nu} \right\} \]
\[ \Gamma^\mu_{\lambda \nu} = \frac{1}{2} g^{\mu \rho} \partial_\nu g_{\rho \lambda} \]

\[ \Gamma^\mu_{\lambda \nu} = \frac{1}{\sqrt{g}} \partial_\nu \sqrt{g} \]
\[ \Rightarrow \nabla^\mu \Gamma_{\lambda \nu} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\mu} \sqrt{g} \nabla^\mu \]
Example: Divergence in 2D Polar Coordinates

\( \partial^m V^m \) is not a scalar under general coordinate transformations. \( \partial^m V^m \) is a scalar—it takes the same value in any coordinate system.

Polar Coordinates:

\[
\begin{align*}
  x &= r \cos \theta \\
  y &= r \sin \theta \\
  \theta &= \frac{x}{\sqrt{x^2 + y^2}} \\
  r &= \sqrt{x^2 + y^2} \\
  \end{align*}
\]

\[
V^m = \frac{\partial}{\partial x^m} V^0 \]

Let \((x, y)\) be the unprimed coords, \((r, \theta)\) the primed coords.

\[
V^r = \frac{\partial}{\partial x^r} V^0 + \frac{\partial}{\partial y^r} V^y = \cos \theta \, V^x + \sin \theta \, V^y
\]

\[
V^\theta = \frac{\partial}{\partial x^\theta} V^0 + \frac{\partial}{\partial y^\theta} V^\theta = -\frac{\sin \theta}{r} \, V^x + \frac{\cos \theta}{r} \, V^y
\]

\[
\partial^m V^m = \frac{1}{\sqrt{g}} \partial^m (\sqrt{g} \, V^m) = \frac{\partial}{\partial x^r} V^0 + \frac{\partial}{\partial y^r} V^y = \nabla \cdot \vec{V}
\]

\[
ds^2 = dr^2 + r^2 \, d\theta^2 \Rightarrow g = r^2 \quad (\det g_{\mu \nu})
\]

\[
\partial^m V^m = \frac{1}{r} \left[ \frac{\partial}{\partial r} (r \, V^r) + r \, \frac{\partial}{\partial \theta} V^\theta \right]
\]

In terms of orthonormal basis vectors \( \hat{e}_r, \hat{e}_\theta \), the vector

\[
\vec{V} = \sqrt{g_{rr}} \, r \hat{e}_r + \sqrt{g_{\theta \theta}} \, V^\theta \hat{e}_\theta
\]

\[
V^r = V^r \\
V^\theta = r \, V^\theta
\]

We recover the usual expression for the divergence in polar coordinates:

\[
\nabla \cdot \vec{V} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 V^r) + \frac{1}{r} \frac{\partial}{\partial \theta} V^\theta
\]
Covariant Laplacian / D'Alembertian

If $\phi(x)$ is a scalar, $\phi_{,\mu} = (g_{\mu\nu} \phi_{,\nu})_{,\mu} = (g_{\mu\nu} \partial_{\nu} \phi)_{,\mu} = \frac{1}{\sqrt{g}} \partial_{\mu} (\sqrt{g} g^{\mu\nu} \partial_{\nu} \phi)$

In flat space these formulas allow us to compute the gradient, divergence, and curl in arbitrary coordinates.

Example: Laplacian in spherical coordinates

\[ ds^2 = dr^2 + r^2 d\Omega^2 + r^2 \sin^2 \theta d\psi^2 \]

\[ g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \sin^2 \theta \end{pmatrix} \quad g^{\mu\nu} = \begin{pmatrix} 1/r^2 & 0 \\ 0 & 1/(r^2 \sin^2 \theta) \end{pmatrix} \]

\[ g = det g_{\mu\nu} = r^4 \sin^2 \theta \]

\[ \nabla^2 \phi = \frac{1}{\sqrt{g}} \partial_{\mu} (\sqrt{g} g^{\mu\nu} \partial_{\nu} \phi) \]

\[ = \frac{1}{r^2 \sin^2 \theta} \left\{ \frac{\partial}{\partial r} \left( r^2 \sin^2 \theta \frac{\partial \phi}{\partial r} \right) + \frac{\partial}{\partial \theta} \left( r^2 \sin \theta \cdot \frac{1}{r} \frac{\partial \phi}{\partial \theta} \right) + \frac{\partial}{\partial \psi} \left( r^2 \sin^2 \theta \cdot \frac{1}{r^2 \sin^2 \theta} \frac{\partial \phi}{\partial \psi} \right) \right\} \]

\[ = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \phi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \phi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \phi}{\partial \psi^2} \]