3.6.2 General change of coordinates

We have seen that is useful to work in a coordinate system appropriate to the properties and symmetries of the system under consideration, using polar coordinates for analyzing a circular drum, or spherical coordinates in analyzing diffusion within a sphere. For arriving at the form of the equation of motion in these coordinate systems we made use of arguments that we generalize below.

Let us start by considering the simple case of *polar coordinates*, (r, ϕ), in the 2D plane $\mathbb{R}^2$ are defined from Cartesian coordinates, (x, y) with $-\infty \leq x, y \leq \infty$, as

$$r = \sqrt{x^2 + y^2}, \qquad \phi = \begin{cases} \arccos\left(\frac{x}{r}\right) & \text{if } y \geq 0 \text{ and } r \neq 0 \\ -\arccos\left(\frac{x}{r}\right) & \text{if } y < 0 \\ \text{undefined} & \text{if } r = 0 \end{cases}, \tag{3.6.9}$$

and conversely

$$x = r\cos(\phi), \qquad y = r\sin(\phi)\,. \tag{3.6.10}$$

The same physical problem can be described in Cartesian coordinates or polar coordinates as they both span the same space (the 2D plane). If the complex number $z = x + iy$ is constructed from the Cartesian coordinates, then $z = r[\cos(\phi) + i\sin(\phi)] = re^{i\phi}$ and $r = |z|$ and $\phi = \arg(z)$ (defined as the principal branch).

The above equations are an example of a coordinate transformation, or change of variables. From Eq. (3.6.11) infinitesimal changes in the two sets of coordinates are related by

$$dx = dr\cos(\phi) - d\phi\; r\sin(\phi) \quad \text{and} \quad dy = dr\sin(\phi) + d\phi\; r\cos(\phi)\,. \tag{3.6.11}$$

The relation between infinitesimal changes in two coordinate systems can thus be expressed as

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \mathbf{J}(r,\phi)\begin{bmatrix} dr \\ d\phi \end{bmatrix}, \tag{3.6.12}$$

in terms of the so-called *Jacobian* (or Jacobi matrix)

$$\mathbf{J}(r,\phi) = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{bmatrix}. \tag{3.6.13}$$

The inverse Jacobian

$$\mathbf{J}^{-1}(x,y) = \begin{bmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{bmatrix}, \tag{3.6.14}$$

provides the transformation in the reverse direction.

There are many other useful coordinate systems. Some such as spherical or cylindrical coordinates are common enough that they bear remembering, but others are specific to a particular problem. More generally, given a set of coordinates $(x_1, \ldots, x_n)$ on $\mathbb{R}^n$, a set of functions $y_i(x_1, \ldots, x_n)$ for $i = 1, \ldots n$ that map $\mathbb{R}^n \to \mathbb{R}^n$, is a coordinate transformation and changes in the new coordinates y_i are related to changes in the original coordinates through the Jacobian

$$\mathbf{J}(y_1, \ldots y_n) = \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{bmatrix} . \tag{3.6.15}$$

The Jacobian is highly useful in computing derivatives and gradient operators in the new coordinate system:

- The change of variables transforms a function $f(\vec{x})$ in the original coordinates to a function $f(\vec{h})$ in the new set of coordinates. By application of the chain rule, the corresponding two vectors of derivatives are related by

$$\frac{\partial f}{\partial y_\alpha} = \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial y_\alpha}, \Rightarrow \quad \frac{\partial f}{\partial \vec{y}} = \mathbf{J}(\vec{y}) \cdot \frac{\partial f}{\partial \vec{x}} . \tag{3.6.16}$$

- The above can be extended to multiple functions that can be collected together as a vector. For a m-component vector function $\mathbf{f}(\vec{x}) = (f_1(\vec{x}), \ldots, f_m(\vec{x}))$ of an n-component coordinate vector $\vec{x}$, the matrix of derivatives with respect to a new coordinate vector $\vec{y}$ is

$$\frac{\partial \mathbf{f}}{\partial \vec{y}} = \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_n} \end{bmatrix} = \mathbf{J}(\vec{y}) \cdot \frac{\partial \mathbf{f}}{\partial \vec{x}} , \tag{3.6.17}$$

  where $\frac{\partial \mathbf{f}}{\partial \vec{y}}$ and $\frac{\partial \mathbf{f}}{\partial \vec{x}}$ are $m \times n$ matrices.

- If only one of the new set of coordinates is changed, say y_1 by dy_1, from Eq. (3.6.15), the corresponding change in the original Cartesian coordinates is

$$d\vec{r}_1 = \begin{bmatrix} \frac{\partial x_1}{\partial y_1} \\ \vdots \\ \frac{\partial x_n}{\partial y_1} \end{bmatrix} dy_1 \equiv (h_1 \ dy_1)\hat{e}_1 . \tag{3.6.18}$$

  Through the above equation we have defined two important quantities: a unit vector $\hat{e}_1$ pointing along the direction of the change caused by dy_1; and a metric factor h_1 quantifying (as $h_1 \ dy_1$) the magnitude of the change. The complete sets of $\{\hat{e}_\alpha\}$ and $\{h_\alpha\}$ quantify variations along all directions.

- The absolute value of the Jacobian determinant represents the change in volume element when making a change of variables while evaluating a integral of a function over a region. To accommodate for the change of coordinates $|\det \mathbf{J}|$ is included as a multiplicative factor within the integral. This is because the n-dimensional dV element is in general a parallelepiped in a new coordinate system made up of the infinitesimal vectors $\{(h_\alpha \hat{e}_\alpha)dy_\alpha\}$. The volume of this infinitesimal parallelepiped is obtained as the determinant of its edge vectors. For the case of Cartesian to polar coordinates, the determinant of Eq. (3.6.13) gives

$$dx\,dy = |\det \mathbf{J}| dr\,d\phi = r(\cos^2\phi + \sin^2\psi) dr\,d\phi = r\,dr\,d\phi\,, \tag{3.6.19}$$

a result we used previously in Eq. (3.5.1).

- The set of metric factors $\{h_\alpha\}$ and corresponding unit vectors $\{\hat{e}_\alpha\}$ can be used to construct the gradient operation in the new coordinates as

$$\vec{\nabla} f = \sum_\alpha \frac{1}{h_\alpha}\frac{\partial f}{\partial y_\alpha}\hat{e}_\alpha\,. \tag{3.6.20}$$

Clearly, along with Eq. (3.6.18), this leads to the expected $df = \frac{\partial f}{\partial y_\alpha} dy_\alpha$. For polar coordinates, Eq. (3.6.13) indicates that

$$d\vec{r} = dr\ \hat{r} + d\phi\ r\hat{\phi}\,, \quad \text{and} \quad \vec{\nabla} f = \frac{\partial f}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial f}{\partial \phi}\hat{\phi}\,. \tag{3.6.21}$$

- To construct the divergence of a vector field $\vec{v}$, we employ the infinitesimal variant of the divergence theorem, Eq. (3.6.8), as in Eq. (3.6.4)[7]

$$\mathrm{div}\vec{v}\ dV = \sum_\alpha dy_\alpha \frac{\partial}{\partial y_\alpha}(dS_\alpha v_\alpha)\,. \tag{3.6.22}$$

The volume and surface elements in the above formula are related by

$$dV = |\det \mathbf{J}| \prod_\alpha dy_\alpha = (h_1\ dy_1) dS_1 = (h_1\ dy_1) dS_1 = \cdots\,. \tag{3.6.23}$$

Dividing by dV thus leads to

$$\mathrm{div}\vec{v} = \vec{\nabla}\cdot\vec{v} = \frac{1}{|\det \mathbf{J}|}\sum_\alpha \frac{\partial}{\partial y_\alpha}\left(\frac{|\det \mathbf{J}|}{h_\alpha} v_\alpha\right)\,. \tag{3.6.24}$$

- Finally, the Laplacian of a scalar field f in general (curvilinear) coordinates is obtained as the divergence of the gradient of f, and given by

$$\nabla^2 f = \frac{1}{|\det \mathbf{J}|}\sum_\alpha \frac{\partial}{\partial y_\alpha}\left(\frac{|\det \mathbf{J}|}{h_\alpha^2}\frac{\partial f}{\partial y_\alpha}\right)\,. \tag{3.6.25}$$

In the next sections we shall use these general formulae to re-derive the explicit forms of these operations in two commonly encountered coordinate systems.

---

[7] The explicit inclusion of the summation symbol $\sum_\alpha$ indicates that the summation convention is no longer used.