Coordinate-Free Representations
of
Vector Derivative Operators
with an
Illustrative Example

The notation used herein will be an amalgam of several sources, made consistent.

- The expressions for the coordinate-free forms of the vector operators is from *Electromagnetic Fields and Waves* by Vladimir Rojansky. If you check the literature, you will find a plethora of volumes with the same title. I have the 1971 edition (this is perhaps the only edition), and I'm citing the forms given on Page 446, with more conventional notation.
- The use of the differential line elements is from *Mechanics* by Fowles; someone boosted my copy a while ago, so I can't give more information, other than what I'm using is from the Appendices. Again, a common title.
- The choice of coordinates and vector notation is that as given in the regular curriculum 8.022 notes **Vector Identities**, linked from the **8.022-ESG page**. Not surprisingly, this is the same notation used by Purcell.
- The Illustrative Example will be spherical polar coordinates, using the "Physics" convention, basically $z = r\cos\theta$. If "Math" coordinates are preferred, with $z = r\cos\phi$, no objections will be raised. You know what needs to be done.

The coordinate-free definitions are:

$$\operatorname{grad} f = \boldsymbol{\nabla} f = \lim_{v\to 0}\frac{1}{v}\oint\!\!\oint_{(S)} \mathbf{n}\, f\, da$$

$$\operatorname{div}\mathbf{F} = \boldsymbol{\nabla}\cdot\mathbf{F} = \lim_{v\to 0}\frac{1}{v}\oint\!\!\oint_{(S)} \mathbf{n}\cdot\mathbf{F}\, da$$

$$\operatorname{curl}\mathbf{F} = \boldsymbol{\nabla}\times\mathbf{F} = \lim_{v\to 0}\frac{1}{v}\oint\!\!\oint_{(S)} \mathbf{n}\times\mathbf{F}\, da$$

In each of the above, $\oint\!\!\oint_{(S)}$ represents a surface integral over a closed surface (S), and $\mathbf{n}$ is the unit outward normal vector. Note the similarity in the expressions.

In fact, you might at some point hear a mathematician say something to the effect that "these are all just exterior derivatives." You might also recoginze that if we express these expressions in words, they would be something similar to "evaluate the function on the boundary of a region and then divide by the size of the region," which is more or less the definition of the derivative of a single-variable derivative.

These definitions may or may not be useful for specific applications. From the first of these expressions, however, we can infer Gauss's Law (the Divergence Theorem),

$$\oiint_{(S)} \mathbf{n} \cdot \mathbf{F}\, da = \iiint_V \boldsymbol{\nabla} \cdot \mathbf{F}\, dv$$

where V is the volume bounded by the closed surface (S).

For use in particular coordinate systems, explicit expressions for the differential surfaces areas and volumes would need to be found. It will, however, be advantageous to find an expression in an *arbitrary* orthonormal coordinate system, one which we can then use in a chosen practical system.

What we do is consider a coordinate system with three coordinates x_1, x_2 and x_3, with associated unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$, assumed mutually orthogonal, with $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$. In general, if we can express the position vector $\mathbf{r}\,(x_1,\, x_2,\, x_3)$ as a vector function of the three coordinates,

$$\mathbf{e}_i = \frac{\frac{\partial}{\partial x_i}\mathbf{r}}{h_i}, \quad \text{where} \quad h_i = \left| \frac{\partial}{\partial x_i}\mathbf{r} \right|.$$

From the above, we can define a differential line element $d\mathbf{r}$ by

$$d\mathbf{r} = h_1\,(x_1,\, x_2,\, x_3)\; \mathbf{e}_1\, dx_1 + h_2\,(x_1,\, x_2,\, x_3)\; \mathbf{e}_2\, dx_2 + h_3\,(x_1,\, x_2,\, x_3)\; \mathbf{e}_3\, dx_3$$

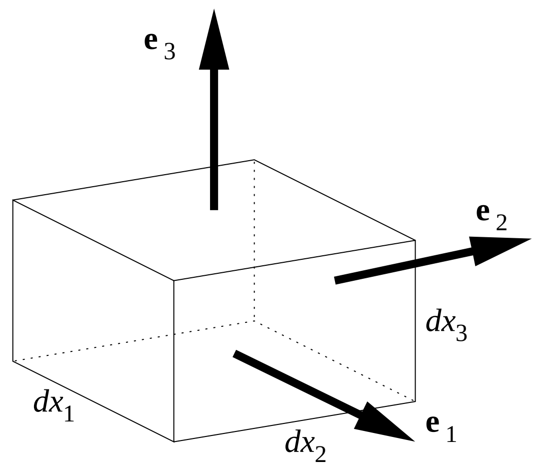
and from this a differential volume element

$$dv = h_1\, h_2\; h_3\, dx_1\, dx_2\, dx_3.$$

To look ahead to the specific cases, in spherical coordinates we have

$$d\mathbf{r} = \;\hat{\mathbf{r}}\, dr + r\, \hat{\boldsymbol{\theta}}\, d\theta + r\, \sin\theta\, \hat{\boldsymbol{\phi}}\, d\phi, \qquad dv = r^2\, \sin\theta\, dr\, d\theta\, d\phi.$$

For the general orthonormal case, the volume v (or dv) will be a rectangular parallelepiped in the limit as $dx_i \to 0$; the normals to the six sides are $\pm\mathbf{e}_i$. A crude figure is on the next page. Hey, Brunelleschi I'm not.



An advantage of using the general form is that we can consider a pair of opposite sides, and then extend the result by permuting the coordinates. So, let's consider the two faces with normals $\mathbf{n} = \pm\mathbf{e}_1$. For the face with normal $\mathbf{n} = -\mathbf{e}_1$, the area of the face will be $h_2(x_1,\, x_2,\, x_3)\, dx_2 \times h_3(x_1,\, x_2,\, x_3)\, dx_3$. For the face with normal $\mathbf{n} = +\mathbf{e}_1$, the area of the face will be $h_2\left(x_1 + dx_1,\, x_2,\, x_3\right)\, dx_2 \times h_3\left(x_1 + dx_1,\, x_2,\, x_3\right)\, dx_3$.

Look at that last line above closely, as this is the crucial part of the derivations, and the advantage. In general, the values of h_i will change as the coordinates vary, and in each of the integrals in the coordinate-free definitions, what's wanted is the *product* of the differential area and the appropriate part of the dependent scalar or vector field.

It turns out that the divergence, being a scalar, is the most convenient of the operators to express in terms of general coordinates. The divergence, as given above, is

$$\operatorname{div}\mathbf{F} = \boldsymbol{\nabla}\cdot\mathbf{F} = \lim_{v\to 0}\frac{1}{v}\oiint_{(S)} \mathbf{n}\cdot\mathbf{F}\, da,$$

where we'll introduce the notation $\mathbf{F} = F_1\,\mathbf{e}_1 + F_2\,\mathbf{e}_2 + F_3\,\mathbf{e}_3$ (this is *not* Purcell's notation!), where the F_i are implied functions of the x_i.

The contribution to the surface integral over the two surfaces perpendicular to $\mathbf{e}_1$ will involve only the component F_1 and how the product $F_1\,h_2\,h_3$ changes as x_1 changes. The contribution to the flux integral is

$$\begin{bmatrix} F_1\left(x_1+dx_1,\,x_2,\,x_3\right)\,h_2\left(x_1+dx_1,\,x_2,\,x_3\right)\,dx_2\,h_3\left(x_1+dx_1,\,x_2,\,x_3\right)\,dx_3 \\ -\,F_1\left(x_1,\,x_2,\,x_3\right)\,dx_2\,h_2\left(x_1,\,x_2,\,x_3\right)\,dx_3\,h_3\left(x_1,\,x_2,\,x_3\right) \end{bmatrix}$$

$$=\frac{\partial}{\partial x_1}\left(F_1\,h_2\,h_3\right)\,dx_1\,dx_2\,dx_3,$$

so that the divergence is

$$\operatorname{div}\mathbf{F}=\boldsymbol{\nabla}\cdot\mathbf{F}=\frac{1}{h_1\,h_2\,h_3}\left[\frac{\partial}{\partial x_1}\left(F_1\,h_2\,h_3\right)+\frac{\partial}{\partial x_2}\left(F_2\,h_3\,h_1\right)+\frac{\partial}{\partial x_3}\left(F_3\,h_1\,h_2\right)\right].$$

In practice, the gradient is often the easiest to calculate and the easiest to motivate in terms of the change of a scalar function of several variables in terms of partial derivatives. The coordinate-free definition,

$$\operatorname{grad}f=\boldsymbol{\nabla}f=\lim_{v\to 0}\frac{1}{v}\oint\!\!\oint_{(S)}\mathbf{n}\,f\,da,$$

while arguably elegant, would give

$$\operatorname{grad}f=\boldsymbol{\nabla}f=\frac{1}{h_1\,h_2\,h_3}\left[\frac{\partial}{\partial x_1}\left(\mathbf{e}_1\,f\,h_2\,h_3\right)+\frac{\partial}{\partial x_2}\left(\mathbf{e}_2\,f\,h_3\,h_1\right)+\frac{\partial}{\partial x_3}\left(\mathbf{e}_3\,f\,h_1\,h_2\right)\right].$$

For coordinate systems other than Cartesian, the unit vectors will themselves be functions of the coordinates, and the above, while true, is of little practical use. With the unit vectors defined in terms of derivatives, the derivatives of unit vectors will in general involve second derivatives of $\mathbf{r}\left(x_1,\,x_2,\,x_3\right)$.

The variation of the coordinate vectors can be taken into account by noticing that $\operatorname{grad}(1)=\boldsymbol{\nabla}1=\mathbf{0}$. That is, with the above definition of the gradient,

$$\operatorname{grad}(1)=\boldsymbol{\nabla}\,1=\lim_{v\to 0}\frac{1}{v}\oint\!\!\oint_{(S)}\mathbf{n}\,da=\mathbf{0}.$$

So far, we haven't shown that the gradient as defined above is the derivative of anything, so we have to show that the claim $\boldsymbol{\nabla}1=\mathbf{0}$ is indeed valid.

What we do is note that $\mathbf{n}\,da=d\mathbf{a}$ and that

$$\oint\!\!\oint_{(S)}d\mathbf{a}=\mathbf{0}$$

for a closed surface. This is necessarily true from differential geometry considerations, but may not be intuitive. A two-dimensional anology would be that $\oint d\mathbf{r} = \mathbf{0}$; by integrating the differential path differences along a closed path, you end up where you started, for a net *vector* displacement of zero. Similarly, integrating the differential area *vectors* over a closed surface gives a *vector* area of zero.

Perhaps more directly, it is immediately clear that $\boldsymbol{\nabla} 1 = \mathbf{0}$ in Cartesian coordinates. Since our definition is coordinate-free, $\boldsymbol{\nabla} 1 = \mathbf{0}$ in *all* coordinate systems.

To find the gradient in a usable form, consider, as for the divergence, the contribution to the integral from the surfaces perpendicular to $\mathbf{e}_1$. (For this part of the derivation, the dependence of the quantities on x_2 or x_3 is suppressed, and the product $\mathbf{e}_1 \, h_2 \, h_3$ is denoted as $\mathbf{e}_1'$.) This contribution is

$$
\begin{aligned}
& f\left(x_1 + dx_1\right) \mathbf{e}_1'\left(x_1 + dx_1\right) - f\left(x_1\right) \mathbf{e}_1'\left(x_1\right) \\
&\qquad = f\left(x_1 + dx_1\right) \mathbf{e}_1'\left(x_1 + dx_1\right) - f\left(x_1 + dx_1\right) \mathbf{e}_1'\left(x_1\right) \\
&\qquad + f\left(x_1 + dx_1\right) \mathbf{e}_1'\left(x_1\right) - f\left(x_1\right) \mathbf{e}_1'\left(x_1\right) \\
&= f\left(x_1 + dx_1\right)\left[\mathbf{e}_1'\left(x_1 + dx_1\right) - \mathbf{e}_1'\left(x_1\right)\right] + \left[f\left(x_1 + dx_1\right) - f\left(x_1\right)\right] \mathbf{e}_1'\left(x_1\right)
\end{aligned}
$$

(if this is reminiscent of the product rule, that's good). What we do now is to rewrite the term on the right above as a partial derivative of f with respect to x_1, but leave the term on the left alone, *except* for replacing $f\left(x_1 + dx_1\right)$ with $f\left(x_1\right)$.

When the terms for the other components of the gradient are added and regrouped, the result is

$$
\begin{aligned}
\boldsymbol{\nabla} f = & \frac{1}{h_1 \, h_2 \, h_3}\left[\mathbf{e}_1 \, h_2 \, h_3 \frac{\partial f}{\partial x_1} + \mathbf{e}_2 \, h_3 \, h_1 \frac{\partial f}{\partial x_2} + \mathbf{e}_3 \, h_1 \, h_2 \frac{\partial f}{\partial x_3}\right] \\
& + f \lim_{v \to 0} \frac{1}{h_1 \, h_2 \, h_3 \, dx_1 \, dx_2 \, dx_3}\left[\begin{array}{r}
\left(\mathbf{e}_1'\left(x_1 + dx_1\right) - \mathbf{e}_1'\left(x_1\right)\right) \\
+\left(\mathbf{e}_2'\left(x_2 + dx_2\right) - \mathbf{e}_2'\left(x_2\right)\right) \\
+\left(\mathbf{e}_3'\left(x_3 + dx_3\right) - \mathbf{e}_3'\left(x_3\right)\right)
\end{array}\right].
\end{aligned}
$$

The second term above is the product of f and $\boldsymbol{\nabla} 1$, believed to be $\mathbf{0}$. More specifically, the term in square brackets involving the variations of the $\mathbf{e}_i'$ is the vector area of the boundary of the differential volume; this boundary must be a closed surface, and the vector area must be zero. Therefore, the useful expression for the gradient in general coordinates is

$$
\operatorname{grad} f = \boldsymbol{\nabla} f = \mathbf{e}_1 \frac{1}{h_1} \frac{\partial f}{\partial x_1} + \mathbf{e}_2 \frac{1}{h_2} \frac{\partial f}{\partial x_2} + \mathbf{e}_3 \frac{1}{h_3} \frac{\partial f}{\partial x_3}.
$$

To the surprise of few, the curl, given above as

$$\operatorname{curl}\mathbf{F} = \boldsymbol{\nabla}\times\mathbf{F} = \lim_{v\to 0}\frac{1}{v}\oiint_{(S)} \mathbf{n}\times\mathbf{F}\,da,$$

is the trickiest, but it is hoped that from the above derivations of the gradient and divergence, confidence exists in expectation of a simple form. Of course. Proceding as before, on the faces perpendicular to $\pm\mathbf{e}_1$, the contribution to the integral will be the change in the vector quantity

$$\mathbf{e}_1\times\mathbf{F}\,h_2\,h_3 = \mathbf{e}_1\times\,(F_1\,\mathbf{e}_1 + F_2\,\mathbf{e}_2 + F_3\,\mathbf{e}_3)\,h_2\,h_3 = \left[\mathbf{e}_3\,(F_2) - \mathbf{e}_2\,(F_3)\right]h_2\,h_3.$$

Thus, the contribution to the integral of $\mathbf{n}\times\mathbf{F}$ from these two faces is

$$\left[\frac{\partial}{\partial x_1}\left(\mathbf{e}_3\,F_2\,h_2\,h_3\right) - \frac{\partial}{\partial x_1}\left(\mathbf{e}_2\,F_3\,h_2\,h_3\right)\right].$$

Similar terms are obtained from the other four surfaces, leading to, after minor rearrangement of terms,

$$\left[\begin{array}{l} \left(\dfrac{\partial}{\partial x_2}\left(\mathbf{e}_1\,F_3\,h_1\,h_3\right) - \dfrac{\partial}{\partial x_3}\left(\mathbf{e}_1\,F_2\,h_2\,h_1\right)\right) \\ + \left(\dfrac{\partial}{\partial x_3}\left(\mathbf{e}_2\,F_1\,h_2\,h_1\right) - \dfrac{\partial}{\partial x_1}\left(\mathbf{e}_2\,F_3\,h_3\,h_2\right)\right) \\ + \left(\dfrac{\partial}{\partial x_1}\left(\mathbf{e}_3\,F_2\,h_3\,h_2\right) - \dfrac{\partial}{\partial x_2}\left(\mathbf{e}_3\,F_1\,h_1\,h_3\right)\right) \end{array}\right]$$

The next step involves a good deal of hindsight, in that we know that the curl tells us something about how a vector field component affects a line integral, hence the variation of the product of the component and the line element. In this case, this product is represented by the terms of the form $F_i\,h_i$, so we rewrite the above expression in the large square braces as

$$\left[\begin{array}{l} h_1\,\mathbf{e}_1\left(\dfrac{\partial}{\partial x_2}\left(F_3\,h_3\right) - \dfrac{\partial}{\partial x_3}\left(F_2\,h_2\right)\right) \\ +h_2\,\mathbf{e}_2\left(\dfrac{\partial}{\partial x_3}\left(F_1\,h_1\right) - \dfrac{\partial}{\partial x_1}\left(F_3\,h_3\right)\right) \\ +h_3\,\mathbf{e}_3\left(\dfrac{\partial}{\partial x_1}\left(F_2\,h_2\right) - \dfrac{\partial}{\partial x_2}\left(F_1\,h_1\right)\right) \end{array}\right] + \left[\begin{array}{l} F_1\,h_1\left(\dfrac{\partial}{\partial x_3}\left(h_2\,\mathbf{e}_2\right) - \dfrac{\partial}{\partial x_2}\left(h_3\,\mathbf{e}_3\right)\right) \\ +F_2\,h_2\left(\dfrac{\partial}{\partial x_1}\left(h_3\,\mathbf{e}_3\right) - \dfrac{\partial}{\partial x_3}\left(h_1\,\mathbf{e}_1\right)\right) \\ +F_3\,h_3\left(\dfrac{\partial}{\partial x_2}\left(h_1\,\mathbf{e}_1\right) - \dfrac{\partial}{\partial x_1}\left(h_2\,\mathbf{e}_2\right)\right) \end{array}\right]$$

Consider the terms on the right involving

$$\frac{\partial}{\partial x_i}\left(h_j\,\mathbf{e}_j\right) - \frac{\partial}{\partial x_j}\left(h_i\,\mathbf{e}_i\right);$$

recalling that $h_i\,\mathbf{e}_i = \dfrac{\partial}{\partial x_i}\mathbf{r}$, these differences are seen to vanish, so that

$$\operatorname{curl}\mathbf{F} = \boldsymbol{\nabla}\times\mathbf{F} = \frac{1}{h_1\,h_2\,h_3}\left[\begin{array}{r} h_1\,\mathbf{e}_1\left(\dfrac{\partial}{\partial x_2}(F_3\,h_3) - \dfrac{\partial}{\partial x_3}(F_2\,h_2)\right) \\ +h_2\,\mathbf{e}_2\left(\dfrac{\partial}{\partial x_3}(F_1\,h_1) - \dfrac{\partial}{\partial x_1}(F_3\,h_3)\right) \\ +h_3\,\mathbf{e}_3\left(\dfrac{\partial}{\partial x_1}(F_2\,h_2) - \dfrac{\partial}{\partial x_2}(F_1\,h_1)\right) \end{array}\right].$$

This form is certainly suggestive of a familiar expression for a curl (or cross product) in terms of a determinant,

$$\operatorname{curl}\mathbf{F} = \boldsymbol{\nabla}\times\mathbf{F} = \frac{1}{h_1\,h_2\,h_3}\left|\begin{array}{ccc} h_1\,\mathbf{e}_1 & h_2\,\mathbf{e}_2 & h_3\,\mathbf{e}_3 \\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \\ F_1\,h_1 & F_2\,h_2 & F_3\,h_3 \end{array}\right|.$$

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

The three expressions are repeated here:

$$\operatorname{grad} f = \boldsymbol{\nabla} f = \mathbf{e}_1\frac{1}{h_1}\frac{\partial f}{\partial x_1} + \mathbf{e}_2\frac{1}{h_2}\frac{\partial f}{\partial x_2} + \mathbf{e}_3\frac{1}{h_3}\frac{\partial f}{\partial x_3}.$$

$$\operatorname{div}\mathbf{F} = \boldsymbol{\nabla}\cdot\mathbf{F} = \frac{1}{h_1\,h_2\,h_3}\left[\frac{\partial}{\partial x_1}(F_1\,h_2\,h_3) + \frac{\partial}{\partial x_2}(F_2\,h_3\,h_1) + \frac{\partial}{\partial x_3}(F_3\,h_1\,h_2)\right].$$

$$\operatorname{curl}\mathbf{F} = \boldsymbol{\nabla}\times\mathbf{F} = \frac{1}{h_1\,h_2\,h_3}\left[\begin{array}{r} h_1\,\mathbf{e}_1\left(\dfrac{\partial}{\partial x_2}(F_3\,h_3) - \dfrac{\partial}{\partial x_3}(F_2\,h_2)\right) \\ +h_2\,\mathbf{e}_2\left(\dfrac{\partial}{\partial x_3}(F_1\,h_1) - \dfrac{\partial}{\partial x_1}(F_3\,h_3)\right) \\ +h_3\,\mathbf{e}_3\left(\dfrac{\partial}{\partial x_1}(F_2\,h_2) - \dfrac{\partial}{\partial x_2}(F_1\,h_1)\right) \end{array}\right].$$

For the promised example, let's return to spherical coordinates, with

$$d\mathbf{r} = \hat{\mathbf{r}}\,dr + r\,\hat{\boldsymbol{\theta}}\,d\theta + r\,\sin\theta\,\hat{\boldsymbol{\phi}}\,d\phi, \qquad dv = r^2\,\sin\theta\,dr\,d\theta\,d\phi,$$

$$h_1 = 1, \qquad h_2 = r, \qquad h_3 = r\,\sin\theta,$$

$$\text{and use} \qquad \mathbf{F} = F_r\,\hat{\mathbf{r}} + F_\theta\,\hat{\boldsymbol{\theta}} + F_\phi\,\hat{\boldsymbol{\phi}}.$$

We then have

$$\operatorname{grad} f = \boldsymbol{\nabla}\,f = \hat{\mathbf{r}}\,\frac{\partial f}{\partial r} + \hat{\boldsymbol{\theta}}\,\frac{1}{r}\frac{\partial f}{\partial \theta} + \hat{\boldsymbol{\phi}}\,\frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}.$$

$$\operatorname{div}\mathbf{F} = \boldsymbol{\nabla}\cdot\mathbf{F} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\,F_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,F_\theta\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}F_\phi.$$

The curl may be calculated similarly from the above expression, but will not be reproduced here.

It may be instructive to note that the term discarded in the derivation of the form for the gradient in general coordinates does indeed vanish in spherical coordinates. The discarded term, had the limits $dx_i \to 0$ been taken, would have been of the form

$$\frac{\partial}{\partial x_1}\left(\mathbf{e}_1\,h_2\,h_3\right) + \frac{\partial}{\partial x_2}\left(\mathbf{e}_2\,h_3\,h_1\right) + \frac{\partial}{\partial x_3}\left(\mathbf{e}_3\,h_1\,h_2\right).$$

In spherical coordinates, the needed derivatives of the coordinate vectors are (as can be shown by purely geometric methods)

$$\frac{\partial\,\hat{\mathbf{r}}}{\partial r} = \mathbf{0}, \qquad \frac{\partial\hat{\boldsymbol{\theta}}}{\partial\theta} = -\,\hat{\mathbf{r}}, \qquad \frac{\partial\hat{\boldsymbol{\phi}}}{\partial\phi} = -\left(\hat{\mathbf{r}}\,\sin\theta + \hat{\boldsymbol{\theta}}\cos\theta\right).$$

Thus,

$$\begin{aligned} &\frac{\partial}{\partial r}\left(\hat{\mathbf{r}}\,r^2\,\sin\theta\right) + \frac{\partial}{\partial\theta}\left(\hat{\boldsymbol{\theta}}\,r\,\sin\theta\right) + \frac{\partial}{\partial\phi}\left(\hat{\boldsymbol{\phi}}\,r\right) \\ &= \hat{\mathbf{r}}\,2\,r\,\sin\theta - \hat{\mathbf{r}}\,r\,\sin\theta + \hat{\boldsymbol{\theta}}\,r\,\cos\theta - \hat{\mathbf{r}}\,r\,\sin\theta - \hat{\boldsymbol{\theta}}\,r\,\cos\theta \\ &= \mathbf{0}. \end{aligned}$$