Let x, y and z be our usual Cartesian coordinates and consider the general change of coordinates

$$\begin{aligned} x &= x(q_1, q_2, q_3) \\ y &= y(q_1, q_2, q_3) \\ z &= z(q_1, q_2, q_3) \end{aligned} \tag{1}$$

or, more succinctly, let $x = x_1$, $y = x_2$, $z = x_3$ so

$$x_i = x_i(q_1, q_1, q_3), \quad i = 1, 2, 3. \tag{2}$$

The position vector $\boldsymbol{r}$ can be expressed in terms of the q_j's through the Cartesian expression:

$$\boldsymbol{r} = \hat{\boldsymbol{x}}\, x(q_1, q_2, q_3) + \hat{\boldsymbol{y}}\, y(q_1, q_2, q_3) + \hat{\boldsymbol{z}}\, z(q_1, q_2, q_3) = \sum_{i=1}^{3} \hat{\boldsymbol{x}}_i\, x_i(q_1, q_2, q_3). \tag{3}$$

Fixing q_2, q_3 and varying q_1, the vector function $\boldsymbol{r} = \boldsymbol{r}(q_1, q_2, q_3)$ describes a curve. The derivative $\partial \boldsymbol{r}/\partial q_1$ is tangent to that curve. Likewise, $\boldsymbol{r} = \boldsymbol{r}(q_1, q_2, q_3)$ with q_1, q_3 fixed and varying q_2 gives another family of curves, one curve for each value of q_1, q_3, with tangents $\partial \boldsymbol{r}/\partial q_2$. Finally, $\boldsymbol{r} = \boldsymbol{r}(q_1, q_2, q_3)$ with q_1, q_2 fixed and varying q_3 is the third family of curves with tangents $\partial \boldsymbol{r}/\partial q_3$. On the other hand, fixing q_1 and varying q_2, q_3, the vector function $\boldsymbol{r} = \boldsymbol{r}(q_1, q_2, q_3)$ now describes a family of surfaces, one for each q_1, while q_2 fixed with varying q_1, q_3 is another surface and q_3 fixed, varying q_1, q_2 is the third family of surfaces.

Line Element (Displacement vector) The general line element $d\boldsymbol{r}$ in curvilinear coordinates is given by (chain rule):

$$d\boldsymbol{r} = \frac{\partial \boldsymbol{r}}{\partial q_1} dq_1 + \frac{\partial \boldsymbol{r}}{\partial q_2} dq_2 + \frac{\partial \boldsymbol{r}}{\partial q_3} dq_3 = \sum_{i=1}^{3} \frac{\partial \boldsymbol{r}}{\partial q_i} dq_i. \tag{4}$$

Surface Elements The surface element on the surface $q_3 = \textit{constant}$, for example, is given by

$$d\boldsymbol{A}_3 = \left(\frac{\partial \boldsymbol{r}}{\partial q_1} \times \frac{\partial \boldsymbol{r}}{\partial q_2} \right) dq_1 dq_2, \tag{5}$$

and likewise for other area elements. An orientation is built into the order of the coordinates.

Volume Element The volume element is given by the mixed (i.e. triple scalar) product

$$dV = \left(\frac{\partial \boldsymbol{r}}{\partial q_1} \times \frac{\partial \boldsymbol{r}}{\partial q_2} \right) \cdot \frac{\partial \boldsymbol{r}}{\partial q_3}\, dq_1 dq_2 dq_3. \tag{6}$$

This volume element is positive if q_1, q_2, q_3 correspond to a right-handed frame, negative otherwise. This important formula, which corresponds to change of variables in multiple integrals, can be written in the form of a determinant – the *Jacobian determinant* or simply *Jacobian*, often denoted J– which generalizes directly to any dimensions

$$dV = \begin{vmatrix} \partial x/\partial q_1 & \partial x/\partial q_2 & \partial x/\partial q_3 \\ \partial y/\partial q_1 & \partial y/\partial q_2 & \partial y/\partial q_3 \\ \partial z/\partial q_1 & \partial z/\partial q_2 & \partial z/\partial q_3 \end{vmatrix} dq_1 dq_2 dq_3 = det \left(\frac{\partial x_i}{\partial q_j} \right) dq_1 dq_2 dq_3. \tag{7}$$

Orthogonal Coordinates

The vectors $\partial \boldsymbol{r}/\partial q_i$ are key to the coordinates. They provide a natural vector basis. The coordinates are said to be orthogonal if these tangent vectors are orthogonal to each other. It is useful to define the unit vector in the q_i coordinate direction by

$$\frac{\partial \boldsymbol{r}}{\partial q_i} = h_i \hat{\boldsymbol{q}}_i \tag{8}$$

where h_i is therefore

$$h_i = ||\frac{\partial \boldsymbol{r}}{\partial q_i}|| = \sqrt{\sum_{k=1}^{3} \left(\frac{\partial x_k}{\partial q_i}\right)^2}. \tag{9}$$

These h_i's are called the *scale* or *metric* factors. The distance traveled in x-space when changing q_i by dq_i, keeping the other q's fixed, is $ds_i = h_i dq_i$.
In curvilinear coordinates, the unit vectors $\hat{\boldsymbol{q}}_i$ depend on the coordinates. We need to know their derivatives with respect to the q_j, $\partial \hat{\boldsymbol{q}}_i/\partial q_j$, for various operations but in particular to determine the appropriate curvilinear expressions for gradient, divergence and curl. These derivative are perpendicular to $\hat{\boldsymbol{q}}_i$ as, by definition, $\hat{\boldsymbol{q}}_i \cdot \hat{\boldsymbol{q}}_i = 1$. Differentiating that expression with respect to q_j yields

$$\hat{\boldsymbol{q}}_i \cdot \frac{\partial \hat{\boldsymbol{q}}_i}{\partial q_j} = 0. \tag{10}$$

The rate of change of a unit vector is always perpendicular to the unit vector.
Orthogonal coordinates are such that $\hat{\boldsymbol{q}}_i \cdot \hat{\boldsymbol{q}}_j = \delta_{ij}$, $\forall i, j$. We'll now derive compact expressions for $\partial \hat{\boldsymbol{q}}_i/\partial q_j$ in terms of the $\hat{\boldsymbol{q}}_i$'s and the h_i's, in the case of orthogonal coordinates.
The key relationship arises from the equality of mixed partials

$$\frac{\partial^2 \boldsymbol{r}}{\partial q_i \partial q_j} = \frac{\partial^2 \boldsymbol{r}}{\partial q_j \partial q_i} \tag{11}$$

if these 2nd derivatives exist and are continuous, which we assume. Expanding these derivatives in terms of $\hat{\boldsymbol{q}}_i$ and $\hat{\boldsymbol{q}}_j$ defined as in (8) gives

$$\hat{\boldsymbol{q}}_i \frac{\partial h_i}{\partial q_j} + h_i \frac{\partial \hat{\boldsymbol{q}}_i}{\partial q_j} = \hat{\boldsymbol{q}}_j \frac{\partial h_j}{\partial q_i} + h_j \frac{\partial \hat{\boldsymbol{q}}_j}{\partial q_i}. \tag{12}$$

Let

$$\frac{\partial \hat{\boldsymbol{q}}_i}{\partial q_j} = \sum_{k=1}^{3} C_{ijk}\, \hat{\boldsymbol{q}}_k, \qquad C_{ijk} = \hat{\boldsymbol{q}}_k \cdot \frac{\partial \hat{\boldsymbol{q}}_i}{\partial q_j}\ . \tag{13}$$

Then (10) implies that

$$C_{iji} = 0, \qquad \forall i, j \tag{14}$$

and projecting (12) onto $\hat{\boldsymbol{q}}_j$, with i, j distinct, gives

$$C_{ijj} = \frac{1}{h_i} \frac{\partial h_j}{\partial q_i}, \qquad \forall i \neq j. \tag{15}$$

as $\hat{\boldsymbol{q}}_j$ is orthogonal to $\partial \hat{\boldsymbol{q}}_j/\partial q_i$. Projecting (12) onto $\hat{\boldsymbol{q}}_k$ now, with $i \neq k$, $j \neq k$, gives

$$h_i C_{ijk} = h_j C_{jik}, \qquad \forall i \neq k, j \neq k. \tag{16}$$

Finally, orthogonality of the unit vectors $\hat{\boldsymbol{q}}_k \cdot \hat{\boldsymbol{q}}_i = \delta_{ik}$, implies $\partial(\hat{\boldsymbol{q}}_k \cdot \hat{\boldsymbol{q}}_i)/\partial q_j = 0$. Expanding this out and using the definition of C_{ijk} yields

$$C_{ijk} = -C_{kji}, \qquad \forall i \neq k. \tag{17}$$

The four relationships (14), (15), (16) and (17) determine all the C_{ijk}'s. There are 27 C_{ijk} coefficients. Eqn. (14) specifies 9 of them and (15) specifies another 6. The remaining 12 coefficients are determined by the 3 equations given by (16) and the 9 equations (17). Indeed (15) and (17) give

$$C_{jji} = -C_{ijj} = -\frac{1}{h_i}\frac{\partial h_j}{\partial q_i}, \qquad \forall i \neq j, \tag{18}$$

so all coefficients with at least one repeated index, i.e. C_{iji}, C_{ijj} and C_{iik} are known. When i, j, k are all distincts then (16) and (17) give

$$\begin{aligned} h_i C_{ijk} &= h_j C_{jik} = -h_j C_{kij}, \\ h_j C_{jki} &= h_h C_{kji} = -h_k C_{ijk}, \\ h_k C_{kij} &= h_i C_{ikj} = -h_i C_{jki}. \end{aligned} \tag{19}$$

The first and last column of these equations give the 3-by-3 system

$$\begin{array}{ccccccl} h_i C_{ijk} & + & & & h_j C_{kij} & = 0, \\ h_k C_{ijk} & + & h_j C_{jki} & + & & = 0, \\ & & h_i C_{jki} & + & h_k C_{kij} & = 0, \end{array} \tag{20}$$

whose unique solution is $C_{ijk} = C_{jki} = C_{kij} = 0$. Hence all coefficients with no repeated indices vanish. The final results are

$$\boxed{\begin{aligned} \frac{\partial \hat{\boldsymbol{q}}_i}{\partial q_j} &= \hat{\boldsymbol{q}}_j \frac{1}{h_i}\frac{\partial h_j}{\partial q_i}, \qquad \forall i \neq j \qquad (21) \\ \frac{\partial \hat{\boldsymbol{q}}_i}{\partial q_i} &= -\hat{\boldsymbol{q}}_j \frac{1}{h_j}\frac{\partial h_i}{\partial q_j} - \hat{\boldsymbol{q}}_k \frac{1}{h_k}\frac{\partial h_i}{\partial q_k}, \quad \forall i, j, k \quad \text{distinct.} \qquad (22) \end{aligned}}$$

For orthogonal coordinates, the surface and volume elements can be expressed in terms of the h_i's as well. For instance, the surface element for the $q_3 = const.$ surface is

$$d\boldsymbol{A}_3 = \hat{\boldsymbol{q}}_3 \, h_1 h_2 \, dq_1 dq_2, \tag{23}$$

and the volume element

$$dV = h_1 h_2 h_3 \, dq_1 dq_2 dq_3, \tag{24}$$

assuming that q_1, q_2, q_3 is right-handed.
ADD: grad, div, curl... so the h's determine everything when the q coords are orthogonal. Moreover they often have simple expressions in terms of the q's. Let's look at some important examples.

Cylindrical coordinates

$$x = \rho \cos\varphi, \quad y = \rho \sin\varphi, \quad z = z. \tag{25}$$

Spherical coordinates

$$x = r \sin\theta \cos\varphi, \quad y = r \sin\theta \sin\varphi, \quad z = r \cos\theta. \tag{26}$$

Elliptical coordinates

$$x = \alpha \cosh u \cos v, \quad y = \alpha \sinh u \sin v, \quad z = z \tag{27}$$

Non-orthogonal Curvilinear Coordinates

When the coordinates are not orthogonal, the length of the natural basis vectors $h_i = ||\partial \boldsymbol{r}/\partial q_i||$ do not fully determine the geometry. We need to know all the lengths and all the angles between the basis vectors, i.e. all the dot-products

$$g_{ij} \equiv \frac{\partial \boldsymbol{r}}{\partial q_i} \cdot \frac{\partial \boldsymbol{r}}{\partial q_j} = \sum_{l=1}^{3} \frac{\partial x_l}{\partial q_i} \frac{\partial x_l}{\partial q_j}. \tag{28}$$

The g_{ij}'s are the *metric* coefficients. Note that $g_{ii} = h_i^2$.
But there is another important set of basis vectors when the coordinates are non-orthogonal: the $\boldsymbol{\nabla} q_i$ vectors, $i = 1, 2, 3$. The $\boldsymbol{\nabla} q_i$ vector is perpendicular to the $q_i =$ *constant* surface. Therefore $\boldsymbol{\nabla} q_1$, for instance, is orthogonal both to $\partial \boldsymbol{r}/\partial q_2$ and $\partial \boldsymbol{r}/\partial q_3$ which are both tangent to the $q_1 =$ *constant* surface, but $\boldsymbol{\nabla} q_i$ is not parallel to $\partial \boldsymbol{r}/\partial q_i$ unless the coordinates are orthogonal. However, their dot product is always 1. Indeed, by the chain rule, we have

$$\frac{\partial \boldsymbol{r}}{\partial q_i} \cdot \boldsymbol{\nabla} q_j = \sum_{l=1}^{3} \frac{\partial x_l}{\partial q_i} \frac{\partial q_j}{\partial x_l} = \frac{\partial q_j}{\partial q_i} = \delta_{ij} \, . \tag{29}$$

When the coordinates are orthogonal, then $\partial \boldsymbol{r}/\partial q_i$ and $\boldsymbol{\nabla} q_i$ are parallel and, moreover, their magnitudes are inverse of one another from (29), i.e. $\partial \boldsymbol{r}/\partial q_i = h_i \hat{\boldsymbol{q}}_i$ and $\boldsymbol{\nabla} q_i = h_i^{-1} \hat{\boldsymbol{q}}_i$. Again, in the orthogonal case the h_i's and the $\hat{\boldsymbol{q}}_i$'s determine everything. But in the non-orthogonal case, the algebra is simpler if we stick with the non-normalized basis vectors $\partial \boldsymbol{r}/\partial q_i$ and $\boldsymbol{\nabla} q_i$, $i = 1, 2, 3$. (Actually, it is useful to normalize those basis vectors and define two sets of non-dimensional unit vectors, as the components of vectors in those normalized bases keep their physical units. One speaks of "physical curvilinear coordinates" when the normalized bases are used).

Note that $\boldsymbol{\nabla} q_i$ refers to the gradient of the inverse functions $q_i = q_i(x_1, \ldots, x_n)$. The **implicit function theorem** states that these inverse functions exists provided the Jacobian $J \equiv det(\partial x_i/\partial q_j)$ is not singular. Geometrically, this means that if the vectors $\partial \boldsymbol{r}/\partial q_i$, $i = 1, \ldots, n$ are not coplanar, then the q_i's are good coordinates (at least locally) and each point in q-space corresponds to one and only one point in x-space. The equation (29) specifies that the *Jacobian matrices*, $\partial x_i/\partial q_j$ and $\partial q_i/\partial x_j$ are inverses of one another. This is true for their determinant $J_x = det(\partial x_i/\partial q_j)$ and $J_q = det(\partial q_i/\partial x_j)$ also: $J_x J_q = 1$. The Jacobian determinant (or simply Jacobian) is the amplification factor of a volume element, *i.e.* the volume element in q-space: $dV_q \equiv dq_1 dq_2 dq_3$ becomes the volume $dV_x \equiv J_x dV_q$ in x-space. Vice-versa: the volume element $dV_x \equiv dxdydz$ in x-space becomes the volume $dV_q = J_q dV_x$ in q-space. We're back in q-space so the volume should be the same as originally and that means $J_x J_q = 1$. This relation is useful in practice as it may be easier to compute one Jacobian than the other (example: try the Carnot cycle problem posted in the exercises). Note that the chain rule also gives

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij} = \sum_{l=1}^{3} \frac{\partial x_i}{\partial q_l} \frac{\partial q_l}{\partial x_j} \, . \tag{30}$$

The $\partial \boldsymbol{r}/\partial q_i$ basis (displacement basis) is particularly appropriate for displacement type vectors such as the velocity. Indeed the velocity components which are $\dot{x}_i$ in Cartesian coordinates are simply $\dot{q}_i$ in the q_i coordinates when using the displacement basis as by the chain rule

$$\frac{d\boldsymbol{r}}{dt} = \sum_i \hat{\boldsymbol{x}}_i \dot{x}_i = \sum_j \dot{q}_j \frac{\partial \boldsymbol{r}}{\partial q_j}. \tag{31}$$

On the other hand, the $\boldsymbol{\nabla} q_i$ basis is particularly appropriate for the gradient operator as by the chain rule again

$$\boldsymbol{\nabla} f = \sum_i \hat{\boldsymbol{x}}_i \frac{\partial f}{\partial x_i} = \sum_j \frac{\partial f}{\partial q_j} \boldsymbol{\nabla} q_j, \tag{32}$$

so the components of the gradient are simply $\partial f/\partial q_i$ in the $\boldsymbol{\nabla} q_i$ basis, how convenient!
In general, any vector $\boldsymbol{v}$ can be expanded in terms of either basis:

$$\boldsymbol{v} \equiv \sum_i \hat{\boldsymbol{x}}_i v_i \equiv \sum_j v'_j \frac{\partial \boldsymbol{r}}{\partial q_j} \equiv \sum_j \bar{v}_j \boldsymbol{\nabla} q_j. \tag{33}$$

Using the relationships (29), one obtains the components of one basis by projecting onto the other, *i.e.*

$$v'_i = \boldsymbol{v} \cdot \boldsymbol{\nabla} q_i = \sum_k v_k \frac{\partial q_i}{\partial x_k} = \sum_j \bar{v}_j \boldsymbol{\nabla} q_i \cdot \boldsymbol{\nabla} q_j , \tag{34}$$

and

$$\bar{v}_i = \boldsymbol{v} \cdot \frac{\partial \boldsymbol{r}}{\partial q_i} = \sum_l v_l \frac{\partial x_l}{\partial q_i} = \sum_j v'_j \frac{\partial \boldsymbol{r}}{\partial q_j} \cdot \frac{\partial \boldsymbol{r}}{\partial q_i} . \tag{35}$$

These two types of vector components, v'_i and $\bar{v}_i$, are usually referred to as the *"contravariant"* and *"covariant"* vectors, respectively. They are usually written v^j and v_j, but I won't use that compact but weird notation here. These two sets of components can be directly related to each other through the *metric coefficients*

$$g_{ij} \equiv \frac{\partial \boldsymbol{r}}{\partial q_i} \cdot \frac{\partial \boldsymbol{r}}{\partial q_j} = \sum_l \frac{\partial x_l}{\partial q_i} \frac{\partial x_l}{\partial q_j}, \tag{36}$$

and

$$g^{ij} \equiv \boldsymbol{\nabla} q_i \cdot \boldsymbol{\nabla} q_j = \sum_l \frac{\partial q_i}{\partial x_l} \frac{\partial q_j}{\partial x_l} . \tag{37}$$

Indeed from (34) and (35) we have

$$v'_i = \sum_j g^{ij} \bar{v}_j, \tag{38}$$

and

$$\bar{v}_i = \sum_j g_{ij} v'_j. \tag{39}$$

The metric coefficients g_{ij} and g^{ij} specify all the lengths and angles of the bases $\partial \boldsymbol{r}/\partial q_i$ and $\boldsymbol{\nabla} q_i$, $i = 1, 2, 3$, respectively. In orthogonal coordinates, we only needed the 3 h's, but in general we need the 6 g's. (The matrices g^{ij} and g_{ij} are inverse of one another again).

We're starting to understand why Einstein introduced the *summation convention* where sums over repeated indices are implicit, *e.g.* Einstein simply writes

$$v'_j \frac{\partial x_l}{\partial q_i}\frac{\partial x_l}{\partial q_j}$$

for the double sum

$$\sum_j \sum_l v'_j \frac{\partial x_l}{\partial q_i}\frac{\partial x_l}{\partial q_j}.$$

The repeated indices j and l implicitly imply a sum over all values of those indices in Einstein's convention. It is a very useful compact notation and you don't quite need to be Einstein to use it. However it does require some care and understanding!